Assessment
1–25

DOI: 10.1177/1073191116645905
asm.sagepub.com
SAGE

Disentangling Components of Flexibility via the Hexaflex Model: Development and Validation of the Multidimensional Psychological Flexibility Inventory (MPFI)

Jaci L. Rolffs[1], Ronald D. Rogge[1], and Kelly G. Wilson[2]

Abstract

The current study developed the 60-item Multidimensional Psychological Flexibility Inventory (MPFI)—a scale assessing the 12 dimensions of the Hexaflex model. We created an exhaustive pool of 554 items including 22 of the most widely used measures from the acceptance and commitment therapy and mindfulness literatures. Exploratory and confirmatory factor analyses were used in combination with item response theory and responsiveness to change analyses in 3,040 online respondents across three studies ($N_{Study\ 1}$ = 372; $N_{Study\ 2}$ = 2,150; $N_{Study\ 3}$ = 518) to create the MPFI. Associations between the MPFI subscales and an array of existing measures supported its convergent and discriminant validities. The MPFI offers acceptance and commitment therapy researchers new tools for elaborating treatment effects.

Keywords

psychological flexibility, Hexaflex model, acceptance and commitment therapy, mindfulness, measure development, item response theory, responsiveness to change

Rooted in relational frame theory (Hayes, Barnes-Holmes, & Roche, 2001), psychological flexibility has been conceptualized as a number of key ways individuals can alter the function of internal experiences by flexibly responding to negative thoughts, feelings, and events, thereby enhancing their well-being. Psychological flexibility helps individuals open up to those experiences, allowing them to be there, while still making behavioral choices in service of the areas of their life that are important to them. Numerous studies have demonstrated that psychological flexibility is linked with better psychological and physical functioning. In fact, the 6-factor theory of psychological flexibility (i.e., the Hexaflex model; Hayes, Strosahl, & Wilson, 1999, 2011) has served as a key element of acceptance and commitment therapy (ACT; Hayes et al., 1999), a treatment that has been shown to be effective across a range of diagnoses (Hayes, Luoma, Bond, Masuda, & Lillis, 2006). Despite the growing body of literature supporting the importance of psychological flexibility, the measurement work on this theory has lagged behind the research examining its practical applications. The most widely used measures of psychological flexibility (the Acceptance and Action Questionnaire, AAQ, Hayes et al., 2004, the Acceptance and Action Questionnaire–II, AAQ-II, Bond et al., 2011, and the Avoidance and Fusion Questionnaire for Youth, AFQ-Y, Greco, Lambert, & Baer, 2008) treat it as a single dimension despite the fact that the theory posits up to 12 distinct dimensions. Work in this area has also developed scales that tap various individual components of the Hexaflex model (e.g., the Mindful Attention and Awareness Scale, MAAS, Brown & Ryan, 2003). However, these scales exist across a diverse array of measures and have yet to be integrated into a comprehensive measure that has a stable 12-factor structure directly mapping onto the Hexaflex model. This has impeded researchers' abilities to examine the individual dimensions of flexibility to determine which ones are most critical to various forms of individual health and well-being, which ones are most directly affected by interventions, and which ones might serve as important moderators of treatments.

The Hexaflex Model of Psychological Flexibility

Psychological Flexibility (Hayes et al., 2011) is composed of 6 distinct components: Acceptance (i.e., willingness to

[1]University of Rochester, Rochester, NY, USA
[2]University of Mississippi, Oxford, MS, USA

Corresponding Author:
Ronald D. Rogge, University of Rochester, 462 Meliora Hall, RC Box 270266, Rochester, NY 14627-0266, USA.
Email: rogge@psych.rochester.edu

contact unwanted experiences fully), Contact with the Present Moment (i.e., being in touch and aware of one's experiences), Self as Context (i.e., keeping perspective of oneself within one's experiences), Defusion (i.e., being able to step back from unwanted experiences without getting stuck in them), Committed Action (i.e., maintaining behaviors that move toward important aspects of life), and Values (i.e., staying connected to the areas of life that are important, giving direction to behaviors). The model also proposes 6 distinct components that make up Psychological *In*flexibility: Experiential Avoidance (i.e., attempts to distance oneself in some way from unwanted experiences), Lack of Contact with the Present Moment (LCPM; i.e., not paying attention to one's experiences in any given moment), Self as Content (i.e., making judgments about experiences resulting in a narrower view of self), Fusion (i.e., getting trapped in unwanted internal experiences), Inaction (i.e., inability to behave in a way that is consistent to what is important in life), and Lack of Contact with Values (i.e., being disconnected from the areas of life that are most meaningful to oneself). The dimensions of flexibility are viewed as critical to promoting individual health and well-being and are therefore promoted within ACT, whereas the dimensions of inflexibility are conceptualized as key elements associated with psychological distress. Although conceptualized as distinct, the 12 components of Flexibility and Inflexibility are also posited to be strongly interconnected with mutual facilitative relationships (Figure 1).

At a practical level, the Hexaflex model provides researchers with a set of 12 potential mechanisms of change to explain the benefits of ACT, offering a theoretically grounded method of deconstructing this therapeutic approach and revealing the key components of it that produce meaningful change. In addition, the Hexaflex model provides a set of 12 individual factors that clients bring into therapy. As a result, the Hexaflex model could offer a deeper conceptual understanding of the differences between closely related disorders (e.g., Major Depressive Disorder vs. Dysthymic Disorder) and also offers 12 dimensions to examine as potential moderators of treatment response (e.g., highlighting the individuals most and least likely to benefit from treatment). Although most directly linked to ACT, the Hexaflex model could be used in the same way to understand the mechanisms of other therapies as well. We would argue that the conceptual architecture of the Hexaflex model could offer a theoretical framework (and a set of meaningful predictors, mechanisms/mediators, and moderators) to help inform basic research examining a broad array of individual and interpersonal outcomes (e.g., health behaviors, well-being). However, to take advantage of these potential benefits of the Hexaflex model, it would be necessary to develop a multidimensional measure of psychological flexibility.

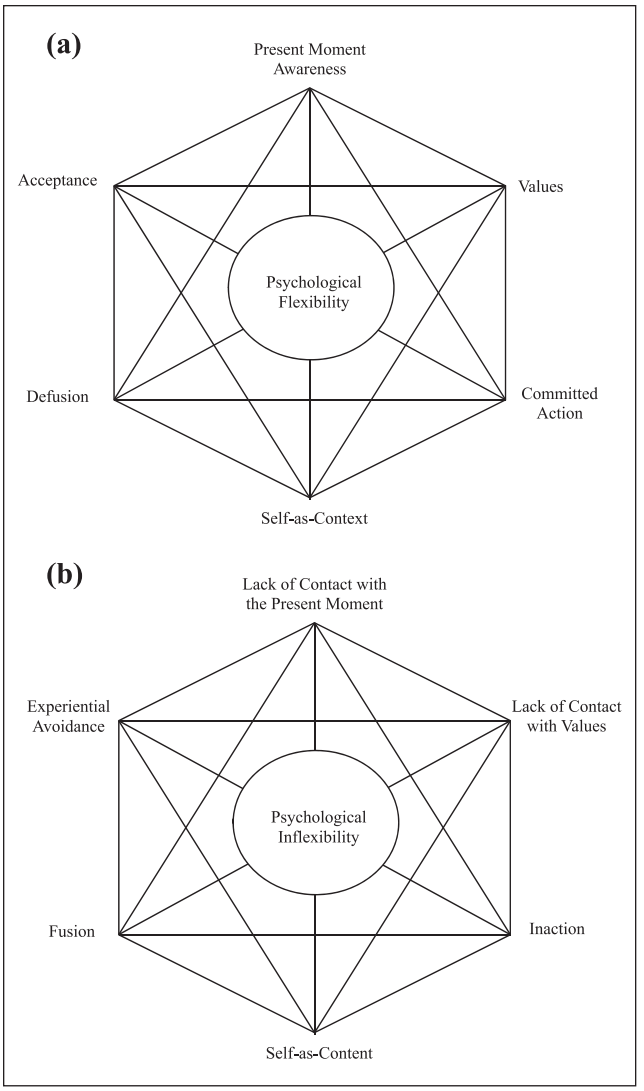

Figure 1. The psychological flexibility (Hexaflex) model.

Measuring Psychological Flexibility/Inflexibility

Previous measures of Psychological Inflexibility (e.g., AAQ, Hayes et al., 2004; AAQ-II, Bond et al., 2011; AFQ, Fergus et al., 2012; Greco et al., 2008) laid the foundation for measurement of flexibility by collapsing the 12 components of the model into a single heterogeneous measure, thereby focusing on the shared variance among the 12 dimensions in the Hexaflex model. This approach was supported by the Hexaflex model as it posits that the distinct dimensions of flexibility will be fairly strongly linked to one another. Although this offered researchers and clinicians critical tools for assessing global levels of psychological inflexibility, these scales fell short of offering a method of examining the specific components proposed in the Hexaflex model. More recently, a number of multidimensional scales have advanced measurement work in this area by tapping into specific components of the Hexaflex model (e.g., the Multidimensional Experiential

Avoidance Questionnaire, MEAQ, Gámez, Chmielewski, Kotov, Ruggero, & Watson, 2011; the Five Facet Mindfulness Questionnaire, FFMQ, Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006; the Self-Compassion Scale, SCS, Neff, 2003). Although these more recent scales offer the possibility of more fine-grained analyses, they were created from different conceptual backgrounds and with markedly different conceptual foci. As a result, these multidimensional measures fail to map clearly onto the 12 dimensions proposed within the Hexaflex model—typically taking one or two dimensions from the Hexaflex model and splitting them into multiple component dimensions (e.g., the MEAQ, the SCS) rather than trying to comprehensively and succinctly assess all 12 dimensions. Thus, researchers interested in assessing the 12 dimensions of the Hexaflex model would currently have to include as many as 22 existing scales comprising 296 items. Researchers would then have to do extensive measurement work in their data sets (e.g., exploratory and confirmatory factor analyses—EFA and CFA—item-level evaluation, scale validation) to identify the 12 dimensions within the diverse array of current measures that most clearly represent the Hexaflex model. Thus, although the fine-grained measurement work in this area has yielded some well-validated scales, it has been so highly focused from so many different conceptual directions that—in terms of the Hexaflex model—it has yielded a fairly fractured set of scales. Specifically, these scales have not been comprehensively examined in comparison with one another (to identify overlap, convergent, and discriminant validity, etc.), and as a result, their exact links to the dimensions of the Hexaflex model remain less clear, failing to offer researchers a viable method of succinctly assessing the conceptual dimensions of the Hexaflex model.

The Current Study

The current study sought to bring conceptual and measurement clarity to the existing large and diverse array of flexibility scales by developing a comprehensive and multidimensional measure of psychological flexibility to assess the 12 key dimensions of the Hexaflex model as separate subscales. Toward this end, we administered a large and diverse combined item pool of 554 possible Hexaflex items (primarily drawn from the comprehensive inclusion of all published measures of mindfulness and flexibility) to a combined online sample of 3,040 respondents. We augmented traditional analyses (e.g., EFA, CFA, item to total correlations, Cronbach alpha coefficients) with a more advanced technique which utilizes a large-sample method to increase the precision/power of scales (item response theory or IRT) along with essential yet more rarely addressed longitudinal measure development analyses (examining responsiveness to change over time). Thus, the current study sought to create a psychometrically optimized, multidimensional measure of psychological flexibility mapping onto the Hexaflex model.

We had six overarching hypotheses guiding this inquiry:

Hypothesis 1: Based on the Hexaflex model, we predicted that at least 12 distinct and separable dimensions would emerge from the massive item pool: 6 dimensions representing flexibility and 6 corresponding dimensions representing inflexibility.
Hypothesis 2: Building on a growing literature suggesting that the positive and negative components of an array of processes (e.g., affect, personality, motivation, cognitive evaluation) form more general appetitive and aversive behavior systems that are related yet meaningfully distinct from one another (Gable, Reis, & Elliot, 2003), we hypothesized that the 12 separate dimensions of psychological flexibility would further demonstrate a higher order structure in which the 6 dimensions of flexibility would form a flexibility composite and separately the 6 dimensions of inflexibility would form an inflexibility composite.
Hypothesis 3: We hypothesized that the previously developed scales for various dimensions of flexibility might have heterogeneous content and items with lower shared variance, rendering the scales including them less precise (more error/noise).
Hypothesis 4: We hypothesized that by starting with a large and diverse item pool and by using a comprehensive set of measurement analyses, the Multidimensional Psychological Flexibility Inventory (MPFI) scales would offer researchers greater amounts of precision and lower amounts of noise for assessing the dimensions of the Hexaflex model.
Hypothesis 5: Based on recent work (Rogge et al., under review), we further hypothesized that the cross-sectional precision of the MPFI would translate into comparable amounts of responsiveness for detecting change in flexibility over time in comparison with the existing scales despite the shorter lengths of the MPFI subscales.
Hypothesis 6: Finally, we hypothesized that the 12 MPFI scales would demonstrate discriminant validity with conceptually distinct constructs from the components of the Hexaflex model by showing appropriately moderate correlations with individual traits (e.g., neuroticism, emotional intelligence) and with indices of individual functioning (e.g., need satisfaction, psychological distress).

Method

Participants

A total of 3,040 respondents participated across the three studies (see Table 1 for demographics). The samples were predominantly female and Caucasian, but their large sizes offered reasonable numbers of male (n = 1,251) and

Table 1. Recruitment Sources and Demographic Characteristics of the Three Studies.

Recruitment/demographic dimension	Study 1	Study 2	Study 3	Overall
	M (*SD*)/%	*M* (*SD*)/%	*M* (*SD*)/%	*M* (*SD*)/%
Sample size	372	2,150	518	3,040
Recruitment sources				
Mechanical Turk	40	20	60	29
ResearchMatch	—	68	—	48
Undergraduate subject pool	44	7.3	40	17
Other	16	4.7	—	—
Demographics				
Female	52	60	59	59
Caucasian	69	84	74	80
Asian/Pacific Islander	18	6.5	13	9.0
African American	8.2	4.9	7.1	5.7
Other	4.8	4.6	5.9	4.8
Hispanic/Latino[a]	6.4	5.5	7.9	6.0
Age	28.7 (12.8)	34.8 (11.8)	31.8 (14.5)	33.5 (12.4)
Income ($)	24,536 (28,802)	44,425 (31,718)	27,470 (28,082)	39,102 (30,741)
Years of education	14.1 (2.3)	15.7 (2.5)	14.2 (2.1)	15.2 (2.4)
High school or less	30	7.20	16	11
Single	45	24	40	29
Dating	32	31	29	31
Engaged	1.6	4.0	2.5	3.5
Married	21	41	30	37
Had children	22	38	32	35
Dissatified in their romantic relationships	38	40	36	29
Currently in counseling	4.9	17	8.7	14.1

[a]Respondents were allowed to report a Hispanic/Latino ethnicity separately from the question assessing race. The cutoff of 13.5 on the Couples Satisfaction Inventory–4 was used to identify respondents currently dissatisfied in their romantic relationships (see Funk & Rogge, 2007).

non-Caucasian (n = 594) respondents in which to develop and validate the MPFI.

Procedure

Recruitment. All procedures were approved by a university institutional review board. As shown in the Table 1, the samples were recruited within the United States and were primarily recruited from three sources: an undergraduate subject pool of students in psychology courses from a small, private university in the North Eastern United States, the ResearchMatch.org system, and Amazon.com's Mechanical Turk service. Participants had to be at least 18 years of age to participate. The survey and all associated materials were presented in English and so participants had to be able to read English to participate. The recruitment materials presented the studies as "The Approaches to Life Study" and informed respondents that they were voluntary online surveys. Individualized feedback at the end of the initial surveys served as the primary recruitment incentive. Students from the undergraduate subject pool received extra credit toward their psychology courses and Mechanical Turk respondents received 20- to 40-cent incentives for participating.

Study 1. The first study focused on refining a large item pool down to a parsimonious set of items and dimensions representing the core of the psychological flexibility model (Hayes et al., 2011). Thus, Study 1 was a 40- to 50-minute online survey that contained a pool of 494 possible items for the MPFI and was given to 372 online respondents. Those items included the items of 22 existing scales from the ACT and mindfulness literatures as well as 84 items written by the authors. As detailed below, the additional items were written by carefully comparing the content of the item pool generated from the 22 scales with conceptual definitions of the dimensions of flexibility and inflexibility within the ACT literature (e.g., Harris, 2009; Hayes et al., 2011; Luoma, Hayes, & Walser, 2007) to ensure that each of the 12 proposed dimensions: (a) had a pool of roughly 20 to 40 potential items, (b) contained a diverse array of items representing key facets from the conceptual definitions, and (c) contained items written in a clear and concise manner requiring no more than an 8th-grade reading level (consistent with previous scales).[1]

Study 2. The second study gave a refined item pool to a large sample ($N = 2,150$) to enable the use of IRT to select the most effective items for each dimension of the MPFI. Study 2 involved a 30- to 35-minute online survey that contained a 288-item pool for the MPFI scale which included 214 items from Study 1's item pool (including the items of 12 existing scales representing the model dimensions) and another 74 items written by the authors. As with the 84 items written for the larger item pool in Study 1, these 74 items were written to ensure that each prospective dimension of the MPFI had a pool of roughly 20 to 30 potential items and contained relatively prototypic items (i.e., clear, concise items consistent with the content of the dimensions). Study 2 also collected 4-month follow-up assessments from respondents to look at the responsiveness of the MPFI for detecting change in flexibility over time.

Study 2: 4-Month Follow-Up. All respondents providing e-mail addresses were sent up to 4 e-mails over the course of roughly 3 weeks that invited them to participate in a 4-month follow-up assessment.[2] This follow-up included questions assessing perceptions of global change as well as the complete MPFI. A total of 2,109 respondents in Study 2 (98%) provided e-mail addresses. Of these participants, 970 (46%) provided follow-up data an average of 4.3 months later ($SD = 0.4$). Analysis of variance (ANOVA) and chi-square analyses contrasting respondents participating in the follow-up to those not participating suggested that respondents participating in the follow-up did not differ on rates of current meditation (see below). However, these analyses suggested that participants providing follow-up data were slightly more likely to be female, $\chi^2(1) = 29.2$, $p < .001$, $\varphi = .117$, Caucasian, $\chi^2(1) = 26.5$, $p < .001$, $\varphi = .111$, single, $\chi^2(1) = 12.4$, $p < .001$, $\varphi = -.076$, and in counseling, $\chi^2(1) = 7.1$, $p < .001$, $\varphi = .058$. Respondents giving follow-up data also tended to be somewhat older, $F(1, 2148) = 112.8$, $p < .001$, partial $\eta^2 = .050$, with slightly higher levels of education, $F(1, 2143) = 74$, $p < .001$, partial $\eta^2 = .033$, and income, $F(1, 2120) = 70.3$, $p < .001$, partial $\eta^2 = .032$). These differences were generally small in magnitude and likely emerged as statistically significant due to the power offered by such a large sample, ultimately suggesting only slight differences in the sample providing follow-up data.

Study 3. The third study focused on evaluating the discriminant validity of the MPFI subscales and placing this new scale firmly within the existing literature by examining its associations with an array of anchor scales from the nomological net surrounding psychological flexibility. Thus, Study 3 was a 25- to 30-minute online survey that contained the MPFI along with the items of an additional 18 anchor scales and was given to 518 online respondents.

Measures

To augment the ability of the scales to detect change in flexibility over time, all potential Hexaflex items were reworded into the past tense and were focused specifically on the past 2 weeks, thereby focusing on the state-like properties of these dimensions. These changes were made with the smallest changes to the actual item wordings as possible in order to maintain the character and item content/language of the original scales. Participants rated each potential Hexaflex item on a 6-point scale ranging from *never* to *always* or from *never true* to *always true*.[3] Internal consistencies are presented in Table 7, and are only presented for scales or subscales that were evaluated at the scale level (as opposed to just having subsets of their items examined within EFA, CFA, and IRT analyses). The superscripted numbers after each scale described below indicate the studies in which they were included. Unless otherwise indicated, all scale scores were created by averaging the items so that higher scores corresponded to higher levels of the construct identified in the title of the scale. To identify existing scales for the item pool, we undertook an exhaustive review of the mindfulness and ACT literatures in an effort to identify and comprehensively include all published scales.

Existing Psychological Flexibility Measures. The studies included the 9-item *Acceptance and Action Questionnaire*[1,2] (Hayes et al., 2004), the 11-item *Acceptance and Action Questionnaire–II*[1,2] (Bond et al., 2011), and the 17-item *Acceptance and Fusion Questionnaire for Youth*[1,2,3] (Greco et al., 2008; validated for use in adults, Fergus et al., 2012). In contrast to an emphasis on flexibility represented by their names, the averages for these three scales were created so that higher scores represented greater levels of psychological inflexibility. We therefore refer to them as measures of global inflexibility as this is consistent with the content and direction of their items and with how they are currently used in the literature. Despite the Hexaflex model positing 6 distinct flexibility and 6 distinct inflexibility dimensions, early measurement work effectively collapsed that down to measuring a single dimension, assuming that flexibility and inflexibility were simply polar opposites of a single spectrum. As mentioned above, our hypotheses and our analyses directly challenged such a unidimensional operationalization. To maintain consistency with its current use in the field, only the seven negatively worded items of the AAQ-II were used when creating a scale score.

Existing Scales Potentially Assessing Components of the Hexaflex Model. As seen in Table 2, the studies included 345 items representing as many as 36 dimensions from 17 existing scales. Due to the large number of overlapping items among the mindfulness measures, only the unique items from each scale were retained.

Table 2. Existing Measures and Subscales Included in the Studies.

Measure/subscale	Reference	*N* of items	Studies		
			1	2	3
Multidimensional scales					
Five Facet Mindfulness Questionnaire	Baer et al. (2006)	39 + 25 test			
Nonreactivity			×	×	
Acting with Awareness			×	×	
Nonjudging			×	×	
Observing			×		×
Description			×		×
Kentucky Inventory of Mindfulness Skills	Baer, Smith, and Allen (2004)	39			
Observe			×		
Describe			×		
Act with Awareness			×		
Accept with Judgement			×		
Toronto Mindfulness Scale	Lau et al. (2006)	15			
Decentering			×	×	
Curiosity			×		×
Philadelphia Mindfulness Scale	Cardaciotto et al. (2008)	20			
Present Moment Awareness			×		
Acceptance			×	×	
Believability of Anxious Feelings and Thoughts[a]	Herzberg et al. (2012)	16			
Somatic Concerns			×		
Emotion Regulation			×		
Negative Evaluation			×		
Difficulties in Emotion Regulation Scale	Gratz and Roemer (2004)	36			
Lack of Emotional Awareness			×	×	
Nonacceptance of Emotional Responses			×		×
Difficulties in Engaging in Goal-Directed Behavior			×		×
Impulse Control Difficulties			×		×
Limited Access to Emotion Regulation Strategies			×		×
Lack of Emotional Clarity			×		×
Multidimensional Experiential Avoidance Questionnaire	Gámez et al. (2011)	32			
Distress Endurance			×	×	
Procrastination			×	×	
Behavioral Avoidance			×	×	×
Distraction and Suppression			×	×	×
Trait Meta-Mood Scale	Salovey, Mayer, Goldman, Turvey, and Palfai (1995)	16			
Attention to Feelings[b]			×		×
Clarity of Feelings			×		×
Self-Compassion Scale	Neff (2003)	18			
Overidentification			×	×	
Self-Judgment			×		×
Mindfulness			×		×
Self-Kindness			×	×	×

(continued)

Table 2. (continued)

Measure/subscale	Reference	*N* of items	Studies		
			1	2	3
Unidimensional scales					
Cognitive and Affective Mindfulness Scale–Revised	Feldman, Hayes, Kumar, Greeson, and Laurenceau (2007)	12	×		
Experiences Questionnaire - Decentering	Fresco et al. (2007)	20	×	×	×
Freiburg Mindfulness Inventory	Walach, Buchheld, Buttenmuller, Kleinknecht, and Schmidt (2006)	22	×		
Southampton Mindfulness Questionnaire	Chadwick et al. (2008)	16	×		
Valued Living Questionnaire	Wilson, Sandoz, Kitchens, and Roberts (2010)	20	×		
Emotion Control Questionnaire–2: Rehearsal	Roger and Najarian (1989)	14	×		×
White Bear Suppression Inventory	Wegner and Zanakos (1994)	15	×		×
Rumination and Reflection Questionnaire: Rumination	Trapnell and Campbell (1999)	12	×		×
Mindful Attention and Awareness Scale[c]	Brown and Ryan (2003)	15	×	×	×
Hope Scale	Snyder et al. (1991)	8	×	×	×

[a]The authors modified the items of the Believability of Anxious Feelings and Thoughts from focusing specifically on anxious feelings to general unwanted negative emotions and thoughts in an effort to align them more closely to the Hexaflex model. As a result of these modifications, two items were dropped from the scale. [b]In contrast to its normal scoring, the Trait Meta-Mood Scale–Attention to Feelings was scored such that higher scores reflected higher inattention to feelings (consistent with the content and direction of its items). [c]In contrast to its typical scoring, the items of the Mindful Attention and Awareness Scale were averaged so that higher scores represented greater lack of contact with the present moment (consistent with the content and direction of its items). Due to the large number of overlapping items among the mindfulness measures, only the unique items from each scale were retained.

Additional Flexibility Items. To diversify the item pool beyond the contents of the existing scales, the authors wrote an additional 84 items for Study 1 and (based on the Study 1 EFA) an additional 74 items for Study 2. All of these items made use of a 6-point response scale (*never true* to *always true*). These items were written from the context of the Hexaflex model. Thus, the authors aimed to write simple and clear items that might serve as prototypical items for each of the dimensions of psychological flexibility. We also kept the items reasonably short using no more than an 8th-grade reading level (consistent with previous work). When writing these items, we drew on the second author's expertise in developing self-report scales (e.g., Funk & Rogge, 2007), the third author's extensive knowledge of ACT and the Hexaflex model, and the language used in the ACT literature to detail and elaborate the dimensions of the Hexaflex model (e.g., Harris, 2009; Hayes et al., 2011; Luoma et al., 2007). We chose to write a larger pool of items and let our EFA and IRT analyses in the 2,150 respondents from Study 2 be the deciding factor for inclusion in the final MPFI scales, obviating the need for any further vetting process. This ensured that any decisions would be based solely on psychometrics, as even decisions by seasoned experts are subject to potential biases and errors. This was an effective process as 48 of the 60 items of MPFI came from this item pool.

Outcome Measures. The studies included 20 items of the *Basic Needs Satisfaction Scale*[1,2] (Johnston & Finney, 2010) with three subscales: Autonomy, Competency, and Relatedness. The studies also included 16 items of the *Mood and Anxiety Symptom Checklist*[1,2,3] (Watson, Clark, et al., 1995; Watson, Weber, et al., 1995) separately assessing Vitality and Psychological Distress. The studies further included the 4-item *Couples Satisfaction Inventory–4* (Funk & Rogge, 2007), completed by participants in romantic relationships. Participants with children also answered a modified version of the Couples Satisfaction Inventory–4 in regard to their satisfaction with their relationship with their children. Participants also answered three health-related questions: "How many times did you get a cold/flu/sinus infection that lasted 2 days or longer?" "How many days of work/school did you

miss due to physical illness?" and "How many times did you visit the doctor due to illness?" To ensure none of the dimensions of inflexibility would inadvertently represent trait negativity, the studies also included the 8-item Neuroticism subscale of the *Big Five Inventory* (John, Donahue, & Kentle, 1991; John, Naumann, & Soto, 2008).

Meditation. To assess levels of meditation in Studies 2 and 3, we asked respondents to indicate if they had ever meditated and 1,268 of the 2,668 individuals in those studies (48%) indicated that they had. Those respondents were then asked, "How often do you currently meditate?" on the following scale (*never*, *rarely*, *once a month*, *two to three times a month*, *once a week*, *two to three times a week*, *four to six times a week, daily*, *multiple times per day*) and, "How important is meditation in your life?" on the following 6-point scale (1 = *not at all*, 2 = *a little*, 3 = *somewhat*, 4 = *quite a bit*, 5 = *very much*, 6 = *extremely*). This identified 1,049 respondents (39%) currently meditating: 379 meditating rarely (43%; $M_{importance}$ = 2.0, *SD* = 0.8), 185 meditating one to three times a month (22%; $M_{importance}$ = 2.9, *SD* = 0.9), 200 meditating one to six times a week (24%; $M_{importance}$ = 3.8, *SD* = 1.0), and 92 meditating one or more times a day (11%; $M_{importance}$ = 5.2, *SD* = 1.0).

Perceptions of Change. In the 4-month follow-up survey of Study 2, respondents answered 12 items asking how much each facet of psychological flexibility might have changed since the initial survey (*−3 got a lot worse to +3 got a lot better*). The items corresponded to each of the dimensions of the MPFI (e.g., Since completing the initial survey, how much have the following changed [if at all]: "How attentive and aware have you been of your daily life," "How committed have you been on a daily basis to taking action toward your deeper values"), and responses to the items were strongly correlated, forming a highly internally consistent scale (α = .935).[4] Responses across the 12 items were averaged, allowing us to identify a subsample of 289 respondents reporting no change on their levels of psychological flexibility/inflexibility.

Results

Identifying 12 Distinguishable Dimensions of Flexibility

Refining the Initial Item Pool. A pool of 494 items spanning 22 existing scales were given to 372 online respondents (Study 1). Items were then grouped into 12 sets based on content, representing each of the dimensions of the Hexaflex model, and each set was subjected to a separate EFA (using principle axis factoring with an oblimin rotation strategy).[5] This allowed us to identify 12 robust (i.e., most likely to replicate—containing a large number of items with strong item loadings and item content the most consistent with the conceptual definitions of the Hexaflex model; typically the largest factor extracted in each analysis) dimensions within that larger 494-item pool that corresponded to each of the conceptual dimensions of the Hexaflex model. These EFAs also identified 10 existing scales (out of the 22 represented in the study) that most directly corresponded to those dimensions of flexibility/inflexibility (as a majority of their items loaded on 1 of the 12 robust dimensions extracted). Those comparison scales are presented in Table 7. Through this process, we were able to trim the original item pool from 494 items down to 214 items. As described above, the authors then wrote an additional 74 items to create the pool for Study 2.

Identifying the Items of the Multidimensional Psychological Flexibility Inventory

Item Response Theory. The revised 288-item pool was given to 2,150 online respondents in Study 2. This large sample allowed us to use IRT (Hambleton, Swaminathan, & Rogers, 1991) to select the items of the MPFI. IRT is a large-sample statistical technique that offers a greater level of precision in creating scales over correlational techniques like factor analysis (see Hambleton et al., 1991), serving to augment those traditional approaches (see Clark & Watson, 1995). IRT is extremely effective at optimizing the properties of scales (e.g., Fraley, Waller, & Brennan, 2000; Funk & Rogge, 2007) thereby offering the possibility of creating scales that offer greater power to detect change over time and to detect treatment effects (e.g., Rogge et al., 2016). To take advantage of the benefits of IRT, each set of items corresponding to a dimension of flexibility was subjected to a separate IRT analysis using the graded response model (Samejima, 1997) in Multilog 7.0 (Thissen, Chen, & Bock, 2002).[6]

IRT Results. As seen in the first six panels of Figure 2, IRT generates item information curves (IICs) for each item in an analysis. These curves are plotted from 3 SDs below the mean on the construct being measured to 3 SDs above the mean, and the curves presented in these plots give researchers an indication of how informative an item is at assessing the construct across different ranges of that construct, with higher curves indicating higher levels of precision. Thus, wherever an IIC demonstrates a lot of height (as the *y*-axis is information), the IRT analyses are suggesting that the item in question is offering a large amount of useful information for discriminating between subjects at that level of the construct being measured. In contrast, wherever the information curves drop down to the floor of the *y*-axis, the IRT analyses are suggesting that the item in question is offering very little useful information at discriminating between subjects at that level of the construct. The first six

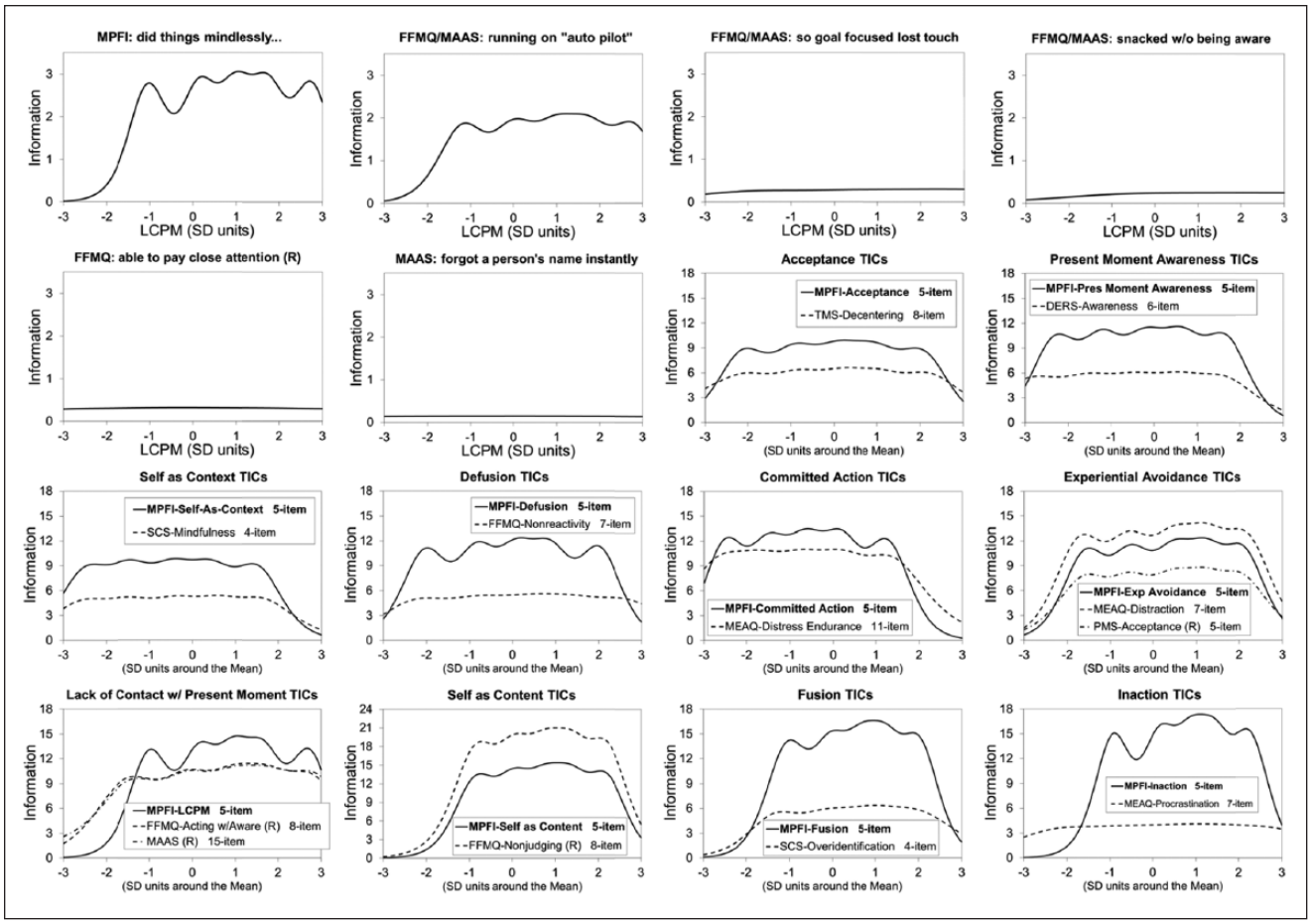

Figure 2. Item and test information curves from the IRT analyses of the 12 Hexaflex dimensions.
Note. MPFI = Multidimensional Psychological Flexibility Inventory; TMS = Toronto Mindfulness Scale; DERS = Difficulties in Emotion Regulation Scale; FFMQ = Five Facet Mindfulness Questionnaire; MEAQ = Multidimensional Experiential Avoidance Questionnaire; PMS = ; MAAS = Mindful Attention and Awareness Scale; SCS = Self-Compassion Scale; IRT = item response theory; TIC = test information curve; LCPM = Lack of Contact with the Present Moment.

panels of Figure 2 present IICs for six of the items in the IRT analysis of the items assessing LCPM. As seen in the first two panels, the MPFI-LCPM item "I did most things mindlessly without paying much attention" offered large amounts of information for assessing LCPM across a wide range of that construct (roughly from 1.5 SDs below the mean to 3 SDs above the mean) as did the item "It seemed I was 'running on automatic' without much awareness of what I was doing," which is shared by the MAAS and the FFMQ Automatic Pilot scales. However, as seen in the four subsequent panels, the IRT analyses identified items on existing scales that offered far less information for the construct being assessed (as their IICs demonstrated very little height on the *y*-axis—low information). These results are consistent with Hypothesis 3, and would suggest that such items are problematic as they would primarily be adding error variance (noise) when included on their parent scales. Although we only present a handful of IICs from the LCPM IRT analysis, all 12 IRT analyses identified problematic items from the existing scales, suggesting the need for IRT-optimized scales.

Creating the MPFI With IRT. To identify the most effective items for assessing each dimension of flexibility/inflexibility, we selected the 5 items that offered the highest levels of information across the broadest range of functioning within each of the IRT analyses. The final 60 items of the MPFI are presented in Table 3. To determine how sets of items will function as a scale, IICs can be summed to form test information curves (TICs). The remaining panels of Figure 2 present the TICs for 10 of the MPFI subscales (solid lines) along with the TICs for their comparison measures (dashed lines). As seen in those panels, for 8 of the dimensions, the IRT-optimized MPFI subscales offered comparable if not higher levels of information than the existing comparison measures despite being much shorter in length than most of those scales. This is evident in these TICs as the information curves for the MPFI subscales evidenced comparable if

Table 3. Exploratory Factor Analysis Results on the 60-Item MPFI Scale in One Random Sample Half (n = 1,079) of Study 2.

	Pattern coefficients											
Factors and items	1	2	3	4	5	6	7	8	9	10	11	12
1. Acceptance												
I made room to fully experience negative thoughts and emotions, breathing them in rather than pushing them away.	**.79**	.02	−.01	.02	.05	−.03	−.05	−.04	−.01	.05	.01	−.02
When I had an upsetting thought or emotion, I tried to give it space rather than ignoring it.	**.73**	.08	.06	−.03	.02	−.01	−.06	.02	.00	−.01	−.08	.04
I was receptive to observing unpleasant thoughts and feelings without interfering with them.[a]	**.64**	−.01	−.03	.19	−.03	.10	.00	−.03	−.03	−.03	.01	.00
I tried to make peace with my negative thoughts and feelings rather than resisting them.[a]	**.56**	.07	.09	.07	.02	.06	.04	−.03	−.06	−.12	.01	−.03
I opened myself to all of my feelings, the good and the bad.	**.46**	.29	−.03	.02	.07	.07	−.10	−.01	.01	.00	.08	−.08
2. Present Moment Awareness												
I paid close attention to what I was thinking and feeling.	−.03	**.83**	.03	−.03	.01	.00	−.05	−.03	.03	.00	.00	.03
I was attentive and aware of my emotions.[a]	.01	**.79**	.01	.00	.03	.02	−.02	−.07	.01	.04	.00	.01
I was in touch with the ebb and flow of my thoughts and feelings.	.06	**.74**	.01	.05	−.08	.07	.04	.02	−.05	.00	−.03	−.05
I was in tune with my thoughts and feelings from moment to moment.[a]	.04	**.70**	−.07	.05	.08	.03	.01	−.01	−.07	.01	−.03	.01
I strived to remain mindful and aware of my own thoughts and emotions.	.10	**.63**	.10	.01	.08	.00	.07	.00	.03	−.03	−.02	.00
3. Self as Context												
I tried to keep perspective even when life knocked me down.	−.05	.08	**.54**	.17	.06	.06	−.02	−.02	−.06	−.03	−.07	−.02
When I was scared or afraid, I still tried to see the larger picture.	−.01	.07	**.46**	.27	.06	.06	−.04	−.09	−.02	.07	−.08	−.01
Even when I felt hurt or upset, I tried to maintain a broader perspective.[a]	.17	.08	**.44**	−.01	.07	.21	.02	−.01	−.01	−.13	.03	−.07
When something painful happened, I tried to take a balanced view of the situation.	.12	.05	**.42**	.07	.11	.07	.02	−.03	−.04	−.05	.04	−.09
I carried myself through tough moments by seeing my life from a larger viewpoint.[a]	.25	.05	**.35**	.04	.12	.15	.03	−.04	−.01	−.07	.01	−.06
4. Defusion												
I was able to let negative feelings come and go without getting caught up in them.[a]	.00	−.01	.02	**.66**	.00	.11	.00	.01	−.06	−.11	.00	−.05
When I was upset, I was able to let those negative feelings pass through me without clinging to them.[a]	.03	.00	.03	**.65**	.02	.08	−.03	.00	−.06	−.08	−.02	−.03
When I was scared or afraid, I was able to gently experience those feelings, allowing them to pass.	.12	.04	.02	**.65**	.09	−.04	−.01	.02	−.07	.00	−.04	−.05
I was able to step back and notice negative thoughts and feelings without reacting to them.	.10	.05	.05	**.61**	.06	−.03	.01	−.06	.03	−.09	.00	.00
In tough situations, I was able to notice my thoughts and feelings without getting overwhelmed by them.	−.01	.09	.19	**.53**	−.01	.06	−.07	−.03	.00	.00	−.05	−.09
5. Values												
I was very in touch with what is important to me and my life.[a]	−.01	.12	−.01	−.05	**.74**	−.01	−.01	.01	−.04	−.06	−.03	−.05
I tried to connect with what is truly important to me on a daily basis.	.00	.08	−.01	.04	**.70**	−.02	−.01	−.05	.06	−.05	.00	−.08
I stuck to my deeper priorities in life.[a]	.05	−.05	−.04	.05	**.70**	.14	.04	−.03	−.02	.07	−.08	−.03
Even when it meant making tough choices, I still tried to prioritize the things that were important to me.	.00	.01	.06	.06	**.69**	.04	−.02	.01	−.03	−.03	−.08	.06
My deeper values consistently gave direction to my life.	.04	−.04	.10	.00	**.64**	.12	−.03	−.06	−.05	.02	.02	−.01
6. Committed Action												
Even when I stumbled in my efforts, I didn't quit working toward what is important.[a]	−.05	.05	.07	.01	.02	**.78**	−.02	−.04	.02	.01	.04	−.04
Even when times got tough, I was still able to take steps toward what I value in life.[a]	.02	.06	.01	.00	.03	**.78**	.02	.00	−.04	−.02	−.05	−.02
Even when life got stressful and hectic, I still worked toward things that were important to me.	−.02	.01	.01	.01	.11	**.75**	.00	−.03	.00	−.01	−.07	.03
I didn't let setbacks slow me down in taking action toward what I really want in life.	.06	−.02	.03	.01	.05	**.74**	−.02	−.01	−.05	.01	−.01	−.04
I didn't let my own fears and doubts get in the way of taking action toward my goals.	.05	.02	−.03	.05	−.01	**.70**	−.02	−.03	.02	−.08	−.04	−.07

(continued)

Table 3. (continued)

Factors and items	Pattern coefficients											
	1	2	3	4	5	6	7	8	9	10	11	12
7. Experiential Avoidance												
When I had a bad memory, I tried to distract myself to make it go away.[a]	.04	−.03	.03	−.03	−.04	.00	**.81**	−.02	.02	.04	−.03	−.02
When unpleasant memories came to me, I tried to put them out of my mind.	−.08	.05	.01	.01	.00	.01	**.80**	.04	.00	−.09	−.01	.00
I tried to distract myself when I felt unpleasant emotions.[a]	.00	−.01	.00	.00	−.04	.03	**.78**	.04	−.03	.08	−.01	.03
When something upsetting came up, I tried very hard to stop thinking about it.	−.04	−.01	−.03	.01	.02	.02	**.77**	−.04	.03	.02	.05	.01
If there was something I didn't want to think about, I would try many things to get it out of my mind.	.03	.00	−.02	.01	.05	−.05	**.76**	.00	.03	.01	.03	.01
8. Lack of Contact with the Present Moment												
I did most things on "automatic" with little awareness of what I was doing.[a]	−.02	.04	−.01	.01	.01	.00	.00	**.92**	.03	−.04	−.04	.02
I went through most days on autopilot without paying much attention to what I was thinking or feeling.	.02	−.06	−.02	.00	.00	.03	.02	**.86**	.01	−.02	.02	−.01
I floated through most days without paying much attention.	.04	−.05	.01	.02	.00	.00	−.02	**.86**	.00	.06	−.01	.02
I did most things mindlessly without paying much attention.[a]	−.02	.03	−.01	−.02	−.04	−.01	.02	**.83**	−.01	.01	.01	.01
Most of the time, I was just going through the motions without paying much attention.	−.02	−.01	.03	.01	.01	−.03	−.01	**.82**	.00	.06	.09	−.04
9. Self as Content												
I thought some of my emotions were bad or inappropriate and I shouldn't feel them.[a]	−.03	.01	−.04	.01	−.02	.03	−.01	.01	**.91**	.00	.03	−.05
I believed some of my thoughts are abnormal or bad and I shouldn't think that way.	−.03	−.02	−.05	.05	−.03	.05	−.02	.00	**.82**	.02	.01	.08
I told myself that I shouldn't be feeling the way I'm feeling.	.03	.01	.04	−.01	.05	−.07	.08	.01	**.80**	−.01	.03	−.03
I told myself I shouldn't be thinking the way I was thinking.	.02	.00	.02	−.04	.00	−.02	.04	.00	**.80**	.03	.00	.00
I criticized myself for having irrational or inappropriate emotions.[a]	.02	−.03	.02	−.05	−.03	.03	−.01	.04	**.78**	.04	−.05	.07
10. Fusion												
Negative thoughts and feelings tended to stick with me for a long time.[a]	.00	−.01	.01	−.06	.01	−.04	.03	.03	.02	**.80**	.04	.00
Distressing thoughts tended to spin around in my mind like a broken record.[a]	.02	.01	.01	−.04	−.04	−.02	.05	.03	.04	**.74**	.03	.06
It was very easy to get trapped into unwanted thoughts and feelings.	−.02	.04	−.08	.01	−.03	−.02	−.01	.05	.10	**.68**	.06	.06
When I had negative thoughts or feelings, it was very hard to see past them.	−.01	.03	−.04	−.03	−.02	.00	.03	.05	.06	**.64**	.02	.15
When something bad happened, it was hard for me to stop thinking about it.	−.03	.01	.05	−.11	−.03	−.02	.07	.08	.08	**.63**	−.01	.03
11. Lack of Contact with Values												
My priorities and values often fell by the wayside in my day-to-day life.[a]	.05	−.08	−.01	−.02	.01	−.03	−.03	.04	.02	−.04	**.72**	.07
The things that I value the most often fell off my priority list completely.	−.06	.05	−.02	.06	−.01	−.11	.02	.02	.05	.01	**.68**	.02
When life got hectic, I often lost touch with the things I value.[a]	−.01	−.01	.00	−.07	−.05	.01	−.05	.02	.01	.02	**.62**	.12
I didn't usually have time to focus on the things that are really important to me.	−.01	−.01	.04	−.05	−.07	.06	.00	.08	.00	.04	**.58**	.02
When times got tough, it was easy to forget about what I truly value.	.01	−.03	−.06	.04	−.06	−.08	−.07	.02	.01	.19	**.50**	−.01
12. Inaction												
Negative feelings often trapped me in inaction.[a]	.02	.03	−.01	−.04	−.06	−.01	.04	.03	.02	−.03	−.02	**.83**
Getting upset left me stuck and inactive.	.04	−.03	.05	−.04	−.01	−.04	.05	.00	−.03	.05	−.01	**.81**
Negative feelings easily stalled out my plans.[a]	−.04	.04	.00	−.02	.04	−.04	−.01	.06	.02	−.01	.07	**.79**
Negative experiences derailed me from what's really important.	−.03	−.03	−.04	.03	−.02	.00	−.03	.00	.05	.06	.12	**.67**
Unpleasant thoughts and feelings easily overwhelmed my efforts to deepen my life.	−.01	.01	−.07	.04	−.02	.00	.01	.00	.10	.12	.06	**.64**

Note. MPFI = Multidimensional Psychological Flexibility Inventory. The final items of the MPFI as selected by 12 separate item response theory analyses were examined with an exploratory factor analysis using the oblimin rotation strategy (to allow the factors to correlate). The scree plot and the Kaiser–Guttman criteria supported a 12-factor solution that accounted for 70% of the variance in this set of items. The strongest pattern coefficients for each item have been bolded for ease of interpretation.

[a]Denotes the top 2 items of each scale as assessed with item response theory: These items can be used to create the 12-item flexibility and 12-item inflexibility composites shown in Table 4.

Table 4. Higher Order Exploratory Factor Analysis Results on the 12 MPFI Subscales in One Random Sample Half (n = 1,079) of Study 2.

	Pattern coefficients	
Factors/subscales	Flexibility	Inflexibility
Flexibility		
Present Moment Awareness	**.89**	.21
Self as Context	**.77**	−.16
Acceptance	**.76**	.01
Contact with Values	**.68**	−.19
Committed Action	**.66**	−.24
Defusion	**.57**	−.33
Inflexibility		
Fusion	−.01	**.90**
Self as Content	.08	**.81**
Inaction	−.10	**.80**
Lack of Contact with Values	−.19	**.63**
Lack of Contact with the Present Moment	−.19	**.54**
Experiential Avoidance	.06	**.53**

Note. MPFI = Multidimensional Psychological Flexibility Inventory. The 12 subscales of the MPFI were examined with a second, higher order exploratory factor analysis using the oblimin rotation strategy (to allow the factors to correlate). The scree plot, the Kaiser–Guttman criteria, and a parallel analysis supported a two-factor solution that accounted for 61% of the variance in this set of items. The strongest pattern coefficients for each subscale have been bolded for ease of interpretation.

not greater height on the information (y-axis) than the curves for the existing scales. Consistent with this, the panel presenting the TICs for LCPM (bottom left panel) demonstrates that the MPFI-LCPM subscale offered comparable if not higher amounts of information to the MAAS and the FFMQ Acting with Awareness (reversed) scales despite being as much as three times shorter in length.[7] The 5-item MPFI Self as Content subscale offered slightly lower information than the 10-item FFMQ Nonjudging subscale. Taken as a set, the TICs presented in Figure 2 support Hypotheses 3 and 4, suggesting that the MPFI subscales represent not only short but highly effective methods of assessing the 12 dimensions of the Hexaflex model, improving on a majority of the existing scales measuring those dimensions. The increased information demonstrated by a majority of the MPFI subscales has been shown to correspond to higher levels of precision and power for detecting group differences (e.g., Funk & Rogge, 2007) and has even been shown to yield stronger treatment effects (e.g., Rogge et al., 2016), underscoring the distinct advantages that IRT can offer.

Verifying the Underlying Correlational Structure of the MPFI

Exploratory Factor Analyses. After creating the MPFI and evaluating it with IRT, we used a combination of EFA and CFA to verify its proposed factor structure. Following current best practices (e.g., Fabrigar, Wegener, MacCallum, & Strahan, 1999), we randomly split the sample of Study 2 into two halves. Table 3 presents the results of an EFA using principle axis factoring with an oblimin rotation strategy (allowing the factors to correlate) in the first random sample half (n = 1,079). Both the scree plot and the Kaiser–Guttman criteria suggested a 12-factor solution.[8] As seen by the pattern coefficients presented in Table 3, each of the 5-item MPFI subscales formed discrete factors when analyzed as a set. This result was consistent with Hypothesis 1 and suggested that our efforts were successful in creating a multidimensional scale that conceptually mapped onto the Hexaflex model. We then entered the MPFI subscale scores into a second, higher order EFA with oblimin rotation to test our hypothesis that the MPFI subscales would form a two-factor higher order structure (Hypothesis 2). Both the scree plot, the Kaiser–Guttman criteria, and a parallel analysis suggested a two-factor solution which accounted for 61% of the variance across the 12 scales. As seen in Table 4, consistent with Hypothesis 2, the 6 dimensions of flexibility formed a discrete factor. This suggests that in addition to using each of those subscales as distinct outcomes in their analyses (a strategy likely to often yield meaningfully different results across the scales), researchers could also collapse those 6 scales into a composite reflecting overall flexibility (a strategy appropriate when the 6 flexibility scales yield similar results). Similarly, the 6 dimensions of inflexibility formed a higher order factor, suggesting that those subscales could either be used separately or could be collapsed into a composite reflecting overall inflexibility.

Confirmatory Factor Analyses. To verify the dimensional structure of the MPFI, we ran CFA models on the 60 items of the MPFI in the second sample half of Study 2 (n = 1,071) using M*plus* 7.11 (Muthén & Muthén, 2012). The CFA model with 12 subscales loading onto 2 higher order factors demonstrated excellent fit, $\chi^2(1697) = 4{,}617$, root mean square error of approximation (RMSEA) = .040, comparative fit index (CFI) = .946, standardized root mean square residual (SRMR) = .060, whereas models testing simpler structures demonstrated unacceptable levels of fit: 2 higher order factors with no subscales, $\chi^2(1709) = 19{,}174$, RMSEA = .098, CFI = .676, SRMR = .085, 6 subscales (collapsing corresponding pairs of flexibility and inflexibility subscales) with 1 higher order factor, $\chi^2(1704) = 18{,}297$, RMSEA = .095, CFI = .693, SRMR = .127, and 1 global factor with no subscales, $\chi^2(1710) = 25{,}922$, RMSEA = .115, CFI = .551, SRMR = .111. As seen in Table 5, all of the items yielded strong path coefficients to their respective latent dimensions of flexibility/inflexibility and, in turn, those latent dimensions yielded strong path coefficients to their respective higher order factors. Thus, consistent with Hypotheses 1 and 2, the CFA results in the second random half of the sample continued to support the hierarchical

Table 5. Confirmatory Factor Analyses Results.

Global dimension/subscale/item	*b*	*SE*	Global dimension/subscale/item	*b*	*SE*
Global Flexibility			Global Inflexibility		
Acceptance (loading on Flexibility)	.763	.016	Experiential Avoidance (loading on Inflexibility)	.540	.025
I tried to make peace with my negative thoughts and feelings rather than resisting them.	.816	.012	I tried to distract myself when I felt unpleasant emotions.	.808	.013
I experienced myself as separate from my changing thoughts and feelings.	.791	.013	When I had a bad memory, I tried to distract myself to make it go away.	.825	.012
I opened myself to all of my feelings, the good and the bad.	.826	.012	When something upsetting came up, I tried very hard to stop thinking about it.	.815	.012
I made room to fully experience negative thoughts and emotions, breathing them in rather than pushing them away.	.784	.014	If there was something I didn't want to think about, I would try many things to get it out of my mind.	.813	.013
When I had an upsetting thought or emotion, I tried to give it space rather than ignoring it.	.796	.013	When unpleasant memories came to me, I tried to put them out of my mind.	.793	.013
Present Moment Awareness (loading on Flexibility)	.742	.017	Lack of Contact with Present Moment (loading on Inflexibility)	.792	.016
I was attentive and aware of my emotions.	.817	.012	I did most things mindlessly without paying much attention.	.870	.009
I was in tune with my thoughts and feelings from moment to moment.	.861	.010	I did most things on "automatic" with little awareness of what I was doing.	.869	.009
I was in touch with the ebb and flow of my thoughts and feelings.	.833	.011	Most of the time, I was just going through the motions without paying much attention.	.888	.008
I paid close attention to what I was thinking and feeling.	.801	.013	I floated through most days without paying much attention.	.889	.008
I strived to remain mindful and aware of my own thoughts and emotions.	.795	.013	I went through most days on autopilot without paying much attention to what I was thinking or feeling.	.877	.008
Self as Context (loading on Flexibility)	.948	.008	Self as Content (loading on Inflexibility)	.768	.016
Even when I felt hurt or upset, I tried to maintain a broader perspective.	.830	.011	I thought some of my emotions were bad or inappropriate and I shouldn't feel them.	.860	.009
I carried myself through tough moments by seeing my life from a larger viewpoint.	.798	.013	I criticized myself for having irrational or inappropriate emotions.	.871	.009
When I was scared or afraid, I still tried to see the larger picture.	.749	.015	I believed some of my thoughts are abnormal or bad and I shouldn't think that way.	.852	.010
When something painful happened, I tried to take a balanced view of the situation.	.745	.015	I told myself that I shouldn't be feeling the way I'm feeling.	.866	.009
I tried to keep perspective even when life knocked me down.	.768	.014	I told myself I shouldn't be thinking the way I was thinking.	.852	.010
Defusion (loading on Flexibility)	.825	.013	Fusion (loading on Inflexibility)	.888	.010
I was able to let negative feelings come and go without getting caught up in them.	.862	.010	Negative thoughts and feelings tended to stick with me for a long time.	.888	.007
When I was upset, I was able to let those negative feelings pass through me without clinging to them.	.840	.011	Distressing thoughts tended to spin around in my mind like a broken record.	.872	.008
When I was scared or afraid, I was able to gently experience those feelings, allowing them to pass.	.807	.012	It was very easy to get trapped into unwanted thoughts and feelings.	.907	.007
In tough situations, I was able to notice my thoughts and feelings without getting overwhelmed by them.	.812	.012	When I had negative thoughts or feelings, it was very hard to see past them.	.897	.007
I was able to step back and notice negative thoughts and feelings without reacting to them.	.800	.013	When something bad happened, it was hard for me to stop thinking about it.	.853	.009
Values (loading on Flexibility)	.837	.013	Lack of Contact with Values (loading on Inflexibility)	.792	.016
I was very in touch with what is important to me and my life.	.838	.011	When life got hectic, I often lost touch with the things I value.	.810	.013
I stuck to my deeper priorities in life.	.833	.011	My priorities and values often fell by the wayside in my day-to-day life.	.754	.016
I tried to connect with what is truly important to me on a daily basis.	.804	.013	The things that I value the most often fell off my priority list completely.	.743	.016
My deeper values consistently gave direction to my life.	.800	.013	When times got tough, it was easy to forget about what I truly value.	.763	.015
Even when it meant making tough choices, I still tried to prioritize the things that were important to me.	.788	.013	I didn't usually have time to focus on the things that are really important to me.	.704	.018
Committed Action (loading on Flexibility)	.877	.010	Inaction (loading on Inflexibility)	.923	.008
Even when times got tough, I was still able to take steps toward what I value in life.	.889	.008	Negative feelings easily stalled out my plans.	.863	.009
Even when I stumbled in my efforts, I didn't quit working toward what is important.	.835	.010	Negative feelings often trapped me in inaction.	.870	.009
Even when life got stressful and hectic, I still worked toward things that were important to me.	.870	.009	Getting upset left me stuck and inactive.	.836	.010
I didn't let setbacks slow me down in taking action toward what I really want in life.	.859	.009	Unpleasant thoughts and feelings easily overwhelmed my efforts to deepen my life.	.841	.010
I didn't let my own fears and doubts get in the way of taking action toward my goals.	.832	.011	Negative experiences derailed me from what's really important.	.858	.009
Correlation between Flexibility and Inflexibility	−.735	.017			

Note. All path coefficients were significant at $p < .001$. Model $\chi^2(1607) = 4{,}617$; comparative fit index = .946; standardized root mean square residual = .060; root mean square error of approximation = .040 (.039-.041).

structure of the MPFI. This offers a cross-validation of that correlational structure, suggesting that the factor structure of the MPFI is likely to continue to replicate in future studies. The results of the CFA model further suggested that global inflexibility and flexibility higher order factors correlate at −.735, suggesting that those two global dimensions share roughly 54% of their variance, a level of collinearity suggestive of distinct yet closely related processes.

Generalizability of Internal Consistency Across Diverse Populations

Although IRT results are reasonably sample invariant and would therefore suggest that the MPFI scales should continue to operate well across a wide range of future samples, we took advantage of the large combined sample size of Studies 2 and 3 (N = 2,668) to more directly examine how the psychometric properties of the MPFI scales would generalize across specific demographic groups. Toward this end, we calculated Cronbach's alpha coefficients for the MPFI scales along with the AAQ, the AAQ-II, and the AFQ-Y within 18 different demographic subgroups. As shown in Table 6, the MPFI scales demonstrated excellent internal consistencies across all of the subpopulations examined. Table 3 identifies the two items of each MPFI subscale identified by IRT as the most informative. This set of 24 items offers researchers and clinicians a shortened version of the MPFI. The generalizability analyses suggested that the resulting 12-item flexibility and 12-item inflexibility composites of that shortened MPFI also offer researchers highly internally consistent scales across a diverse array of subpopulations. When statistically compared with the AAQ, the AAQ-II, and the AFQ-Y in each of the subsamples (see Abd-El-Fattah & Hassan, 2011), the 30-item MPFI Inflexibility composite demonstrated significantly higher levels of internal consistency than the existing scales across all 18 subgroups.

Evaluating Responsiveness to Detecting Change Over Time

Although dimensions of psychological flexibility and inflexibility undoubtedly contain some trait-like properties, when developing the MPFI we assumed that all of these dimensions would also have state-like properties and chose to focus the measure on those. Supporting this approach, work in the mindfulness literature has demonstrated that, even without specific interventions, components of the Hexaflex model (e.g., lack of contact with present moment) can demonstrate meaningful change over time (e.g., Weinstein, Brown, & Ryan, 2009). This work highlights that the Hexaflex components can be conceptualized as dynamic processes with meaningful state-like fluctuations on as brief as a day-to-day basis.[9] To examine the responsiveness of the MPFI scales, we estimated the Minimal Detectible Change index (MDC_{95}; Stratford, Finch, et al., 1996; a statistic indicating the number of points that an individual would need to change on each scale for that change to be statistically significant) for (a) the MPFI subscales, (b) the existing global measures of inflexibility, and (c) the comparison scales that our analyses identified as assessing distinguishable components of the Hexaflex model. These MDC_{95} statistics provide researchers with concrete methods of identifying clinically significant individual change. As clinical journals (e.g., *Journal of Consulting and Clinical Psychology*) are now calling on researchers to include estimates of clinically significant change in their treatment articles, these MDC_{95} statistics will provide critical information to allow ACT researchers to answer that call. To estimate the MDC_{95} statistics for the scales in our study, we used the respondents retrospective self-reports of global change in the 4-month follow-up to identify a subpopulation of 289 individuals reporting no change in their psychological flexibility/inflexibility. Analyses within these individuals allowed us to estimate the standard error of repeated measurement for each scale (SE_{RM}), and in turn, MDC_{95} statistics.[10] As seen in the MDC_{95} column of Table 7, these estimates suggested that an individual's flexibility and inflexibility averages would need to increase roughly between 1 and 1.5 points (on the 1- to 6-point scale) between two assessments for that change to be statistically significant improvement or deterioration. The MDC_{95} estimates further revealed that longer scales (i.e., the 21-item FFMQ Automatic Pilot subscale, the 30-item MPFI Flexibility composite) tended to be more responsive to detecting individual change (requiring a change of only .94 and .90 average points between assessments) than shorter scales. This is consistent with findings in the responsiveness literature (e.g., Rogge et al., 2016), and suggests that the 30-item MPFI Inflexibility composite scale (requiring only a shift of .90 average points) could potentially offer ACT researchers a greater ability to detect individual improvement in inflexibility than existing scales like the AAQ-II and the AFQ-Y (requiring shifts of 1.26 and 1.06 average points, respectively). Consistent with Hypothesis 5, the MDC_{95} estimates also indicated that despite being notable shorter than a majority of the existing comparison scales, the individual MPFI subscales offer researchers comparable levels of responsiveness to detecting change, representing short yet highly sensitive scales.

Grounding the MPFI Within the Current Literature

Convergent Validity: Linking the MPFI Subscales to Current Measures. As seen in the "Global Inflexibility scales" section of Table 7, the MPFI Inflexibility composite

Table 6. Internal Consistencies (Cronbach's Alphas) Within Specific Demographic Groups Across the 2,668 Respondents of Studies 2 and 3.

	Gender		Age (years)				Race/ethnicity				Education level				Clinical status		Currently meditating	
Scales	Male	Female	18-24	25-34	35-44	45-73	Caucasian	African American	Asian	Hispanic/ Latino	≤High school	Some college	Bachelor's degree	Graduate degree	In treatment	Not in treatment	No	Yes
Subsample size	1,070	1,598	717	846	507	598	2,185	142	202	160	240	860	898	664	401	2,246	1,619	1,049
Global Flexibility scale																		
MPFI Flexibility composite (30-item)	.966	.965	.963	.964	.969	.967	.965	.971	.959	.962	.960	.968	.968	.961	.966	.963	.964	.967
MPFI Flexibility composite (12-item)	.903	.902	.891	.900	.916	.903	.903	.911	.882	.885	.886	.907	.910	.890	.908	.895	.898	.908
Global Inflexibility scales																		
MPFI Inflexibility composite (30-item)	**.961**	**.962**	**.955**	**.962**	**.965**	**.959**	**.962**	**.962**	**.953**	**.954**	**.948**	**.962**	**.962**	**.961**	**.959**	**.958**	**.961**	**.962**
MPFI Inflexibility composite (12-item)	.904	.908	.890	.907	.918	.902	.908	.913	.890	.888	.876	.909	.911	.902	.902	.898	.904	.911
AAQ	.736	.824	.757	.803	.822	.769	.806	.651	.670	.752	.694	.794	.788	.807	.826	.764	.788	.800
AAQ-II	.929	.926	.919	.927	.933	.928	.929	.917	.900	.921	.915	.936	.919	.924	.919	.919	.928	.924
AFQ-Y	.929	.924	.917	.933	.930	.918	.925	.931	.926	.918	.918	.930	.924	.920	.917	.921	.927	.924
MPFI Flexibility subscales																		
Acceptance	.868	.888	.859	.888	.895	.876	.885	.874	.803	.865	.960	.968	.968	.961	.886	.875	.869	.892
Present Moment Awareness	.894	.913	.896	.901	.916	.910	.909	.895	.885	.886	.875	.906	.915	.902	.897	.906	.901	.908
Self as Context	.874	.894	.885	.878	.892	.889	.889	.888	.848	.850	.875	.906	.915	.902	.893	.877	.884	.886
Defusion	.896	.906	.877	.898	.920	.912	.909	.880	.835	.887	.867	.909	.908	.900	.917	.894	.901	.905
Values	.905	.903	.891	.906	.902	.912	.904	.911	.888	.921	.901	.902	.908	.902	.895	.900	.902	.904
Committed Action	.926	.930	.913	.932	.930	.936	.930	.943	.890	.919	.905	.931	.932	.927	.936	.922	.927	.930
MPFI Inflexibility subscales																		
Experiential Avoidance	.879	.903	.885	.893	.917	.875	.897	.883	.857	.880	.863	.899	.882	.903	.904	.889	.890	.898
Lack of Contact with Present Moment	.935	.947	.934	.940	.945	.949	.946	.929	.915	.913	.925	.942	.940	.950	.943	.939	.938	.948
Self as Content	.913	.936	.907	.932	.940	.923	.929	.927	.887	.913	.886	.926	.929	.936	.934	.921	.928	.926
Fusion	.939	.943	.923	.943	.954	.944	.944	.949	.913	.924	.977	.917	.943	.944	.942	.935	.944	.938
Lack of Contact with Values	.853	.878	.855	.862	.881	.872	.870	.890	.818	.844	.860	.865	.860	.883	.867	.861	.872	.860
Inaction	.929	.928	.901	.931	.942	.936	.932	.931	.890	.904	.907	.928	.934	.924	.933	.919	.927	.931

Note. MPFI = Multidimensional Psychological Flexibility Inventory; AAQ = Acceptance and Action Questionnaire; AFQ-Y = Avoidance and Fusion Questionnaire for Youth. The internal consistencies of the MPFI 30-item Inflexibility composite were compared (using equations from Abd-El-Fattah & Hassan, 2011) with the internal consistencies of the existing inflexibility measures (i.e., AAQ, AAQ-II, and AFQ). The bolded alpha coefficients of the 30-item MPFI Inflexibility composite indicate significantly higher levels of internal consistency as assessed by pairwise comparisons ($p < .001$, for all tests) with the other inflexibility composite scales.

Table 7. Descriptives and Correlations Among Measures.

Scale	N of items	Men M	Men SD	Women M	Women SD	α	MDC_{95}	MPFI scale intercorrelations[a] 1	2	3	4	5	6	F	7	8	9	10	11	12	IN
MPFI Flexibility subscales (Studies 2 and 3)																					
1. Acceptance	5	3.5	1.0	3.4	1.0	.90	1.54	—	**.70**	**.62**	**.60**	**.52**	**.56**	**.79**	**−.18**	**−.24**	**−.23**	**−.28**	**−.19**	**−.25**	**−.29**
2. Present Moment Awareness	5	3.9	1.0	3.8	1.0	.91	1.70	**.68**	—	**.67**	**.56**	**.63**	**.62**	**.83**	−.03	**−.36**	**−.14**	**−.23**	**−.30**	**−.27**	**−.28**
3. Self as Context	5	4.0	1.0	4.0	1.0	.89	1.48	**.62**	**.56**	—	**.75**	**.71**	**.75**	**.89**	**−.07**	**−.37**	**−.28**	**−.42**	**−.42**	**−.46**	**−.43**
4. Defusion	5	3.6	1.1	3.4	1.0	.91	1.57	**.63**	**.46**	**.76**	—	**.60**	**.64**	**.82**	**−.18**	**−.35**	**−.36**	**−.53**	**−.39**	**−.49**	**−.49**
5. Values	5	4.1	1.1	4.1	1.0	.91	1.48	**.52**	**.60**	**.70**	**.59**	—	**.75**	**.84**	−.06	**−.39**	**−.25**	**−.40**	**−.47**	**−.45**	**−.42**
6. Committed Action	5	4.1	1.1	4.1	1.0	.93	1.66	**.53**	**.52**	**.75**	**.64**	**.75**	—	**.86**	**−.12**	**−.42**	**−.33**	**−.47**	**−.50**	**−.56**	**−.50**
F. Flexibility composite	30	3.9	0.9	3.8	0.8	.91	1.14	**.80**	**.77**	**.88**	**.82**	**.84**	**.85**	—	**−.13**	**−.42**	**−.32**	**−.47**	**−.45**	**−.50**	**−.48**
MPFI Inflexibility subscales (Studies 2 and 3)																					
7. Experiential Avoidance	5	3.0	1.1	3.2	1.0	.90	1.55	**−.34**	**−.14**	**−.19**	**−.29**	**−.16**	**−.19**	**−.26**	—	**.31**	**.52**	**.43**	**.33**	**.39**	**.62**
8. Lack of Contact with the Present Moment (LCPM)	5	2.5	1.1	2.6	1.1	.95	1.69	**−.36**	**−.45**	**−.46**	**−.43**	**−.52**	**−.51**	**−.55**	**.34**	—	**.51**	**.62**	**.61**	**.61**	**.77**
9. Self as Content	5	2.6	1.2	2.7	1.2	.94	1.46	**−.34**	**−.21**	**−.44**	**−.52**	**−.35**	**−.42**	**−.46**	**.47**	**.45**	—	**.71**	**.52**	**.63**	**.82**
10. Fusion	5	2.8	1.3	3.0	1.3	.95	1.47	**−.37**	**−.23**	**−.57**	**−.69**	**−.47**	**−.53**	**−.57**	**.41**	**.56**	**.70**	—	**.64**	**.77**	**.88**
11. Lack of Contact with Values	5	2.6	1.0	2.5	0.9	.87	1.52	**−.33**	**−.36**	**−.51**	**−.47**	**−.66**	**−.59**	**−.59**	**.31**	**.62**	**.50**	**.60**	—	**.72**	**.79**
12. Inaction	5	2.5	1.1	2.5	1.1	.93	1.33	**−.37**	**−.28**	**−.60**	**−.65**	**−.54**	**−.64**	**−.62**	**.38**	**.59**	**.66**	**.78**	**.68**	—	**.87**
IN. Inflexibility composite	30	2.7	0.9	2.7	0.9	.90	0.90	**−.45**	**−.35**	**−.59**	**−.65**	**−.57**	**−.61**	**−.65**	**.61**	**.75**	**.82**	**.88**	**.78**	**.87**	—

	N of items	Men M	Men SD	Women M	Women SD	α	MDC_{95}	Correlations with MPFI subscales 1	2	3	4	5	6	F	7	8	9	10	11	12	IN
Global Inflexibility scales (Study 2)																					
AAQ	9	3.2	0.7	3.2	0.7	.79	—	**−.34**	**−.21**	**−.42**	**−.49**	**−.36**	**−.44**	**−.45**	**.51**	**.50**	**.68**	**.73**	**.53**	**.68**	(.77)
AAQ-II	7	2.6	1.2	2.5	1.1	.93	1.25	**−.40**	**−.28**	**−.57**	**−.61**	**−.49**	**−.58**	**−.59**	**.47**	**.59**	**.74**	**.82**	**.65**	**.83**	(.87)
AFQ-Y	16	2.7	1.0	2.7	0.9	.93	1.06	**−.38**	**−.28**	**−.54**	**−.58**	**−.48**	**−.56**	**−.54**	**.47**	**.59**	**.73**	**.80**	**.65**	**.81**	(.87)
Comparison Flexibility scales (Study 2)																					
Acceptance (TMS Decentering)	8	3.5	0.9	3.3	0.8	.93	1.49	(.82)	**.65**	**.68**	**.68**	**.56**	**.61**	**.80**	**−.21**	**−.35**	**−.34**	**−.43**	**−.32**	**−.40**	**−.44**
Present Moment Awareness (DERS Awareness)	6	3.9	1.0	4.0	0.9	.88	1.58	**.60**	(.73)	**.58**	**.45**	**.56**	**.48**	**.68**	**−.17**	**−.38**	**−.24**	**−.22**	**−.32**	**−.27**	**−.33**
Self as Context (SCS Mindfulness)	4	3.8	1.0	3.7	0.9	.83	1.57	**.61**	**.57**	(.84)	**.70**	**.64**	**.65**	**.81**	**−.20**	**−.41**	**−.40**	**−.50**	**−.44**	**−.51**	**−.52**
Defusion (FFMQ Nonreactivity)	7	3.3	0.9	3.2	0.9	.85	1.46	**.56**	**.42**	**.64**	(.73)	**.48**	**.52**	**.67**	**−.22**	**−.29**	**−.35**	**−.49**	**−.35**	**−.45**	**−.46**
Committed Action (MEAQ Distress Endurance)	11	4.0	0.9	4.0	0.8	.90	1.34	**.45**	**.43**	**.66**	**.52**	**.63**	(.73)	**.69**	**−.09**	**−.33**	**−.27**	**−.34**	**−.43**	**−.48**	**−.41**
Comparison Inflexibility scales (Study 2)																					
Experiential Avoidance (MEAQ Distraction)	7	3.1	1.0	3.1	1.0	.92	1.44	**−.31**	**−.13**	**−.17**	**−.26**	**−.12**	**−.17**	**−.23**	(.94)	**.31**	**.48**	**.39**	**.30**	**.37**	**.58**
Experiential Avoidance (PMS Acceptance—reversed)	10	3.1	1.0	3.1	1.0	.91	1.28	**−.37**	**−.21**	**−.36**	**−.46**	**−.29**	**−.35**	**−.41**	(.88)	**.45**	**.71**	**.64**	**.48**	**.58**	**.79**
LCPM (FFMQ Acting with Awareness—reversed)	21	2.8	0.7	3.0	0.7	.94	0.99	**−.29**	**−.30**	**−.36**	**−.43**	**−.35**	**−.38**	**−.43**	**.41**	(.73)	**.58**	**.63**	**.57**	**.61**	**.74**
LCPM (MAAS—reversed)	15	2.7	0.8	2.8	0.7	.90	0.93	**−.33**	**−.37**	**−.41**	**−.45**	**−.40**	**−.41**	**−.48**	**.42**	(.76)	**.58**	**.62**	**.60**	**.63**	**.76**
Self as Content (FFMQ Nonjudging—reversed)	11	2.7	1.0	2.7	1.1	.94	1.18	**−.32**	**−.18**	**−.41**	**−.49**	**−.30**	**−.39**	**−.42**	**.51**	**.46**	(.96)	**.72**	**.49**	**.64**	**.81**
Fusion/Inaction (SCS Overidentification)	4	2.7	1.1	3.0	1.2	.87	1.34	**−.36**	**−.26**	**−.55**	**−.64**	**−.42**	**−.50**	**−.55**	**.44**	**.54**	**.67**	(.81)	**.58**	(.75)	**.81**
Inaction (MEAQ Procrastination)	7	3.0	1.0	3.0	0.9	.89	1.30	**−.37**	**−.33**	**−.46**	**−.48**	**−.49**	**−.56**	**−.54**	**.44**	**.55**	**.46**	**.55**	**.57**	(.64)	**.67**

(continued)

Table 7. (continued)

	N of items	Men		Women		α	MDC_{95}	Correlations with MPFI subscales													
		M	*SD*	*M*	*SD*			1	2	3	4	5	6	F	7	8	9	10	11	12	IN
Conceptually Distinct scales (Study 3)																					
BFI Neuroticism	8	2.9	0.9	3.3	1.0	.89	—	**−.21**	**−.22**	**−.46**	**−.60**	**−.44**	**−.52**	**−.49**	**.18**	**.44**	**.49**	**.69**	**.47**	**.66**	**.63**
TMM Emotional Intelligence: Inattention to Feelings	7	3.1	0.7	2.7	0.7	.69	—	**−.33**	**−.42**	**−.11**	.05	**−.19**	**−.10**	**−.21**	**.10**	**.26**	**.12**	**−.09**	**.10**	.00	**.10**
TMM Emotional Intelligence: Clarity of Feelings	5	3.9	0.9	3.8	0.9	.79	—	**.42**	**.61**	**.50**	**.49**	**.57**	**.56**	**.62**	−.08	**−.40**	**−.30**	**−.31**	**−.33**	**−.36**	**−.38**
SCS Kindness to Self (Study 2)	5	3.4	1.0	3.5	1.0	.87	—	**.48**	**.48**	**.55**	**.58**	**.52**	**.53**	**.62**	.02	**−.32**	**−.30**	**−.39**	**−.32**	**−.37**	**−.36**
SCS Judgement of Self	5	3.0	1.1	3.1	1.1	.88	—	**−.15**	**−.22**	**−.31**	**−.38**	**−.34**	**−.38**	**−.36**	**.23**	**.55**	**.66**	**.69**	**.58**	**.67**	**.72**
FFMQ Observing	15	3.6	1.0	3.6	1.0	.94	—	**.52**	**.58**	**.46**	**.41**	**.45**	**.42**	**.56**	.08	**−.16**	.02	−.05	−.08	−.07	−.06
FFMQ Description	10	3.3	0.6	3.4	0.6	.90	—	**.47**	**.53**	**.41**	**.37**	**.43**	**.41**	**.52**	.04	**−.09**	.01	.00	−.01	−.01	−.01
EQ Rumination	6	3.7	0.9	3.8	0.8	.77	—	**.19**	**.22**	.06	**−.11**	.02	−.06	.05	**.21**	**.31**	**.42**	**.48**	**.35**	**.41**	**.46**
TMS Curiosity	6	3.1	1.0	3.2	1.0	.91	—	**.49**	**.41**	**.30**	**.24**	**.23**	**.16**	**.36**	.04	**.12**	**.24**	**.17**	**.15**	**.20**	**.20**
Individual Functioning Indices (Study 2)																					
Currently in counseling	1	10%	—	21%	—	—	—	**−.18**	**−.10**	**−.23**	**−.25**	**−.20**	**−.23**	**−.24**	**.15**	**.20**	**.25**	**.31**	**.22**	**.32**	**.30**
Vitality	8	3.3	1.0	3.2	0.9	.94	—	**.42**	**.41**	**.58**	**.55**	**.60**	**.61**	**.64**	**−.22**	**−.46**	**−.44**	**−.57**	**−.51**	**−.56**	**−.59**
Psychological Distress	8	2.2	1.1	2.1	1.0	.95	—	**−.33**	**−.28**	**−.51**	**−.54**	**−.48**	**−.56**	**−.54**	**.33**	**.53**	**.61**	**.74**	**.58**	**.71**	**.75**
BNSS Autonomy	6	4.0	0.9	4.1	0.9	.81	—	**.44**	**.46**	**.60**	**.56**	**.66**	**.61**	**.67**	**−.30**	**−.52**	**−.48**	**−.60**	**−.64**	**−.62**	**−.66**
BNSS Relatedness	8	4.2	0.9	4.5	0.8	.87	—	**.34**	**.36**	**.53**	**.45**	**.58**	**.53**	**.56**	**−.18**	**−.44**	**−.37**	**−.49**	**−.52**	**−.52**	**−.53**
BNSS Competency	6	4.0	0.9	4.1	0.9	.82	—	**.41**	**.41**	**.60**	**.56**	**.67**	**.67**	**.67**	**−.25**	**−.53**	**−.49**	**−.61**	**−.63**	**−.64**	**−.67**
In a romantic relationship	1	60%	—	72%	—	—	—	.04	.03	**.04**	**.05**	**.06**	**.05**	**.05**	−.04	−.03	**−.07**	**−.05**	−.02	**−.07**	**−.06**
Relationship satisfaction with partner (n = 1,460)	4	14.0	4.9	14.0	4.7	.95	—	**.19**	**.20**	**.26**	**.24**	**.29**	**.28**	**.29**	**−.09**	**−.23**	**−.18**	**−.26**	**−.27**	**−.27**	**−.28**
Has children	1	38%	—	37%	—	—	—	.01	.00	.04	**.06**	**.07**	.03	**.04**	.02	.00	−.02	−.03	.02	−.03	−.01
Relationship satisfaction with children (n = 795)	4	15.1	4.4	14.9	4.2	.87	—	**.22**	**.22**	**.29**	**.28**	**.34**	**.29**	**.32**	−.06	**−.18**	**−.16**	**−.22**	**−.20**	**−.25**	**−.22**
N of colds in past 6 months	1	0.6	1.8	0.6	1.1	—	—	**−.04**	**−.04**	**−.06**	**−.06**	**−.06**	**−.07**	**−.07**	**.06**	**.08**	**.11**	**.09**	**.08**	**.10**	**.11**
N of days of work missed due to illness in past 6 months	1	1.1	6.9	1.3	5.0	—	—	−.02	−.02	−.02	−.03	−.03	−.04	−.03	**.08**	**.05**	**.08**	**.07**	**.06**	**.09**	**.09**
N of doctor visits in the past 6 months	1	0.5	1.7	0.9	2.3	—	—	−.02	−.01	−.01	−.04	.01	**−.05**	−.02	**.07**	**.06**	**.10**	**.11**	**.09**	**.11**	**.12**
Currently meditating	1	35%	—	43%	—	—	—	**.09**	**.13**	**.09**	.03	**.06**	**.06**	**.09**	−.04	−.02	.00	.01	−.02	.01	−.01
Frequency of current meditation per month (n = 856)	1	4.1	5.9	3.5	5.8	—	—	**.13**	**.10**	**.10**	**.13**	**.09**	.05	**.12**	**−.08**	−.06	−.02	−.06	−.05	−.06	−.07

Note. MPFI = Multidimensional Psychological Flexibility Inventory; AAQ = Acceptance and Action Questionnaire; AFQ-Y = Avoidance and Fusion Questionnaire for Youth; TMS = Toronto Mindfulness Scale; DERS = Difficulties in Emotion Regulation Scale; FFMQ = Five Facet Mindfulness Questionnaire; MEAQ = Multidimensional Experiential Avoidance Questionnaire; PMS = ; MAAS = Mindful Attention and Awareness Scale; SCS = Self-Compassion Scale; BFI = Big Five Inventory; TMM = Trait Meta-Mood Scale; EQ = Experiences Questionnaire; BNSS = Basic Needs Satisfaction Scale. For ease of interpretation, all statistically significant ($p < .05$) correlations have been bolded. Correlations in parentheses represent validity correlations between the MPFI subscales and their conceptual counterparts from the existing literature.

[a]For the MPFI subscale intercorrelations, correlations within men are presented above the diagonal and correlations in women are presented below the diagonal. The remaining correlations were calculated across both genders.

MDC_{95} = Minimal Detectible Change index (Stratford, Finch, et al., 1996) and refers to the number of points an individual's average score must change (on the 1- to 6-point scale employed in the current study) between assessments for that change to be statistically significant ($p < .05$).

demonstrated appropriately strong correlations with the three most widely used measures of inflexibility (i.e., the AAQ, the AAQ-II, and the AFQ-Y; cells of the correlations in question were are presented in parentheses for clarity), suggesting that the MPFI Inflexibility composite is indeed measuring a common construct. Thus, any future findings with the psychometrically optimized MPFI Inflexibility scale would represent direct extensions of the work done with these three scales. In contrast, the MPFI Flexibility composite demonstrated only moderate correlations with the existing global scales, and therefore represents a novel contribution to the field.

As mentioned above, the EFA results of Study 1 identified existing scales whose items mapped onto 10 dimensions of the MPFI. These existing scales often had different labels (derived from the differing theoretical approaches underlying their original development) than the Hexaflex dimension our results suggested they were assessing. For example, the existing scale that most directly assessed the MPFI dimension of acceptance was the (Toronto Mindfulness Scale) TMS Decentering subscale. However, despite this different label, an inspection of the specific items suggested the items of the TMS Decentering subscale (e.g., "I approached each experience by trying to accept it, no matter whether it was pleasant or unpleasant") were highly similar in content to the MPFI Acceptance items (e.g., "I opened myself to all of my feelings, the good and the bad"), helping explain why the items of the TMS Decentering subscale loaded onto the same factor as the MPFI Acceptance items in the EFAs in both Studies 1 and 2.[11] To facilitate the presentation of these comparison scales, the "Comparison Flexibility scales" and "Comparison Inflexibility scales" sections of Table 7 identify the Hexaflex dimensions our results suggested they were assessing. The cells presenting correlations between these existing scales and their MPFI counterparts are presented within parentheses for clarity. As seen in Table 7, the MPFI subscales indeed demonstrated appropriately strong correlations with their comparison scales, suggesting convergent validity.

Discriminant Validity of the MPFI. To further validate the MPFI subscales, we examined how they correlated with conceptually distinct constructs.[12] Four of the constructs in the "Conceptually Distinct scales" section of Table 7 were clearly distinct from the Hexaflex model based on previous research and theory (emotional intelligence inattention to and clarity of feelings, neuroticism, and curiosity) as were the 15 distinct indices of individual functioning in the final section of Table 7. These 21 scales were included in Studies 2 or 3 for the express purpose of demonstrating discriminant validity of the MPFI scales. However, as described above, we were purposefully overinclusive in creating the original 494-item pool (drawing from 22 different scales). Thus, the EFAs in Study 1 helped identify 5 additional scales that were less centrally linked to the Hexaflex model and could therefore be considered additional external correlates for these discriminant validity analyses (e.g., FFMQ Description, FFMQ Observing, SCS Kindness to Self). As seen in the "Conceptually Distinct scales" and the "Individual Functioning Indices" sections of Table 7, the MPFI subscales demonstrated low to moderate correlations (in the expected directions) with all of the conceptual boundary constructs examined. Consistent with Hypothesis 6, these results suggest that the MPFI subscales are not inadvertently measuring constructs outside of the Hexaflex model and are yielding results consistent with the predictions of the Hexaflex model. For example, the MPFI subscales of Values and Committed Action showed the strongest correlations (as tested with the Meng, Rosenthal, & Rubin, 1992, approach) with current levels of vitality. This suggests that individuals who generally stick to their deeper priorities and who are able to keep taking steps toward important aspects of their lives despite setbacks also tended to feel more energized and happy in their lives. Similarly consistent with expectations, the MPFI Fusion subscale demonstrated the strongest correlation with current levels of psychological distress, suggesting that individuals who often find themselves getting trapped in negative thoughts and feelings also tended to report higher levels of feeling dissatisfied, sad, discouraged, and depressed. Furthermore, the MPFI Fusion and Inaction subscales demonstrated stronger associations with neuroticism than the other MPFI inflexibility subscales, suggesting that individuals who often found themselves getting stuck in negative thoughts (i.e., high fusion) and individuals who often found themselves getting trapped in inaction by negative thoughts (i.e., high inaction) tended to be more globally negative toward life (i.e., high neuroticism). The MPFI Present Moment Awareness subscale demonstrated one of the strongest correlations with the TMM Clarity of Feelings subscale, suggesting that individuals who tended to be very attentive and aware of their own feelings also tended to report high levels of clarity toward their own feelings. The distinct patterns of correlations across the MPFI subscales suggests that those subscales will often offer distinct patterns of results when treated as separate constructs in analyses. These results also ground the new MPFI scales within the larger literatures on this wide range of existing scales.

Discussion

Examining a massive item pool across a combined sample of 3,040 respondents allowed us to develop the first measure (the MPFI) that comprehensively assesses the 12 dimensions of psychological flexibility and inflexibility posited by the Hexaflex model.

Implications

The MPFI Integrates and Extends Previous Measurement Work. The current study sought to build on the groundbreaking measurement work that has resulted from the rise of ACT (Hayes et al., 1999) as a promising, third-wave, evidence-based, behavioral therapy. As mentioned above, much of this foundational work focused on creating global scales to assess inflexibility (e.g., the AAQ, Hayes et al., 2004; the AAQ-II, Bond et al., 2011; the AFQ-Y, Greco et al., 2008) whose global nature failed to offer researchers a method of examining the specific dimensions posited within the Hexaflex model. Although a number of multidimensional scales have also been developed in this field, most of those focus on specific subcomponents of the Hexaflex model rather than attempting to assess the model in its entirety. By incorporating all of these measures into the item pool for the MPFI, the current study not only developed a scale that comprehensively assesses the dimensions of the Hexaflex model but it also brought some conceptual clarity to the literature by synthesizing these diverse scales into a cohesive framework grounded in the Hexaflex model.

Flexibility and Inflexibility Represent as Many as 12 Distinct Processes That Can Vary Independently of One Another. Consistent with the Hexaflex model, the 12 dimensions of the MPFI were significantly correlated with one another, suggesting that they tended to be linked with one another across individuals. However, a closer inspection of the intercorrelations of the MPFI subscales in Table 7 reveals a number of key points. Turning to the dimensions of flexibility, those MPFI subscales correlated from .46 to .76 with each other, yielding an average intercorrelation of .63. This suggests that those scales share anywhere from 21% to 57% of their variance, and on average share about 40% of their variance with one another. Thus, although those dimensions of flexibility are reasonably correlated with one another, each of those dimensions also contains quite a bit of unique variance. As a result, each of the dimensions of flexibility could contribute unique information to models and could likely change and vary independent of one another (e.g., an individual could show increases in acceptance without necessarily showing corresponding increases in committed action). Similarly, the dimensions of inflexibility correlated from .31 to .78 with each other, indicating that those dimensions share from 10% to 61% of their variance, and on average share about 30% of their variance. This continues to suggest that although those dimensions are correlated, an individual could likely show improvements on one dimension of inflexibility (e.g., lack of present moment awareness) without necessarily showing improvements on other inflexibility dimensions (e.g., fusion).

The correlations between the sets of flexibility and inflexibility subscales are even more modest in size, ranging from −.03 to −.69 with an average flexibility subscale to inflexibility subscale correlation of −.37. This suggests that the individual flexibility and inflexibility subscales share anywhere from 0.1% to 47% of their variance, and on average only share about 14% of variance. This echoes our EFA and CFA results, and at a practical level this suggests that flexibility and inflexibility are not simply two extremes of a single dimension. In fact, these results would suggest that an individual could likely improve on a specific dimension of inflexibility (e.g., becoming less experientially avoidant) without necessarily showing comparable improvement on the corresponding dimension of flexibility (e.g., without showing a marked increase in acceptance). Although linked, the correlational results with the MPFI would suggest that those could be two distinct processes, worthy of being assessed and modeled independently. This is consistent with findings for negative and positive affect suggesting that they are distinguishable processes rather than extremes of a common dimension as the absence of negative affect does not necessarily imply high levels of positive affect (e.g., Watson, Clark, & Tellegen, 1988). In fact, evidence suggests that the positive and negative components of constructs like affect, personality, cognitive evaluation, and motivation might form more general appetitive and aversive behavioral systems that can vary independently of one another (Gable et al., 2003). The current findings would suggest that flexibility and inflexibility might also fit into those larger appetitive and aversive behavioral systems.

An alternative explanation for the more modest correlations seen could be that individuals might vary in their specific dimensions of flexibility and inflexibility across different domains of their lives. For example, and individual might have high levels of committed action at work but at the same time have high levels of inaction in their interpersonal relationships with friends and family. Such an individual might appear to have somewhat elevated levels of both committed action and inaction, thereby helping lower the magnitude of the negative correlation between those two scales in the sample. Regardless of the underlying cause of the more modest correlations, the initial correlational results with the MPFI would suggest that researchers and clinicians could gain useful insights by assessing and modeling flexibility and inflexibility as distinct (yet related) processes.

At a broader theoretical level, we would argue that these correlational results, combined with the EFA and CFA results, offer some conceptual clarity as to the higher order structure of the components of the Hexaflex model. Although previous work has conceptualized flexibility as simply as a single dimension, or as six distinct dimensions, the results presented here offer compelling evidence to support 12 distinct dimensions forming two higher order composites. Thus, rather than suggesting that committed action is the polar opposite of inaction, for example, our results

suggest that the 6 dimensions of flexibility are more closely related to one another than they are to the dimensions of inflexibility, and vice versa. The development of the MPFI allowed such questions to be comprehensively examined for the first time in the literature. These results therefore help clarify the theoretical foundations of the Hexaflex model by highlighting the nature of its underlying correlational structure, thereby providing critical information for researchers interested in fine-grained analyses of flexibility and inflexibility in their models.

The MPFI Is Sensitive to Change. Most clinical scales are developed primarily in cross-sectional data sets and yet are routinely used as outcome measures in longitudinal treatment studies. This practice assumes that the cross-sectional properties of a scale will translate into that scale being appropriately responsive to detecting change over time. Unfortunately, this is not always a fair assumption as scales can vary widely in their abilities to detect change, leading researchers to highlight the evaluation of responsiveness to change as a critical (yet often ignored) measure development step (see Husted, Cook, Farewell, & Gladman, 2000). Consistent with this, leading clinical journals like the *Journal of Consulting and Clinical Psychology* have begun insisting that treatment researchers translate their findings into clinically meaningful rates of individual improvement and deterioration using statistics like the Reliable Change index (Jacobson & Truax, 1991) or its more recent iteration, the MDC_{95} (Stratford, Finch, et al., 1996). However, to be able to offer clinically meaningful information on individual change, it is first necessary to evaluate the responsiveness to change of the outcome measures being used (i.e., estimating the MDC_{95} for specific outcome scales—a statistic indicating the number of points that an individual would need to change on each scale for that change to be statistically significant). The current article is one of the first in the psychological flexibility literature to address this critical aspect of measure development and provide researchers with MDC_{95} estimates that could be used in future longitudinal studies across all of the flexibility/inflexibility scales examined. The results suggested that the 5-item MPFI subscales offered comparable levels of responsiveness to that seen with the existing measures, despite being much shorter in length. The results also support the state-like conceptualization of the dimensions of psychological flexibility and inflexibility in the MPFI.

The MPFI Is Flexible. Both researchers and clinicians can use the MPFI in a variety of ways to assess the psychological flexibility and inflexibility of individuals. Our EFA, CFA, and correlational results suggested that the MPFI subscales are separable and distinct but at the same time they share as much as 40% to 50% of their variance with each other. This is an optimal level of overlap between subscales that are designed to be used both separately and occasionally as a composite (see Clark & Watson, 1995). At a practical level, this suggests that if the 12 individual subscales are used (and/or modeled) separately to form a more complex psychological flexibility profile for an individual, which will yield interesting and distinct results for that individual or in that model. Thus, a fairly inflexible client might specifically be very high on fusion and inaction but notably less inflexible on the dimensions of experiential avoidance, LCPM, self as content, and lack of contact with values. Using the MPFI subscales would therefore provide a clinician (or a researcher) with that clinically useful information. However, the correlations among the subscales also indicate that often these different dimensions simply go hand in hand with one another. Thus, some individual flexibility/inflexibility profiles might just show elevations across all 6 dimensions of flexibility or inflexibility. Similarly, although the individual dimensions of the MPFI will often give unique results in models, sometimes they will converge on a common inflexibility or flexibility finding. In both instances, the composite scores could be used. To provide even more flexibility in the use of the MPFI, Table 3 identifies the 2 items of each MPFI subscale identified as most effective in the IRT analyses which can be used in a 12-item flexibility and a separate 12-item inflexibility composite. Those 12-item composites demonstrated high internal consistency across a diverse range of demographic subgroups. Given their shorter lengths, these two 12-item scales would offer lower precision and responsiveness to change than their 30-item counterparts. Despite this, they offer researchers and clinicians short, optimized scales, strongly grounded in the Hexaflex model that would be more manageable for purposes like weekly assessments within clinical settings, daily diary studies, or national surveys.

The MPFI Opens New Lines of Research. Given the conceptual grounding of the MPFI within the Hexaflex model, the MPFI offers researchers a set of 14 new outcome measures (i.e., 6 dimensions of flexibility, 6 dimensions of inflexibility, and 2 global composite dimensions) to examine the benefits of ACT on psychological flexibility more precisely. The MPFI will therefore provide ACT researchers the tools they need to start deconstructing the benefits of ACT across the dimensions of the Hexaflex model, allowing them to more clearly examine the associations within the Hexaflex model as well as its links to individual functioning and well-being. Building on a recent work examining individual components of the Hexaflex model as potential mechanisms of action within treatment studies (e.g., Arch et al., 2012; Wicksell, Olsson, & Hayes, 2010), the subscales of the MPFI offer researchers a set of potential mechanisms of change (i.e., mediators) within ACT and cognitive behavioral therapies (e.g., Arch et al., 2012), conceptually grounded within the Hexaflex model. Thus, the use of the

MPFI to explore such mechanisms of change could help theoretically organize these inquiries. In addition, the subscales of the MPFI could be examined as individual characteristics of clients that might affect their response to treatment (i.e., moderators). Thus, psychological flexibility profiles generated by the MPFI could offer clinicians critical information for effectively tailoring treatment to the needs of individual clients.

The MPFI also offers researchers the first global measure of psychological flexibility (in its 30-item flexibility composite). Scales like the AAQ-II and AFQ-Y have demonstrated that ACT can reduce clients' psychological inflexibility, thereby enhancing quality of life and reducing symptoms across a variety of disorders (see Hayes et al., 2006, for a review). However, with the advent of the MPFI, it will now be possible to demonstrate that ACT is likely to also increase individuals' psychological flexibility. Use of the MPFI would further allow researchers to determine which global dimension (flexibility or inflexibility) is more strongly linked to various aspects of quality of life for individuals with different disorders. Thus, we would argue that the MPFI opens a wide range of future lines of research within the ACT field.

Limitations and Future Directions

Despite the promising results with the MPFI scale presented, the results of the current study were limited by a number of factors. First, the sample was primarily Caucasian and female. This was somewhat ameliorated by the sheer size of the sample as it still included 1,251 men and 594 non-White participants. Despite that, future studies should seek more diverse samples to ensure that the findings presented here will generalize. Second, the study was conducted entirely online; as this requires Internet access, it might have biased the sample toward higher levels of education and income. Future work should examine similar hypotheses with methods that reach a broader cross-section of society (e.g., national samples from random digit dialing). Third, the study used only self-report data. Diversifying this methodology in future work might actually yield stronger results. For example, the MPFI could be evaluated in comparison with therapist ratings on various dimensions of client flexibility/inflexibility in future studies to further validate the MPFI subscales. Fourth, the MPFI was developed and validated within community samples. Future work should seek to validate its use within clinical samples to ensure that it will still offer exceptional psychometric properties in clinical settings. Along these same lines, following the groundbreaking work on the AFQ-Y, future studies could seek to validate the use of the MPFI in adolescent samples. Fifth, although it is likely that this sample was predominantly U.S. citizens who spoke English as a first language, we did not specifically assess nationality or primary language. Thus, future studies should specifically evaluate the ability of the MPFI scales to work in respondents of other nationalities and for whom English might be a second language. Sixth, overall 29% of the respondents were paid incentives for completing the surveys, incurring the risk that those payments might have influenced survey responses. In the current study, those incentives were markedly small (20-40 cents) and only offered to a fraction of the sample, thereby reducing the likelihood that the payments would have unduly swayed the findings presented. Finally, although the responsiveness analyses yielding MDC_{95} statistics offer researchers a concrete method of assessing clinically significant individual change in treatment studies, they relied on only two waves of assessment (mirroring the pre–post design of most treatment studies). Future studies should seek to augment the responsiveness analyses presented here with more sophisticated approaches that make use of multiwave assessments (e.g., 28-35 days of daily diary data) to carefully model the degree to which the items of subscales reliably show change in the same direction across waves of assessment (e.g., Cranford et al., 2006).

Despite these limitations, the current study provided compelling evidence to support the MPFI as a method of comprehensively assessing the components of the Hexaflex model. By incorporating and directly building on previous measurement work in this area, the MPFI integrates that previous work into a single, unified scale more parsimoniously tied to the contents of the Hexaflex model. By using IRT in a large online sample, the MPFI was psychometrically optimized so that each of the 12 dimensions could be measured by short 5-item (or even 2-item) scales without losing precision or responsiveness to change over time, thereby minimizing the length of the overall scale as much as possible. Taken as a set, the results presented here offer compelling evidence supporting the MPFI as an effective measure of the Hexaflex model, offering a critical first step in being able to explore the numerous ways the Hexaflex model can inform our understanding of both ACT therapy and individual well-being.

Acknowledgments

We would like to thank our respondents for supporting this research.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was funded by an internal grant at the University of Rochester, Department of Clinical and Social Sciences in Psychology.

Notes

1. We used three readability tools to evaluate the readability of the item pool: the Flesch–Kincaid Grade Level Formula (Kincaid, Fishburne, Rogers, & Chissom, 1975), the Mc Laughlin SMOG Formula (Mc Laughlin, 1969), and the Linsear Write Readability Formula (Scott, 2016). Fifteen groups of roughly 30 to 40 items (roughly 400 words) were submitted to readability analyses to comprehensively examine reading levels required across the entire item pool. These analyses suggested that on average the MPFI item pool required about a 6th-grade reading level (Flesch–Kincaid grade level = 6.11, SMOG grade level = 6.21, Linsear Write grade level = 6.57), and all sets of items evaluated (including the items of the legacy scales) fell below an 8th-grade reading level. The items of the MPFI were also submitted to these same readability analyses. Consistent with the larger item pool, the items of the MPFI required about a 6th- or 7th-grade reading level (Flesch–Kincaid grade level = 6.40, SMOG grade level = 6.60, Linsear Write grade level = 7.15).
2. We chose a 4-month follow-up interval based on analyses in previous data sets exploring change over time on globally evaluative measures of individual, relationship, and family functioning (e.g., Funk & Rogge, 2007; Saavedra, Chapman, & Rogge, 2010; Peltz, Rogge, Sturge-Apple, O'Connor, & Pigeon, in press). These analyses have tended to suggest that a 3- to 4-month follow-up interval strikes a nice balance, allowing researchers to capture both meaningful change and meaningful stability on such outcome constructs.
3. We chose these two response scales as the vast majority of the items of the existing scales were written in a way that was compatible with one of the two sets of responses. Thus, selecting these response scales allowed us to put the items of the item pool onto common response sets with only minimal changes in the wording of the original items of those existing scales. Placing the items onto common response sets facilitated cross-measure analyses and the creation of the MPFI without the need for a large array of different response sets across its items. Thus, all items in the pool were presenting using one of these two 6-point response scales instead of their original response options (if their original response options differed).
4. Although this represents a new scale, the items were written to ask about change using prototypical language mapping directly onto each dimension of flexibility/inflexibility. Similarly novel global change scales and, in fact, this larger approach of using self-reported retrospective reports of global change to identify clinically meaningful change groups have been successfully used to examine responsiveness of self-report measures across a diverse range of constructs (e.g., lower back pain, Stratford, Finch, et al., 1996; fatigue and emotional functioning, Jaeschke, Singer, & Guyatt, 1989; quality of life, Juniper Guyatt, Willan, & Griffith, 1994; relationship satisfaction, Rogge et al., 2016). Despite the retrospective nature of this approach, scales demonstrating higher responsiveness as assessed by these methods have also been shown to yield stronger prospective treatment effect sizes (Rogge et al., 2016), suggesting that this approach offers clinically useful information. It should be further noted that in the current article, these retrospective reports were only used to identify what would be termed a "clinically stable" population of individuals whose psychological flexibility did not noticeably change over the 4-month interval—restricting our use of this new scale to the most concrete answer choice (i.e., no change).
5. Although it would have been preferable to run one EFA on the entire 494-item pool, this was simply not possible. That many highly correlated items in a single analysis created too much multicollinearity in the correlation matrices to allow the analyses to run. Thus, the items were split into 12 groups and run in 12 separate analyses. These analyses in smaller sets of items allowed the EFAs to run and to identify 12 dominant (i.e., the dimensions with the largest numbers of items) dimensions that conceptually mapped onto the dimensions of the Hexaflex model. The factor scores for the primary factors were saved from those EFAs. We then examined how each item out of the 494 correlated with all 12 resulting dimensions (essentially creating the equivalent of a structural matrix), allowing us to examine potential cross-loadings of items with other factors than their primary factor. This step also helped us ensure that the items had been appropriately grouped when running the 12 separate EFAs. The results suggested that only a small handful of items had been initially misclassified, and those items were promptly shifted to their correct groups and the corresponding EFAs were run again.
6. IRT assumes: (a) that the items being analyzed within a specific analysis are measuring the same construct (i.e., unidimensional) and (b) that the items being analyzed are not excessively redundant with one another (i.e., locally independent). Before running the IRT analyses on each set of items, we subjected each set to an EFA to ensure that set was unidimensional and examined item-to-item partial correlations within each set (partialling out the sum of that set) to ensure that the items in the set were locally independent. This yielded the 12 distinct pools of items ranging in size from 14 to 36 items.
7. As TICs are the sum of IICs, the more items a scale has, the greater the information it can be expected to offer in its TIC, particularly when it is made up of items that all offer high levels of information. Thus, a scale with 15 items could be expected to yield a much higher TIC than an 8- or 5-item scale, consistent with the knowledge that longer scales tend to offer more information and precision (e.g., Rogge et al., 2015). The fact that the 5-item MPFI subscales offered comparable if not higher information to scales that were up to three times longer in length suggests not only that the MPFI subscales have been psychometrically optimized (i.e., made up of particularly effective items) but also that the existing scales must contain suboptimal and even problematic items (i.e., items that fail to offer any information relevant to the construct being assessed by the scale—essentially, items uncorrelated with the rest of the scale).
8. This solution accounted for 70% of the variance in the 60 items of the MPFI and mapped onto the dimensions of the Hexaflex model (see Table 3) demonstrating both statistical (i.e., the Kaiser–Guttman criteria, % of variance) and conceptual (the Hexaflex model) support for such a solution.

This 12-factor solution also yielded strong factor loadings for each of the 5-item subscales. In contrast, when solutions with fewer factors were attempted, those reductions in the number of factors yielded increasingly large numbers of items with poor (e.g., below .4) loadings on their primary factors. Thus, the EFA results suggested that a 12-factor solution represented the most parsimonious description of the correlational structure within the MPFI items. Consistent with this, the 12-factor solution was also robustly supported by a CFA (see Table 5) in the second random sample half (described below), further suggesting that it represents the most appropriate solution.

9. In fact, we would assert that the various dimensions of psychological flexibility and inflexibility likely influence and are influenced by other interindividual processes (e.g., interpersonal conflict) and intraindividual processes (e.g., perceived stress; see Weinstein et al., 2009). Consistent with this dynamic conceptualization, we drew on methodology commonly used for other treatment outcome measures (e.g., the PANAS, Watson et al., 1988; also see Fredrickson & Joiner, 2002) to focus the scales on the state-like properties of outcome constructs by presenting all of the items in the MPFI item pool in the past tense and asking respondents to focus on the past 2 weeks when answering the items. This helped ensure that the MPFI would be maximally useful to future longitudinal researchers by focusing the scales on the aspects of psychological flexibility and inflexibility that could change over time, either in response to other processes or in response to interventions.
10. To evaluate the responsiveness to change of the MPFI subscales, we first estimated the SE_{RM} (an estimate of noise over time) for each scale by running repeated measures ANOVAs for each dimension of flexibility/inflexibility within the 289 respondents reporting no change in their flexibility/inflexibility at 4 months. As those individuals did not perceive any change in psychological flexibility, any change in scores between the two assessments can be assumed to be due to noise in measurement over time. We therefore estimated the SE_{RM} by plugging the mean squared error across time (MSE) from each repeated measure ANOVA into the following equation: SE_{RM} = SQRT (2 * MSE) (see Guyatt, Walter, & Norman, 1987). We then used the following formula, MDC_{95} = 1.96 * SE_{RM}, to calculate each scale's MDC_{95} (see Stratford, Finch, et al., 1996). Such retrospective self-reports are a widely used method in the responsiveness to change literature (for a review, see Stratford, Binkley, & Riddle, 1996), particularly when a well-validated objective measure for evaluating change on the construct of interest does not yet exist (as is the case for psychological flexibility). More important, higher levels of responsiveness based on these retrospective ratings have been shown to translate into stronger treatment effects when several outcome measures are directly compared (e.g., Rogge et al., 2016), underscoring the clinical utility of this information.
11. More specifically, this suggests that the aspect of acceptance that robustly emerged in our Study 1 EFAs and our Study 2 EFA and CFA was in fact originally termed decentering in the literature and assessed with the TMS Decentering subscale. The items of the TMS Decentering subscale loaded on the same factor as the items forming the MPFI Acceptance subscale in Study 1. As shown in Table 7, those two scales also correlated .82 in Study 2. Taken as a set, these results indicate that the TMS Decentering subscale is measuring the same construct as the MPFI Acceptance subscale. Although we also describe the item content overlap between the two scales for readers to help ground these findings back into the specific items used, it was the statistical findings (not the item content) that drove our identification of the TMS Decentering subscale as the comparison flexibility scale that directly corresponded to the MPFI Acceptance subscale. As shown in the TICs for these two scales in Figure 2, although the MPFI acceptance scale was shorter (5-item) than its equivalent TMS subscale (8-item), our IRT analyses suggested that the MPFI scale offered higher levels of information/precision for assessing this construct.
12. To clearly delineate the construct being measured by a scale, it is critical to evaluate the limits of what it is measuring by showing that it demonstrates appropriately moderate links to measures of related yet conceptually distinct scales (i.e., its discriminant validity; see Clark & Watson, 1995). In fact, Cronbach and Meehl (1955) compellingly argued that researchers most fully demonstrate the construct validity of a scale (ensuring that it is measuring the construct it was designed to assess) by demonstrating its discriminant links to anchor scales in the nomological net, a web of conceptually distinct constructs that have been theoretically and/or empirically linked to the construct being measured by the scale, a point echoed by Campbell and Fiske (1959).

References

Abd-El-Fattah, S. M., & Hassan, H. K. (2011). Dependent-alpha calculator: Testing the differences between dependent coefficients alpha. *Journal of Applied Quantitative Methods*, *6*, 59-61.

Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Assessment of mindfulness by self-report. *Assessment*, *11*, 191-206. doi: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*, 27-45. doi:10.1177/1073191105283504

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., . . . Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: A revised measure of the psychological inflexibility and experiential avoidance. *Behavior Therapy*, *42*, 676-688. doi:10.1016/j.beth.2011.03.007

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*, 822-848. doi:10.1037/0022-3514.84.4.822

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105. doi:10.1037/h0046016

Cardaciotto, L., Herbert, J. D., Forman, E. M., Moitra, E., & Farrow, V. (2008). The assessment of present-moment

awareness and acceptance. *Assessment*, *15*, 204-223. doi:10.1177/1073191107311467

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*, 451-455. doi:10.1348/014466508X314891

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 309-319. doi:10.1037/1040-3590.7.3.309

Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin*, *32*, 917-929.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302. doi:10.1037/h0040957

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*, 272-299.

Feldman, G., Hayes, A., Kumar, S., Greeson, J., & Laurenceau, J. (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*, 177-190. doi:10.1007/s10862-006-9035-8

Fergus, T. A., Valentiner, D. P., Gillen, M. J., Hiraoka, R., Twohig, M. P., Abramowitz, J. S., & McGrath, P. B. (2012). Assessing psychological inflexibility: The psychometric properties of the Avoidance and Fusion Questionnaire for Youth in two adult sample. *Psychological Assessment*, *24*, 402-408. doi:10.1037/a0025776

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology*, *78*, 350-365. doi:10.1037//0022-3514.78.2.350

Fredrickson, B. L., & Joiner, T. (2002). Positive emotions trigger upward spirals toward emotional well-being. *Psychological Science*, *13*, 172-175.

Fresco, D. M., Moore, M. T., van Dulmen, M. H. M., Segal, Z. V., Helen Ma, S., Teasdale, J. D., & Williams, J. M. G. (2007). Initial psychometric properties of the Experiences Questionnaire: Validation of a self-report measure of decentering. *Behavior Therapy*, *38*, 234-246. doi:10.1016/j.beth.2006.08.003

Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology*, *21*, 572-583. doi:10.1037/0893-3200.21.4.572

Gable, S. L., Reis, H. T., & Elliot, A. J. (2003). Evidence for bivariate systems: An empirical test of appetition and aversion across domains. *Journal of Research in Personality*, *37*, 349-372.

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*, 692-713. doi:10.1037/a0023242

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, *26*, 41-54.

Greco, L. A., Lambert, W., & Baer, R. A. (2008). Psychological inflexibility in childhood and adolescence: Development and evaluation of the Avoidance and Fusion Questionnaire for Youth. *Psychological Assessment*, *20*, 93-102. doi:10.1037/1040-3590.20.2.93

Guyatt, G. H., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Disease*, *40*, 171-178.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Harris, R. (2009). *Act made simple*. Oakland, CA: New Harbinger.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York, NY: Kluwer Academic/Plenum.

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*, 1-25. doi:10.1016/j.brat.2005.06.006

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York, NY: Guildford Press.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2011). *Acceptance and commitment therapy: The process and practice of mindful change*. New York, NY: Guildford Press.

Hayes, S. C., Strosahl, K., Wilson, K. G., Bissett, R. T., Pistorello, J., Toarmino, D., . . . McCurry, S. M. (2004). Measuring experiential avoidance: A preliminary test of a working model. *Psychological Record*, *54*, 553-578.

Herzberg, K. N., Sheppard, S. C., Forsyth, J. P., Credé, M., Earleywine, M., & Eifert, G. H. (2012). The Believability of Anxious Feelings and Thoughts Questionnaire (BAFT): A psychometric evaluation of cognitive fusion in a nonclinical and highly anxious community sample. *Psychological Assessment*, *24*, 877-891. doi:10.1037/a0027782

Husted, J. A., Cook, R. J., Farewell, V. T., & Gladman, D. D. (2000). Methods for assessing responsiveness: A critical review and recommendations. *Journal of Clinical Epidemiology*, *53*, 459-469.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12-19.

Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status: Ascertaining the minimal clinically important difference. *Controlled Clinical Trials*, *10*, 407-415.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory: Versions 4a and 54*. Berkeley: University of California.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement,

and conceptual issues. In O. P. John, R. W. Robins & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). New York, NY: Guilford Press.

Johnston, M. M., & Finney, S. J. (2010). Measuring basic needs satisfaction: Evaluating previous research and conducting new psychometric evaluations of the Basic Needs Satisfaction in General Scale. *Contemporary Educational Psychology*, *35*, 280-296.

Juniper, E. F., Guyatt, G. H., Willan, A., & Griffith, L. E. (1994). Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *Journal of Clinical Epidemiology*, *47*, 81-87.

Kincaid, J. P., Fishburne, R. P., Jr., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for navy enlisted personnel (No. RBR-8-75)*. Millington, TN: Naval Technical Training Command.

Lau, M. A., Bishop, S. R., Segal, Z. V., Buis, T., Anderson, N. D., Carlson, L., . . . Carmody, J. (2006). The Toronto Mindfulness Scale: Development and validation. *Journal of Clinical Psychology*, *62*, 1445-1467. doi:10.1002/jclp.20326

Luoma, J. B., Hayes, S. C., & Walser, R. D. (2007). *Learning ACT: An acceptance & commitment therapy skills-training manual for therapists*. Oakland, CA: New Harbinger.

Mc Laughlin, G. H. (1969). SMOG Grading—A new readability formula. *Journal of Reading*, *12*, 639-646.

Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlation coefficients. *Quantitative Methods in Psychology*, *111*, 172-175. doi:10.1037/0033-2909.111.1.172

Muthén, L. K., & Muthén, B. O. (2012). M*plus* (Version 7.11) [Computer software]. Los Angeles, CA: Author.

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, *2*, 223-250.

Peltz, J. S., Rogge, R. D., Sturge-Apple, M. L., O'Connor, T. G., & Pigeon, W. R. (in press). Reciprocal influences among family processes and Toddler's Sleep Problems. *Journal of Family Psychology*.

Roger, D., & Najarian, B. (1989). The construction and validation of a new scale for measuring emotion control. *Personality and Individual Differences*, *10*, 845-853. doi:10.1016/0191-8869(89)90020-2

Rogge, R. D., Crasta, D., Maniaci, M. R., Funk, J. L., & Lee, S. (under review). *How well can we detect shifts in relationship satisfaction over time? Evaluating responsiveness to change in relationship satisfaction scales*. Manuscript submitted for publication.

Saavedra, M. C., Chapman, K. E., & Rogge, R. D. (2010). Clarifying links between attachment and relationship quality: Hostile conflict and mindfulness as moderators. *Journal of Family Psychology, 24*, 380-390.

Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., & Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale. In J. W. Pennebaker (Ed.), *Emotion, disclosure, and health* (pp. 125-154). Washington, DC: American Psychological Association.

Samejima, F. (1997). Graded response model. In W. J. Van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 85-100). New York, NY: Springer.

Scott, B. (2016). *Free Text Readability Consensus Calculator*. Retrieved from http://www.readabilityformulas.com/freetests/six-readability-formulas.php

Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., . . . Harney, P. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology*, *60*, 570-585. doi:10.1037/0022-3514.60.4.570

Stratford, P. W., Binkley, J. M., & Riddle, D. L. (1996). Health status measures: Strategies and analytic methods for assessing change scores. *Physical Therapy*, *76*, 1109-1123.

Stratford, P. W., Finch, E., Solomon, P., Binkley, J., Gill, C., & Moreland, J. (1996). Using the Roland-Morris Questionnaire to make decisions about individual patients. *Physiotherapy Canada*, *48*, 107-110.

Thissen, D., Chen, W. H., & Bock, D. (2002). Multilog user's guide: Multiple, categorical item and test scoring using item response theory (Version 7.0) [Computer software]. Lincolnwood, IL: Scientific Software.

Trapnell, P. D., & Campbell, J. D. (1999). Private self-consciousness and the five-factor model of personality: Distinguishing rumination from reflection. *Journal of Personality and Social Psychology*, *76*, 284-304. doi:10.1037/0022-3514.76.2.284

Walach, H., Buchheld, N., Buttenmuller, V., Kleinknecht, N., & Schmidt, S. (2006). Measuring mindfulness—The Freiburg Mindfulness Inventory (FMI). *Personality and Individual Differences*, *40*, 1543-1555. doi:10.1016/j.paid.2005.11.025

Watson, D., Clark, L., Weber, K., Assenheimer, J. S., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: II. Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. *Journal of Abnormal Psychology*, *104*, 15-25. doi:10.1037/0021-843X.104.1.15

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*, 1063-1070.

Watson, D., Weber, K., Assenheimer, J. S., Clark, L., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology*, *104*, 3-14. doi:10.1037/0021-843X.104.1.3

Wegner, D. M., & Zanakos, S. (1994). Chronic thought suppression. *Journal of Personality*, *62*, 615-640. doi:10.1111/j.1467-6494.1994.tb00311.x

Weinstein, N., Brown, K. W., & Ryan, R. M. (2009). A multi-method examination of the effects of mindfulness on stress attribution, coping, and emotional well-being. *Journal of Research in Personality*, *43*, 374-385.

Wicksell, R. K., Olsson, G. L., & Hayes, S. C. (2010). Psychological flexibility as a mediator of improvement in acceptance and commitment therapy for patients with chronic pain following whiplash. *European Journal of Pain*, *14*, 1059.e1-1059.e11. doi:10.1016/j.ejpain.2010.05.001

Wilson, K. G., Sandoz, E. K., Kitchens, J., & Roberts, M. (2010). The Valued Living Questionnaire: Defining and measuring valued action within a behavioral framework. *Psychological Record*, *60*, 249-272.