# Dispositional Mindfulness in Context: Cultural and Individual Perspectives

**By**

**Johannes Alfons Karl**

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Doctor of Philosophy in Psychology

**Victoria University of Wellington**

**2021**

# Abstract

Mindfulness, which was derived from Buddhist philosophy and practice, is often defined as "paying attention in a particular way, on purpose, in the present moment, and nonjudgmentally". The practice of secular mindfulness exercises has received substantial interest in psychology over the last decade and mindfulness-based practices are now widely implemented in clinical interventions. Previous research has identified stable individual differences in mindfulness which are present even in non-practitioners. My research builds on this body of work and explores (i) the current state and directions in the literature on trait mindfulness research; (ii) the relationship between trait mindfulness and established individual differences such as personality and reinforcement sensitivity; and (iii) the cross-cultural applicability of current mindfulness measures.

In the first study in this thesis, I used recent developments in bibliometric analysis to examine the development of the field of trait mindfulness, identifying important research areas in this line of work and patterns of cross-national collaboration. I found 1229 documents in the time span from 2005 to 2021 using a search in the Web of Science. Examining the complete corpus of literature that referenced trait mindfulness, I found that current research approaches focus more on clinically relevant outcomes than on potential predictors of mindfulness, which manifested in substantial clusters of themes around well-being and treatment. I also found substantively more articles published by authors working in Western countries than in the majority world. This indicates that research appears to be biased both towards clinical outcomes of mindfulness and skewed towards Western cultural contexts and concerns.

In my next study, I examined the replicability of the Five Facet Mindfulness Questionnaire (FFMQ) to explore whether the same five major dimensions of mindfulness emerge in a different sample 15 years later. The FFMQ contains five facets: Non-Judging (non-evaluation of thoughts and feelings), Non-Reacting (ability to not act on negative thoughts and emotions), Acting with Awareness (awareness of self in the moment), Describing (labelling and expressing experiences), and

Observing (awareness of sensory experiences). Following the overall protocol of the original study and using a range of currently available mindfulness measures, I found that the facets of the FFMQ could largely be retrieved in this conceptual replication. In addition, new measures of “Western” mindfulness were empirically separable from measures based in Buddhist conceptualizations. This supports the use of multi-facetted mindfulness measures to capture self-reported mindfulness.

In the second part of my thesis, I focused on potential individual-level predictors of the facets of mindfulness. In Study 3, I joined two previously separated lines of research by jointly examining the relationship between mindfulness, reinforcement sensitivity, and personality. In contrast to previous studies, I found that the facets of mindfulness might be differentially related to supposed biological (reinforcement sensitivity) and cognitive (personality) individual differences while accounting for their overlap. Specifically, Neuroticism, which in past studies was related to Non-Judging and Non-Reacting, was only related to Non-Reacting. In turn, Non-Judging was predicted by behavioral inhibition, but Non-Reacting was not.

In Study 4, I moved from cross-sectional analyses to a 4-month longitudinal investigation, using recent advances in modelling to separate within and between-individual relationships. In contrast to the cross-sectional investigation, I found a more complex pattern of relationships, including potential feedback loops between individual differences and mindfulness. Specifically, I found that the expression of supposed biological differences in long-term orientation predicted individuals’ level of awareness, but in turn higher awareness also predicted greater long-term orientation. This provides a tentative mechanistic explanation of the link between Acting with Awareness and health-behaviors identified in previous studies.

In the third part of the thesis, I focus on the applicability of mindfulness measures across cultures. As indicated above, mindfulness emerged in Eastern contexts but is currently studied in Western societies. Hence, I test how well the FFMQ as the gold standard of mindfulness trait measures performs across cultures. To provide a toolkit for cross-cultural researchers, I present a synthesis of standards for cross-cultural comparisons and developed a proto-type of an R-package that implements various methodological advances and analytical tools. In the final study, I applied these tools to examine the suitability of the FFMQ for cross-cultural comparisons across 16 countries. Overall, I found that the FFMQ is substantially biased towards higher income and more individualistic contexts and shows substantial variation across cultures. This finding implies that the FFMQ might not be suitable in its current form for cross-cultural comparisons, possible due to cultural differences in the understanding of Acting with Awareness, which in an exploratory study is separated into awareness of mind and body. This indicates that additional research is necessary to ensure the cross-cultural comparability of mindfulness and to advance research.

In my general discussion, I explore both methodological and conceptual avenues for future research in trait mindfulness. Returning to questions of individual differences in mindfulness, I highlight how recent advances in network modelling might allow researchers to untangle the differences in between and within-individual relationships observed in this thesis. I present some evidence of the application of network models from research on personality, to highlight the usefulness of this technique for future research on mindfulness. Focusing on cultural differences in structure and functionality, I review various lines of research that indicate that mindfulness-like features may be found in various cultural contexts, but may be differently experienced and expressed, as indicated by my psychometric examination of the FFMQ. I outline how researchers taking a functionalist approach might link current mindfulness approaches with different philosophical and cultural approaches to enrich the nomological network and present initial evidence on these relationships.

# Acknowledgements

A thesis is as much about the final project as it is about the way. For both I owe thanks and gratitude to many.

To my supervisors Ron and Paul: For providing me with a scientific role model, the chance to indulge in my passion projects, and the opportunity to teach alongside you.

To my friends and colleagues at Victoria, Kealagh, Torven, Nuzha, Inkuk, Tram, Fin, and Gerard: For putting up with me and giving me the opportunity to learn from you.

To my parents Christa and Eberhard: For enabling me to follow a path that was not open to you and your trust in me successfully walking it.

To my wife Tanja: For your unwavering love, the inspiration you provide me, and the future you created with me.

# Publications During Candidature

**2021**

Fischer, R., Karl, J. A., & Pilati, R. (2021). Cultural syndromes in a changing world: A longitudinal investigation of brazilian jeitinho social problem-solving strategies. Personality and Social Psychology Bulletin. https://doi.org/10.1177/01461672211043385

Fischer, R., Karl, J., Fontaine, J., & Poortinga, Y. (2021). Evidence of validity does NOT rule out systematic bias: A commentary on nomological noise and cross-cultural invariance. PsyArXiv. https://doi.org/10.31234/osf.io/k9wbj

Fischer, R., Sinwongsuwat, K., Tepsing, P., & Karl, J. A. (2021). Mapping the minds of spectators during extreme ritual: A psychological network perspective. https://doi.org/10.31234/osf.io/sgyz4

Karl, J. A., Fischer, R., & Jose, P. E. (2021). The development of mindfulness in young adults: The relationship of personality, reinforcement sensitivity, and mindfulness. Mindfulness, 12, 1103–1114. https://doi.org/10.1007/s12671-020-01576-3

Karl, J. A., Johnson, F. N., Bucci, L., & Fischer, R. (2021). In search of mindfulness: A review and reconsideration of cultural dynamics from a cognitive perspective. Journal of the Royal Society of New Zealand, 1–24. https://doi.org/10.1080/03036758.2021.1915804

Karl, J., & Fischer, R. (2021). Affect and state mindfulness. PsyArXiv. https://doi.org/10.31234/osf.io/jqhu7

**2020**

Fischer, R., Bortolini, T., Karl, J. A., Zilberberg, M., Robinson, K., Rabelo, A., Gemal, L., Wegerhoff, D., Nguyễn, T. B. T., Irving, B., & others. (2020). Rapid review and meta-meta-analysis of self-guided interventions to address anxiety, depression, and stress during COVID-19 social distancing. Frontiers in Psychology, 11. https://doi.org/10.3389/fpsyg.2020.563876

Fischer, R., & Karl, J. A. (2020a). Experimental methods in cross-cultural management. The SAGE Handbook of Cross-Cultural Management, 111–126.

Fischer, R., & Karl, J. A. (2020b). Predicting behavioral intentions to prevent or mitigate COVID-19: A cross-cultural meta-analysis of attitudes, norms, and perceived behavioral control effects. Social Psychological and Personality Science, 19485506211019844. https://doi.org/10.1177/19485506211019844

Fischer, R., & Karl, J. A. (2020c). The network architecture of individual differences: Personality, reward-sensitivity, and values. Personality and Individual Differences, 160, 109922. https://doi.org/10.1016/j.paid.2020.109922

Fischer, R., Karl, J. A., Luczak–Roesch, M., Fetvadjiev, V. H., & Grener, A. (2020). Tracing personality structure in narratives: A computational Bottom–Up approach to unpack writers, characters, and personality in historical context. European Journal of Personality, 34(5), 917–943. https://doi.org/10.1002/per.2270

Karl, J. A., & Fischer, R. (2020). Revisiting the five-facet structure of mindfulness. Measurement Instruments for the Social Sciences, 2(1), 7. https://doi.org/10.1186/s42409-020-00014-3

Karl, J. A., Méndez Prado, S. M., Gračanin, A., Verhaeghen, P., Ramos, A., Mandal, S. P., Michalak, J., Zhang, C.-Q., Schmidt, C., Tran, U. S., Druica, E., Solem, S., Astani, A., Liu, X., Luciano, J. V., Tkalčić, M., Lilja, J. L., Dundas, I., Wong, S. Y. S. Y., … Fischer, R. (2020). The cross-cultural validity of the Five-Facet Mindfulness Questionnaire across 16 countries. Mindfulness, 11, 1226–1237. https://doi.org/10.1007/s12671-020-01333-6

Karl, J., Verhaeghen, P., Aikman, S. N., Solem, S., Lassen, E., & Fischer, R. (2020). Stoicism and wellbeing. PsyArXiv. https://doi.org/10.31234/osf.io/6rtny

Ko, A., Pick, C. M., Kwon, J. Y., Barlev, M., Krems, J. A., Varnum, M. E., Neel, R., Peysha, M., Boonyasiriwat, W., Brandstätter, E., & others. (2020). Family matters: Rethinking the psychology of human social motivation. Perspectives on Psychological Science, 15(1), 173–201. https://doi.org/10.1177/1745691619872986

Singh, P., Tewari, S., Kesberg, R., Karl, J. A., Bulbulia, J., & Fischer, R. (2020). Time investments in rituals are associated with social bonding, affect and subjective health: A longitudinal study of Diwali in two Indian communities. Philosophical Transactions of the Royal Society B, 375(1805), 20190430. https://doi.org/10.1098/rstb.2019.0430

Ward, C., Kim, I., Karl, J. A., Epstein, S., & Park, H.-J. (2020). How normative multiculturalism relates to immigrant well-being. Cultural Diversity and Ethnic Minority Psychology, 26(4), 581. https://doi.org/10.1037/cdp0000317

Ward, C., Watters, S. M., Stuart, J., & Karl, J. A. (2020). Normative multiculturalism in socio-political context. Wiser World with Multiculturalism. 24th Congress of the International Association for Cross-Cultural Psychology, Guelph, Canada.

**2019**

Chrystal, M., Karl, J. A., & Fischer, R. (2019). The Complexities of "Minding the Gap": Perceived Discrepancies Between Values and Behavior Affect Well-Being. Frontiers in Psychology, 10, 736. https://doi.org/10.3389/FPSYG.2019.00736

Karl, J. A., & Fischer, R. (2019). Individual differences and mindfulness. PsyArXiv. https://doi.org/10.31234/OSF.IO/Z2CX6

# Table of Contents

# Table of Figures

# Table of Tables

# Statement of Authorship

Four studies of the current thesis have been published prior in peer-reviewed journals. Study 4 and 6 have been published in the journal *Mindfulness*. Study 5 was published in *Frontiers of Psychology.* Study 2 was published in *Measurement Instruments in the Social Sciences.* In this thesis, a revised version of each manuscript is presented with minor stylistic changes and with added context to add coherence to the overall thesis. I contributed at least 50% of the work for all manuscripts. Study 3 has been published as pre-print and Study 1 is currently under review for publication. Parts of the argument presented in the general introduction and discussion have been published as part of an article in the *Journal of the Royal Society of New Zealand*; I am primary author on this article. With the help of my supervisors, I conceptualized the research, designed the surveys, collected the data, analyzed and interpreted the data, and prepared the manuscripts.

# General Introduction[1]

Mindfulness is often defined as: “paying attention in a particular way, on purpose, in the present moment, and nonjudgmentally” (Kabat-Zinn, 1994). Academic interest in mindfulness has increased exponentially in the past two decades as can be seen in Figure 1.1 A. As of May 9 2021, the Web of Science (WOS) records 18,376 studies when using the search term “mindfulness”. Importantly, this contains a range of different approaches to mindfulness. The most common focus is on mindfulness interventions such as mindfulness-based stress reduction. A second emergent focus of mindfulness research is the investigation of mindfulness as trait or disposition. In contrast to state perspectives on mindfulness that focus on situational changes and interventions, trait perspectives examine stable between-individual differences. Some of the most commonly used scales to asses individual differences in trait mindfulness are the Five-Facet Mindfulness Questionnaire (Baer et al., 2006) and the Mindful Attention and Awareness Scale (Brown & Ryan, 2003). The trait mindfulness literature focuses on cognitive processes that can be the result of sustained mindfulness practice (Kiken et al., 2015), but also show substantial individual differences in non-practitioners (Baer et al., 2008; Pang & Ruch, 2018). In this thesis I take a more general perspective of trait mindfulness, conceptualizing mindfulness in line with Krägeloh (2020, p. 64) as “the general tendency of a person to show characteristics of nonjudgmental awareness of present-moment experience in their everyday life”. This view is reflective of the whole trait theory (Fleeson, 2001; Fleeson & Jayawickreme, 2015) which understands traits as descriptions of underlying density distributions of related states.

[1] Parts of this section have been published elsewhere: Karl, J. A., Johnson, F. N., Luisa, B., & Fischer, R. (2021). In search of mindfulness: A review and reconsideration of cultural dynamics from a cognitive perspective. Journal of the Royal Society of New Zealand. https://doi.org/10.1080/03036758.2021.1915804

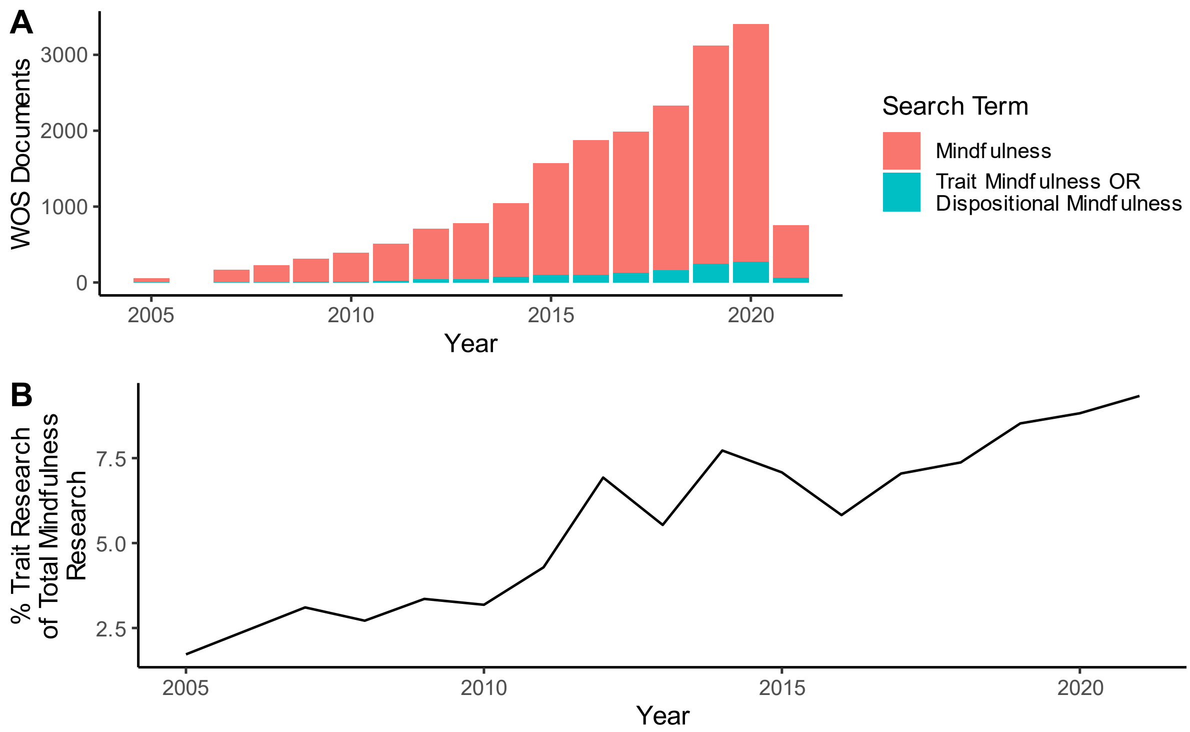



Figure 1.1 Documents indexed in the Web of Science on mindfulness (A); relative change of documents on trait or dispositional mindfulness research compared to overall research on mindfulness(B)

Importantly, trait mindfulness has shown beneficial effects for individuals in similar ways to mindfulness as a mental health or spiritual practice (Quaglia et al., 2016; Tomlinson et al., 2018). Possibly as a result of this finding, the trait approach has seen a substantial increase in interest. As can be seen in Figure 1.1B, articles including the terms “dispositional mindfulness” or “trait mindfulness” increased their share in all documents on mindfulness indexed on the Web of Science from below 3% in 2010 to around 10% in 2021. Nevertheless, as I explore in Chapter 1 of this thesis, this increase might have been lopsided, prioritizing outcomes of trait mindfulness over potential sources of variation in trait mindfulness. The current thesis focuses on addressing this relative lack of information on potential predictors of trait mindfulness. I present a series of studies which attempt to embed trait mindfulness more firmly in both macro-level processes and also in established individual differences such as personality. In the following section I provide a short historical overview of the concept of mindfulness from a Western psychological perspective and outline why personality and cultural dynamics are worth exploring.

**A short history of mindfulness from a psychological perspective**

Our contemporary understanding of mindfulness in psychology is derived from Buddhist philosophy (Baer et al., 2006). Yet, it is unclear how far the current conceptualizations strayed from their original understanding and whether currently employed mindfulness interventions may be mislabeled as Buddhist (Purser, 2019). First, we need to remind ourselves that Buddhism is not a monolithic philosophy or religion. Most writers on mindfulness over the last 150 years seem to have been influenced by Theravāda Buddhist philosophy (Gethin, 2011; Gilpin, 2008), a Buddhist tradition that can substantially differ from earlier Buddhist writings (Anālayo, 2004, 2018, 2019) and from other strands of Buddhism. Some writers have termed this school of thought 'Buddhist modernism' (McMahan, 2008; Sharf, 1995), given the widespread contemporary usage. Even within this specific type of Buddhist philosophy, it is important to acknowledge that the term "mindfulness" is not a literal translation of the original term used in Pali (*sati*), but rather represents one approximation of the meaning of *sati* (Sun, 2014). To illustrate, Rhys Davids (cited in: J. M. G. Williams & Kabat-Zinn, 2013, p. 23) in his translation of the term in 1910 wrote:

"Etymologically *Sati* is Memory. But as happened at the rise of Buddhism to so many other expressions in common use, a new connotation was then attached to the word, a connotation that gave a new meaning to it, and renders 'memory' a most inadequate and misleading translation. It became the memory, recollection, calling-to-mind, being-aware-of, certain specified facts. Of these the most important was the impermanence (the coming to be as the result of a cause, and the passing away again) of all phenomena, bodily and mental. And it included the repeated application of this awareness, to each experience of life, from the ethical point of view."

It is important to note that Rhys seems to have arrived at this translation only in 1910 and his previous translation attempts showed substantial uncertainty about the term, translating it either as "mental activity" or "thought" (Gethin, 2011). By using the translation of *sati* as mindfulness, the term became placed in and understandable from a long-standing contemplative practice.

**Western shifts to Mindfulness as 'bare attention'**

The second major event that shaped the understanding of mindfulness in Western psychology can be traced back to is to the publication of Nyanaponika Thera's book "The Heart of Buddhist Meditation" in 1954 which defined it as moment-to-moment, lucid, non-reactive, non-judgmental awareness of whatever appears to consciousness (Thera, 1998). This conceptualization together with the introduction of de-contextualized awareness practices at the same time has made mindfulness practice accessible to lay practitioners without a background in Buddhist ethics and profoundly shaped the development of later psychological and medical implementations. It is difficult to overstate the importance that this conceptualization of mindfulness had on the Western understanding of mindfulness, with most interventions and mindfulness measures being aligned with this definition of mindfulness (Curtis, 2019; Nilsson & Kazemi, 2016). The conceptual basis of current mindfulness interventions and measurement can in large parts be traced to a radical re-definition of mindfulness as bare attention, intentionally stripping away complex ethical notions to make it approachable to lay-practitioners. Could we state that "mindfulness" as currently conceptualized in psychological and clinical practice is following Buddhist principles? This is a question that might have no conclusive answer and will depend on whether one aligns with an interpretation of mindfulness as "bare attention" (Sharf, 1995). This issue is further complicated by that fact that many mindfulness-based interventions (MBI) now contain material taken from other cultural traditions (e.g., yoga, meditation). Up until now I have discussed mindfulness as a unitary construct within clinical practice. While this might be the case on a conceptual level, both in its operationalization in Western definitions where several skills are subsumed under a general factor of mindfulness (Baer et al., 2006) and psychological measures where researchers often extract a single score of mindfulness, mindfulness is likely to have different components (Blanke & Brose, 2017; Lau et al., 2006). This is not only true for mindfulness as state that is altered during practice, but also in individual differences in mindfulness that might arise from differential experiences of states (Karl & Fischer, 2021; Kiken et al., 2015).

These different components of mindfulness are most broadly captured by the Five-Factor Mindfulness Questionnaire, which is based on empirical analyses of multiple existing trait measures of mindfulness (Baer et al., 2006) and captures five facets: Acting with Awareness, Non-Judging, Non-Reacting, Describing, and Observing. These stable individual differences are thought to be reflective of individual differences in momentary, state like processes which can be captured in state measures (Blanke & Brose, 2017). These instruments were inspired by Buddhist philosophy and the five components are therefore supposed to capture central elements within this original tradition. There are also distinct efforts to define mindfulness as a Western concept, more focused on openness and being receptive to new ideas and experiences (Pirson et al., 2018). In a later chapter of this thesis I report a recently published analysis (Karl & Fischer, 2020) which demonstrated that more Buddhist inspired instruments were empirically distinct from measures that operationalize mindfulness as “Western Mindfulness” (Pirson et al., 2018). However, this study also indicated that current mindfulness measures might be affected by wording effects, so that negatively worded items are understood and responded to differently than positively worded items. Therefore, cultural worldviews and their expression in either positive or negative terms may influence how people understand and report mindfulness experiences.

**Thesis Structure**

The main body of the current thesis is clustered into three major chapters, each consisting of two studies (I show a schematic overview in Figure 1.2). Chapter one (The Current State of Mindfulness Research), presents two studies examining the current state of trait mindfulness research and measurement. The first study uses recently developed bibliometric approaches, to examine the corpus of published trait mindfulness studies. My goal is to identify research topics and gaps in this literature. In line with a close-reading of the literature, this psychometric analysis indicated relative paucity of research on cultural understandings of dispositional mindfulness and a strong focus on outcomes such as psychological well-being. Moving to the operationalization of mindfulness traits, the second study examined the applicability and replicability of the Five-Facet Mindfulness Questionnaire as a measure of mindfulness. The study overall suggests replicability of the broad five facets of mindfulness, but also an empirical separation from more recently proposed non-Buddhist inspired measures of mindfulness. I also identify potential measurement issues.

The second chapter focuses on individual-level predictors of mindfulness, one of the gaps identified in the bibliometric study. I examine the relationship between mindfulness, personality, and reinforcement sensitivity both cross-sectionally (Study 3) and longitudinally (Study 4). In my cross-sectional data I found that trait mindfulness is substantially related to both personality (especially neuroticism) and reinforcement sensitivity (especially behavioral inhibition). Nevertheless, relationships varied substantially across facets, pointing to potential differences in the underlying generative processes. In the longitudinal follow up study, I found complex relationships between the constructs, including potential reciprocal loops between mindfulness and reinforcement sensitivity. Taken together, these two studies highlight the need for further research on the relationship of trait mindfulness and established individual differences, providing an initial step to untangle their causal relationship.

The final chapter focuses on the applicability of current mindfulness measures across cultures, a second issue that I identified in the bibliometric analysis. First, I present a summary and new proto-package (Study 5) that summarizes current best practices for cross-cultural comparisons. I then apply selected methods in a study using secondary data sources from 16 countries (Study 6). I examined the cross-cultural applicability of the FFMQ and explored potential sources of variation across these cultures. I found substantial variation across cultures in the conceptualization of mindfulness which can be partially explained by underlying cultural dimensions such as individualism-collectivism. I finish my thesis with a general discussion, laying out avenues to advance the research on trait mindfulness from both an individual difference and cultural perspective.

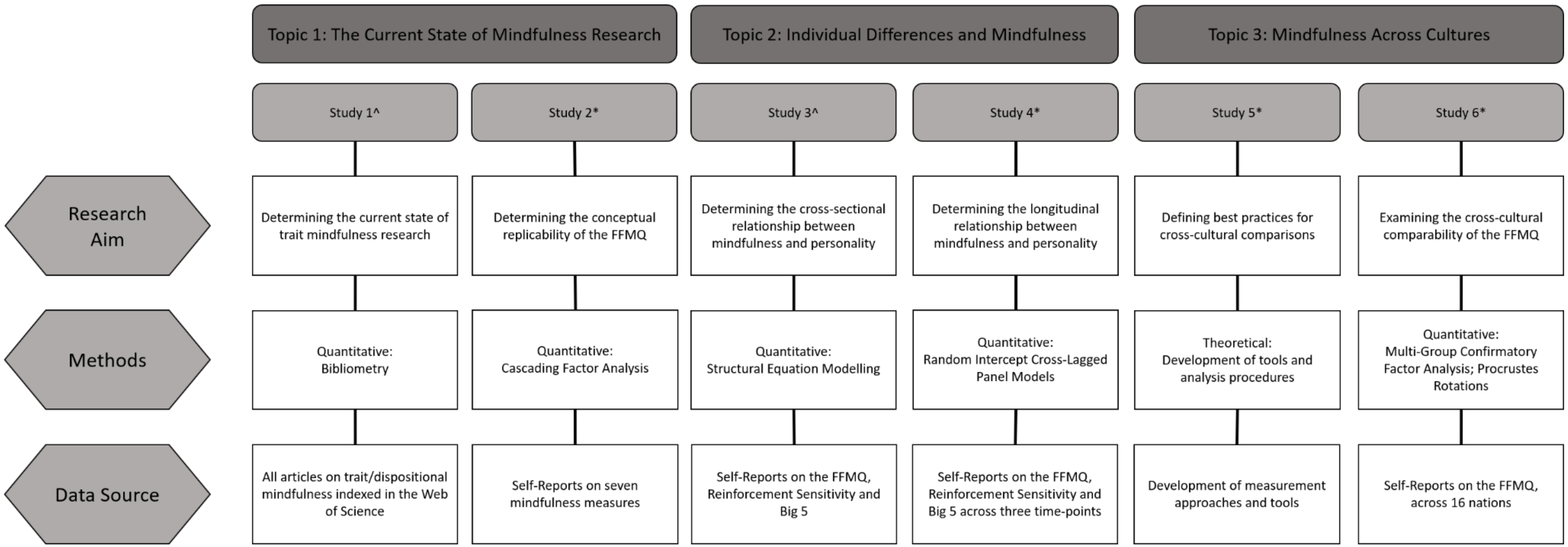



*Figure 1.2* Structure of the current thesis.

*Note.* For brevity I refer to **trait** mindfulness as mindfulness. Studies marked ^ are publicly available as pre-prints, studies marked * have been published in peer-reviewed journals. FFMQ – Five Factor Mindfulness Questionnaire

# Chapter 1 The Current State of Dispositional Mindfulness Research

# Study 1 Trait Mindfulness in Psychological Research: A Bibliometric Review[2]

## Preface

Trait mindfulness is a steadily growing field with diverse research foci and directions. In study one I aimed to provide a first high-level overview of the current state of the literature. With the bibliometric approach exemplified in this study I opted for a computerized bottom-up approach rather than a qualitative systematic review. This bottom-up approach allowed me to utilize the whole available corpus on the Web of Science which focused on trait mindfulness, identifying broad trends and topics. Regarding the literature on trait mindfulness the aim of this study is to stimulate a debate and interconnection of different research traditions, both conceptually and across cultures. In the context of the current thesis, this study provides an empirical groundwork for the gaps in the literature addressed in the individual studies of this thesis. To address the identified need for greater empirical evidence on potential predictors of trait mindfulness, I conducted study three and four which focused on personality and reinforcement sensitivity as predictors of trait mindfulness. To address the need for extended cultural perspectives on trait mindfulness identified in this study, I provide theoretical approaches in study five and present a first large scale investigation into the cross-cultural applicability of the FFMQ in study six.

[2] This study has been submitted to *Mindfulness* for review. Minor revisions and stylistic changes have been made to the manuscript to establish coherence with the rest of the thesis.

Mindfulness is a concept that originally formed part of a wider philosophical and spiritual belief system associated with Buddhism. Core ideas inspired and derived from this Buddhist philosophy were imported into Western psychology and medical practice during the second half of the 20th century (Baer et al., 2006). While mindfulness had been discussed for the better part of the last century by academics interested in Buddhism, discussions of mindfulness in Western psychology got traction with the publication of Nyanaponika Thera's book "The Heart of Buddhist Meditation" in 1954, which defined it as moment-to-moment, lucid, non-reactive, non-judgmental awareness of whatever appears to consciousness (Thera, 1998). Based on these conceptualizations of mindfulness, the practice was introduced into the psychological/medical community as stress reducing intervention, most notably exemplified by Kabat-Zinn's mindfulness-based stress reduction (MBSR: Kabat-Zinn, 2011). These interventions allowed non-Buddhists to engage in mindfulness practices without the need to endorse or engage in the wider Buddhist ethical belief system (Curtis, 2019; for a review highlighting the absence of ethical components in current definitions see: Nilsson & Kazemi, 2016). Mindfulness-based interventions have been widely incorporated into clinical practice (Karl, Johnson, et al., 2021) and have been shown to be effective in addressing a wide range of mental health issues (Fischer et al., 2020). This ease of implementation and efficacy in addressing mental health issues prominent in Western societies (especially anxiety and depression) has led to a substantial growth in the field (Creswell, 2017; Van Dam et al., 2018).

Building on this research on mindfulness interventions a separate body of literature has emerged that focuses on stable, trait-like individual differences in mindfulness, i.e., reflecting one's general propensity to be mindful in daily life (Siegling & Petrides, 2014). While trait mindfulness has been shown to be influenced by consistent mindfulness practices, it can also be present in non-practitioners (Baer et al., 2008). Importantly, meta-analytic evidence indicates that increased trait mindfulness has a host of beneficial outcomes ranging from positive work outcomes to psychological well-being (Carpenter et al., 2019; Karyadi et al., 2014; Mesmer-Magnus et al., 2017; Sala et al., 2020). This focus on individual differences created a second and rather distinct research line in addition to clinical intervention studies, focusing on trait mindfulness that aim to identify outcomes, predictors, and underlying mechanisms of such stable interindividual differences. Research on trait mindfulness opens up new avenues to understand mindfulness from neuroscientific, biological, and individual difference perspectives, as well as providing opportunities for identifying possible cultural differences. The last decades have been marked by a diversification and broadening of this new body of inquiry beyond clinical interventions. A number of good overviews over specific topics in the field of mindfulness are available already (Baer et al., 2006; Chiesa & Malinowski, 2011; Keng et al., 2011). Given the diversity of the theories and approaches, researchers have started to use meta-science approaches such as bibliometry and topic modelling to generate overviews of the field of mindfulness (Karl, Johnson, et al., 2021; Kee et al., 2019), and to identify the relationships of empirical research with Buddhist theoretical foundations (Valerio, 2016). Nevertheless, currently no high-level overview of the field of trait mindfulness is available. The absence of authoritative overviews of research trends makes it difficult to track the development and current state of this specific subfield. This is particularly important because the study of trait mindfulness allows building new bridges to different areas of psychology and clinical practices.

In this review we aim to provide a first systematic meta-science documentation of the trait mindfulness research field using a bibliometric approach. This approach represents an empirical meta-review of the existing literature, rather than a close-reading approach commonly adopted in reviews (Aria & Cuccurullo, 2017). Bibliometrics is a branch of information science aimed at analyzing bibliographic sources (such as books and journal articles) using a quantitative approach. While the exact definition of the term bibliometrics is subject of debate it can be broadly defined as: “Bibliometrics is the quantitative study of literatures as they are reflected in bibliographies. Its task, immodestly enough, is to provide evolutionary models of science, technology, and scholarship.” (White & McCain, 1989, p. 119). While still an emergent technique in psychological research, bibliometric methods have long been used in natural sciences and business research to document the structure and development of research topics. At a certain maturity of a research field the corpus reaches a volume that makes a complete review using manual approaches untenable. For example, in 2017 a search for the term “mindfulness’ yielded 5947 documents (Kee et al., 2019). Consequently, researchers conducting a review of a field are forced to narrow the scope of their investigation, potentially missing valuable interconnections between topics. In contrast to traditional reviews that rely on the selection of key works by the authors of the review, bibliometric approaches utilize the whole corpus identified in a search as data source. From this corpus, researchers can first extract descriptive features such as year of publication, authors, and country of origin. This allows for insight into country collaborative networks, important outlets, and influential authors. Beyond these descriptives of a field, researchers have also begun to use bibliometric methods combined with advanced statistical modelling approaches to identify clustering of research topics within a field and historical development of research topics in the field over time. The bibliometric approach allows for such a high-level of synthesis of the material present in a field, examining trends and connections in research foci, which would be difficult to complete using a more traditional review approach.

In the current study we aim to advance four major goals: First, we map out the research-space around trait mindfulness and take stock of current research fields, important publications, and authors, identifying broad research trends. This allows us to get a sense of the general impact of the field, and helps to identify key authors and relevant publication outlets. The presentation of this data can help novices in this field to identify key resources and authors to read and identify possible outlets for publications. For more seasoned researchers, it helps them to get a sense of the larger impact that specific publications and authors may have had.

Second, what are central themes of research across the corpus and how have they evolved over time? This is of great relevance for understanding the historical development of a research field and examining possible trends, opportunities, and research gaps for further development. With the increasing specialization of researchers and the institutional creation of academic silos, it is important to step back and examine broader trends. How narrowly or broadly defined are research topics and how central are individual research foci within the larger research community? How are studies on dispositional mindfulness connected to other work on individual differences – what broad research trends can be identified that link individual differences in mindfulness to other theoretical frameworks to identify possible origins and predictors of such dispositional differences. From a meta-science perspective, how do research topics change over time – do they become more narrowly focused on specific topics? What are the core themes over time? What is the impact of broader vs more narrowly defined research topics? The renewed calls for interdisciplinarity and transdisciplinary research in science in general (Bernstein, 2015) and health research in particular (Gehlert et al., 2010) require better understanding of larger underlying research trends and research communities.

Third, how is research on trait mindfulness distributed globally? Given that the origins of mindfulness lie in Buddhist thought it is interesting and theoretically important to examine whether the field is characterized by a more diverse authorship compared to the wider field of psychology. How international is the field of trait mindfulness? What is the representativeness of contemporary trait mindfulness research and how well does it capture cultural variations that are important to track (Henrich, 2020; Henrich et al., 2010). Finally, and related to the previous point, given that mindfulness has originated in an explicit Buddhist context, it raises the question whether the research foci differ between countries and world regions with a Buddhist vs non-Buddhist history. These last two questions are highly exploratory and driven by the recognition that contemporary research is likely to be biased to specific paradigms and models (Hendriks et al., 2019).

## Methods

### Data Source

We used a broad search strategy using the Web of Science (WOS). In order to identify the maximum possible records, we used the search strings: “Dispositional Mindfulness” OR “Trait

Mindfulness". We initially started our search from 1970 to the present day but found the first explicit mention of either search term in 2005. We therefore restricted the year range to 2005-2021. We downloaded all articles as bibtex files, including all available information such as keywords, abstracts, and authorship information. We combined all files into one master-database representing the full corpus and transformed the files into processable dataframes using the bibliometrix package (Aria & Cuccurullo, 2017) in R (R Core Team, 2020). All data and analysis code can be found on the OSF (https://osf.io/84m95/?view_only=a565011c959648a2a3cdccd5486d100d)

# Results

## Descriptives

Research on individual differences in dispositional mindfulness has increased substantially over the last two decades (Figure 1 in Appendix A) with an average growth rate of 17.39% per year. Overall, we found 1229 documents (1022 articles, 64, early access articles, 5 proceeding papers, 4 editorials, 1 letter, 65 meeting abstracts, 7 proceeding papers, 58 reviews, 1 book chapter review, 2 early access reviews) in 442 unique sources (Journals, Books, etc.). Table 2.1 shows the 10 most common sources, which unsurprisingly were headed by the Journal *Mindfulness.* Importantly, the second most important outlet for trait mindfulness research is *Personality and Individual Differences*. Most articles were authored by multiple authors with an average of 2.81 authors per document. To examine important papers in the global citation network (all articles available on the WoS) and our slice of the citation network (studies within our corpus of articles), we extracted the 10 most important papers based on their global and local citation score. We found a substantial correlation of number of citations of a document within our network with its overall citations (indicating that our network captures a representative slice of the overall citation network). At the same time, when using rank correlations, the relative order changed which indicates differential importance of papers in our network compared to broader research trends. Focusing on specific features, the 10 most important papers in the general citation network focused largely on scale development and conceptual definitions (Table 2.2) and were exclusively written by North American-based first authors. In contrast, papers in our local network focused largely on clinically relevant outcomes, such as well-being or psychological issues. Taken together this indicates two issues: first, the primacy of North American measures and the importance of the conceptualizations of mindfulness in the literature which are widely influential in the wider literature; and second the overall focus of the trait mindfulness field on outcomes such as well-being. Work on possible mechanisms of trait mindfulness and predictors of individual differences in mindfulness take a second seat to possible health and clinical outcomes. This suggests that there is space for greater exploration of underlying mechanisms and predictors that explain these dispositional differences.

**Identifying research themes and trends**

To clarify the concepts researched jointly with dispositional mindfulness we extracted the 20 most common keywords applied by authors to their articles (Table 2.3). To provide a better high-level overview over author keywords, we applied Multiple-Correspondence Analysis based on a co-occurrence matrix of keywords. Importantly, this analysis tried to extract clusters of keywords, representing archetypical lenses of examination of trait mindfulness. Individual documents can contain keywords from multiple clusters. As this method yields exploratory results that are dependent on the included variables, we decided to implement three different cut-offs for minimal degrees of keyword associations included in the analysis. To capture the broad range of topics addressed by authors we decided to cut at 5, 25 (Shown in Figure 2 in Appendix A), and 50 minimum degrees (Figure 2.1). Cutting at a minimum of 5 degrees we found a split in the keywords separating out clusters of general research on trait mindfulness, focusing on topics such as anxiety and depression (Frewen et al., 2008) or meditation (Creswell et al., 2007), and cognitive-neuroscience research on mindfulness, focusing on terms such as amygdala (Way et al., 2010) or default mode network (Wang et al., 2014). Cutting at 25 degrees, a cluster split from the main cluster encapsulating discussion on different therapeutic approaches such as mindfulness-based cognitive therapy (Farb et al., 2013). Finally, cutting at 50 degrees the main cluster separated into a general cluster containing terms such as stress (Carmody & Baer, 2008) or health (Creswell & Lindsay, 2014), a cluster focused on anxiety and depression (Frewen et al., 2008), and one research cluster on emotion regulation (Creswell et al., 2007).

**Global changes in trait mindfulness research**

To examine the change in the research of mindfulness we examined the evolution of themes across the last decade of research. We split the dataset into two blocks running from 2005 to 2010 and a second block from 2011 to 2021. We extracted all publisher keywords, to capture higher order themes, that occurred at least 30 times (we show all themes in Table 2.4 and in Figures 2.2a-b). Following the approach proposed by Cobo et al. (2011), we divided the research space into four quadrants based on a theme's centrality (capturing a theme's connection with other themes, with greater connection implying greater embedding in the thematic field) and density (capturing the strength of the interconnection of key-terms within a theme, expressing a field's development in terms of internal coherence) in the overall network. This gives rise to a four-quadrant system of 1) Niche Themes (well-developed internal ties but unimportant external ties), 2) Motor Themes (well developed and important for the structuring of a research field), 3) Basic Themes (important for a research field but are not developed), and 4) Emerging/Declining Themes (weakly developed and marginal). In the network structure covering 2005-2010 we found three niche themes: reduction (key terms: reduction, generalized anxiety disorder; focusing on reducing general anxiety disorder); intervention (key terms: intervention, information; focusing on general interventions); and anxiety disorders (key terms: anxiety, disorders, follow-up; focusing on longitudinal research on anxiety). Personality was the sole motor theme, containing terms such as personality, consciousness, and model. We found four motor themes within this specific thematic field: 1) inventory (key terms: inventory , 5-factor model, representing psychometric properties of the FFMQ), 2) cognitive therapy (key terms: cognitive therapy, major depression, stress reduction; representing CBT and clinical depression|stress), 3) meditation (key terms: meditation, awareness, amygdala, prefrontal cortex; focusing on meditation, including neuroscience topics), and 4) prevention (key terms: prevention, relapse, depression, rumination; focusing on prevention of psychological illnesses and substance abuse).

Examining the thematic networks during the period of 2011-2021, we found a lower number of clusters compared to the previous timeframe (indicating a consolidation of research topics into larger connected areas that are internally less coherent). This was reflected in the clear presence of four basic themes: 1) dispositional mindfulness (key terms: dispositional mindfulness, emotion regulation, stress, health ; capturing research on dispositional mindfulness and well-being), 2) stress reduction (key terms: stress reduction, benefits, behavior, personality; focusing on research on stress reduction and personality), 3) questionnaire (key terms: questionnaire, validation, self-report, psychometric properties; focusing on psychometric validations and developments), and 4) depression (key terms: depression, anxiety, intervention, cognitive therapy; focusing on clinical interventions and randomized trials). We further found one theme bordering the emergent/basic quadrant focusing on meditation and cognitive processes (key terms: meditation, attention, mechanisms, performance, awareness). Last, we found one theme bordering motor/basic quadrants, focusing on dispositional mindfulness and depressive symptoms (key terms: mindfulness, individual-differences, dispositional, depressive symptoms). We show the change of terms between categories together with the overlap of categories in Table 2.5 (visualized in Figure 2.3). Interestingly, we see on the one hand consolidation with research topics focusing on anxiety, and therapeutic approaches to depression merging into one theme, but on the other hand we see an emergence of a theme focusing on stress reduction, subsuming aspects of individual difference research such as personality. Similarly, the meditation theme split and the core theme became more refined, retaining keywords focusing on awareness and attention.

**Global distribution of trait mindfulness research**

Looking at the geographic distribution of first authors’ institutions, we found that the publications on dispositional mindfulness were substantially biased towards Europe, Australia, and North America (Figure 2.4, Table 1 in Appendix A). The USA was the most productive country, accounting for 43.16% of all published documents, followed by China (9.12%), Canada (7.85%), Australia (6.33%), and the UK (5.41%). Importantly the USA also had the lowest rate of multi-country studies (9%), indicating that the majority of scientific output on mindfulness focuses on USA specific samples and issues. Interestingly, China had the second highest output of published documents on this topic of trait mindfulness (9.12%), but also showed a relatively high percentage (29.60%) of multi-country collaborations. To clarify the relationship between countries, we examined the collaboration network between countries based on co-authorships (Figure 2.5). Overall, we found that the nodes with the highest strength were the USA (160), the UK (81), China (59), Netherland (49), and Australia (38), indicating that most cross-country collaborations included authors from these countries.

**Cultural differences in mindfulness research**

As China was the only non-American/non-European country among the top 10 countries, we compared the keywords applied by authors in China to keywords applied by US authors. Using only keywords that were present in both samples we found a high correlation in usage frequency: r(114) = .93, $p < .001$. To examine the correlation of the relative importance of keywords across countries we transformed the frequencies within countries into rank-orders with ties broken at random (to increase the robustness, we bootstrapped the analysis 1000 times). Overall, we found a low correlation between ranks across countries: $r(114) = .204[.202, .207]$, $p < .043[.041, .046]$. Taken together this indicates that the terms that can be matched are of similar absolute importance, but their relative importance within samples might differ in the two countries.

Comparing the most frequently used keywords in the USA and China, we find similar patterns with sample descriptors and specific indicators of ill-being such as Anxiety and Depression, but also marked differences such as Substance Abuse taking a higher place in the USA. The potentially most striking difference is 1) the absence of *Meditation* in the Chinese sub-network, which ranks relatively highly in the USA network and 2) the strong presence of statistical features such as *Mediation* in Chinese articles. One potential reason for this could be found in the apparent difference in outcomes. Publications by first authors based in the USA center largely around clinically relevant variables, which are also often the main targets of meditation interventions. In contrast, Chinese output seem to be more focused on a better understanding of abstract emotion regulation processes. It is also interesting to note that in Chinese outputs, greater distinctions are made between Mindfulness, Dispositional Mindfulness, and Trait Mindfulness. Overall, this indicates that the research focus between the largest non-western and western producer on this topic seem to differ, with the US potentially prioritizing mental health related research, whereas Chinese-based researcher may focus relatively more on processes.

We examined the country differences further by examining the 20 most cited papers in the reference section in each country's corpus. This helps us to understand whether USA and Chinese-based first authors rely on different sources for developing their research. We show the results in Table 2.6 (we also list the top cited papers in the Chinese corpus in Table 2.7, indicating the papers by China-based first authors that had the largest impact on the field). Overall, we found an overlap of 50% in the top cited documents, suggesting that the research in both countries draws on somewhat similar sources. Documents of unique high importance in China focused around adaptation of scales (Deng et al., 2012), methodological and statistical questions (Hayes, 2013; Hu & Bentler, 1999; Podsakoff et al., 2003), well-being (Coffey & Hartman, 2008; Weinstein et al., 2009), social anxiety (Goldin & Gross, 2010; Rasmussen & Pidgeon, 2011), and mindfulness theory (Garland, Farb, et al., 2015). Documents of unique high importance in the US were focused on mindfulness scales (Baer et al., 2004), cognitive neuroscience (Brown et al., 2012; Creswell et al., 2007), mechanisms of meditation (Hölzel et al., 2011), romantic relationships (Barnes et al., 2007), and well-being (Carmody & Baer, 2008; Grossman et al., 2004; Kabat-Zinn, 1982, 1990; Segal et al., 2002). Further, we examined the top five articles that were not shared between the countries for each country. In the US, these were mindfulness scale psychometrics (Bohlmeijer et al., 2011), substance abuse (Bowen et al., 2009, 2014; Fernandez et al., 2010), mindfulness and culture (Grossman & Dam, 2011). In China, the articles were focused on mindfulness scale psychometrics (S.

Chen et al., 2012), PTSD in children (Foa et al., 2001), well-being (Creswell, 2015), natural disasters (indicating a specific applied focus: Lyu et al., 2017), and health measures (Yang et al., 2003).

## Discussion

Our current study aimed to provide a mapping of the research spaces investigating trait mindfulness. Our main findings reveal two important considerations about the current research on trait mindfulness, one being the unequal distribution of mindfulness research globally and possible implications for our understanding of trait mindfulness, the other is the focus on clinical and health outcomes.

### Global distribution of mindfulness research

First, resembling psychology as a wider field (Henrich, 2020), we found a dominance of US and European researchers in the field of mindfulness. Especially US-based first authors showed a low likelihood to collaborate with other colleagues internationally. If they did, they enjoyed a high centrality in the collaborators network, indicating that a substantial body of work on mindfulness is exclusively focused on US samples and furthermore, research on trait mindfulness in other countries often includes US perspectives. Given the historical origin of mindfulness as a Buddhist philosophical construct, this raises questions and possible challenges about the current conceptualization and authenticity of the construct of mindfulness, which already has received some discussion (Grossman & Dam, 2011). An encouraging trend is that this is being recognized as seen by the central position of this paper in studies being published by US-based authors. Furthermore, there is increasing awareness that individual difference measures aimed at capturing trait mindfulness such as the FFMQ may perform sub optimally in non-WEIRD populations (Christopher, Charoensuk, et al., 2009; Karl et al., 2020), which in turn has led to the development of alternative and more culturally aligned measures by researchers (Ng & Wang, 2021). If research on trait mindfulness fails to incorporate more diverse non-WEIRD perspectives this might not only result in operational definitions of trait mindfulness that are not universally accessible, but also may fail to meet the different needs of populations around the globe. This becomes apparent looking at the different use of keywords in the two biggest producers of research, one based in the Global West (USA) and the other based in the Global East (China).

While in both countries researchers focused on well-being outcomes, the priority was markedly different. Whereas in the USA substance abuse was a major target of research, this was absent in China when examining the major research trends as indicated by keywords. Additionally, while meditation was an important research topic connected to trait mindfulness in the US, it was again absent in China. Interestingly, a common term in both countries was *Mediation,* indicating that a substantial portion of research on trait mindfulness does not research direct relationships, but rather tests more complex path models (for the most highly cited examples in the current set see: Demarzo et al., 2014; Iani et al., 2017; Nitzan-Assayag et al., 2015). When examining the unique high impact citations driving research in China and the US, we found a reflection of this pattern with a substantial number of references in the US focusing on meditation or mindfulness practice (Carmody & Baer, 2008; Grossman et al., 2004; Hölzel et al., 2011), whereas uniquely important references in China focused on methodological concerns (Deng et al., 2012; Hayes, 2013; Hu & Bentler, 1999; Podsakoff et al., 2003).

**Mindfulness research themes and trends**

Second, while examining key-terms and their development we found an increasing consolidation of trait-mindfulness research into distinct clusters and a strong focus on outcomes compared to predictors. Examining the thematic maps of overall keywords related to trait mindfulness, we found that a major split in the research field exists between cognitive/neuroscience investigations into trait mindfulness (for example: Wang et al., 2014; Way et al., 2010) and a more diverse field containing personality, clinical, and positive psychology. Narrowing down further this central cluster separated into two clusters, the emergent cluster captured different clinical approaches such as CBT, or general research on emotion regulation. Overall, breaking apart the keywords used in conjunction with trait mindfulness reveal that the field is mostly split into neuroscience/cognitive research and outcome-focused clinical research.

Zooming out to the broader trends in the literature, the result of the keyword analysis is mirrored in the development of the themes overall across the time period. Research on mindfulness has consolidated into distinct subfields over the last two decades. Research topics such as personality and individual differences, over time, became less of an individual focus and merged with the wider literature on stress reduction. Interestingly, a distinct subfield has emerged that focused on psychometric approaches to mindfulness, indicating the increasing emphasis in the field to consolidate and validate measurements of mindfulness (Andrei et al., 2016; Karl et al., 2020; Karl & Fischer, 2020; Siegling & Petrides, 2014, 2016). Overall, research on trait mindfulness has consolidated around psychometric issues and outcome focused topics such as stress, well-being, and clinical interventions. This reveals a potential imbalance within the field, with increasing focus on outcomes, while less research is conducted on potential predictors of mindfulness. To achieve a fuller understanding of dispositional mindfulness, it is essential to address potential predictors given the complex causal interplay between mindfulness and established individual differences such as personality (Karl, Fischer, et al., 2021) and situational variables such as affect (Karl & Fischer, 2021; Mahlo & Windsor, 2021).

## Limitations

One major limitation of our current work lies in the database (Web of Science) used. Our current source might miss papers that are not indicated in the WoS or not formally published (so-called grey literature). Our search also relied on author (or publisher) assigned key terms to identify articles of interest. Additionally, our use of English language search terms leaves open the question of how terms used in other languages map onto our selected key-terms. Last, given the substantial body of literature resulting from our search we focus on broad trends that do not allow for a narrative review or the identification of more qualitative and nuanced trends of the research field.

We provide our full database on the Open Science Framework to allow interested researchers to explore more narrow sub-topics.

## Conclusions and further research

Our current research presents a first high-level overview over the topic of trait or dispositional mindfulness. Overall, our research indicates that the field is maturing and quite distinct areas focusing on cognitive attentional processes and clinical interventions have emerged, with a strong focus of the field on outcomes of mindfulness, including both applied and basic attentional processes. In contrast, potential predictors of trait mindfulness, such as cultural and individual differences are less developed in recent thematic networks. The increasing interest in measurement and validity of current mindfulness constructs (manifested in the emergent themes around scale validity) might present an opportunity to more closely examine the nomological network of mindfulness and individual differences, as well as cultural differences in mindfulness.

*Table 2.1 Top 10 Research outlets*

| Sources | Documents |
|---|---|
| Mindfulness | 231 |
| Personality and Individual Differences | 86 |
| Frontiers in Psychology | 43 |
| Plos One | 18 |
| Psychosomatic Medicine | 16 |
| Annals of Behavioral Medicine | 13 |
| Social Cognitive and Affective Neuroscience | 13 |
| Addictive Behaviors | 11 |
| International Journal of Environmental Research and Public Health | 11 |
| Journal of American College Health | 11 |

*Table 2.2 Top cited documents in the global and local network.*

| First Author | Year | DOI/ISBN | WOS Citations/Local Citations | Topic | ISO3 |
|---|---|---|---|---|---|
| Top Documents by WOS citations | | | | | |
| Brown KW | 2003 | 10.1037/0022-3514.84.4.822 | 776 | Scale Development | USA |
| Baer RA | 2006 | 10.1177/1073191105283504 | 630 | Scale Development | USA |
| Bishop SR | 2004 | 10.1093/CLIPSY/BPH077 | 413 | Mindfulness Measurement Conceptualization | CAN |
| Brown KW | 2007 | 10.1080/10478400701598298 | 278 | Mindfulness and Well-Being | USA |
| Kabat-Zinn J | 2003 | 10.1093/CLIPSY/BPG016 | 268 | Mindfulness in Clinical Contexts | USA |
| Kabat-Zinn J. | 1990 | 385298978 | 252 | Mindfulness Theory | USA |
| Baer RA | 2008 | 10.1177/1073191107313003 | 251 | Scale Validation | USA |
| Kabat-Zinn J. | 1994 | 9781400000000 | 224 | Mindfulness Theory | USA |
| Baer RA | 2003 | 10.1093/CLIPSY/BPG015 | 213 | Mindfulness and Well-Being | USA |
| Keng SL | 2011 | 10.1016/J.CPR.2011.04.006 | 196 | Mindfulness and Well-Being | USA |
| Top Documents in the Local Network | | | | | |
| Creswell JD | 2007 | 10.1097/PSY.0b013e3180f6171f | 414/140 | Affective Labelling | USA |
| Barnes S | 2007 | 10.1111/j.1752-0606.2007.00033.x | 289/106 | Relationship Well-Being | USA |
| Shapiro SL | 2011 | 10.1002/jclp.20761 | 173/70 | Stress-Reduction | USA |
| Tomlinson ER | 2018 | 10.1007/s12671-017-0762-6 | 83/62 | Well-Being | GBR |
| Way BM | 2010 | 10.1037/a0018312 | 104/57 | Depression | USA |
| Hulsheger UR | 2013 | 10.1037/a0031313 | 405/56 | Well-Being | NLD |
| Heppner WL | 2008 | 10.1002/ab.20258 | 135/45 | Aggression | USA |
| Mrazek MD | 2012 | 10.1037/a0026678 | 227/42 | Construct Validation | USA |
| Arch JI | 2010 | 10.1016/j.brat.2010.02.005 | 88/41 | Anxiety | USA |
| Eisenlohr-Moul TA | 2012 | 10.1177/1073191112446658 | 70/41 | Substance Use | USA |

*Table 2.3* Top 20 keywords in the general corpus, the USA, and China

| General | General Frequency | USA Keywords | USA Frequency | China Keywords | China Frequency |
|---|---|---|---|---|---|
| Mindfulness | 827 | Mindfulness | 371 | Mindfulness | 59 |
| Dispositional Mindfulness | 100 | Stress | 39 | Dispositional Mindfulness | 22 |
| Depression | 94 | Anxiety | 34 | Trait Mindfulness | 10 |
| Stress | 88 | Depression | 33 | Emotion Regulation | 9 |
| Anxiety | 85 | Emotion Regulation | 31 | Adolescents | 8 |
| Emotion Regulation | 79 | Meditation | 31 | Mediation | 8 |
| Meditation | 69 | Substance Use | 23 | Anxiety | 6 |
| Trait Mindfulness | 53 | Trait Mindfulness | 22 | Depression | 6 |
| Well-Being | 43 | Dispositional Mindfulness | 20 | Mediating Effect | 5 |
| Attention | 40 | Attention | 18 | Mental Health | 5 |
| Adolescents | 35 | College Students | 17 | Perceived Stress | 5 |
| Rumination | 35 | Well-Being | 17 | Rumination | 5 |
| Self-Compassion | 31 | Coping | 12 | Ambulatory Assessment | 4 |
| Mediation | 30 | Emotion | 12 | Life Satisfaction | 4 |
| Emotion | 29 | Cortisol | 11 | Psychological Distress | 4 |
| Mental Health | 29 | Mediation | 11 | Sleep Quality | 4 |
| Acceptance | 27 | Adolescence | 10 | Stress | 4 |
| Substance Use | 27 | Acceptance | 9 | Adolescent | 3 |
| Adolescence | 21 | Adolescents | 9 | Firefighters | 3 |
| Depressive Symptoms | 20 | Aggression | 9 | Growth | 3 |

*Table 2.4 Themes extracted from the dispositional mindfulness literature 2005-2010|2011-2021*

| 2005-2010 | |
|---|---|
| Label | Terms |
| inventory | inventory, 5-factor model |
| personality | personality, consciousness, model, esteem, individual-differences |
| prevention | prevention, depression, relapse, experiential avoidance, rumination, symptoms, parasuicide, therapy |
| reduction | reduction, generalized anxiety disorder |
| cognitive therapy | cognitive therapy, major depression, stress reduction |
| anxiety | anxiety, disorders, follow-up |
| intervention | intervention, information |
| meditation | meditation, self-report, awareness, metaanalysis, validation, amygdala, attention, dispositional mindfulness, prefrontal cortex, responses |
| 2011-2021 | |
| stress reduction | stress reduction, benefits, behavior, personality, trait mindfulness, satisfaction |
| mindfulness | mindfulness, individual-differences, dispositional, depressive symptoms |
| depression | depression, anxiety, intervention, cognitive therapy, therapy, symptoms, rumination, facets, quality-of-life, randomized controlled-trial |
| meditation | meditation, attention, mechanisms, performance, awareness |
| questionnaire | questionnaire, validation, self-report, psychometric properties, interventions, scale, model, acceptance, validity, metaanalysis |
| dispositional mindfulness | dispositional mindfulness, emotion regulation, stress, health, reduction, responses, mental-health, self-compassion, life, college-students |

*Table 2.5 Keywords changing cluster between time-blocks*

| Themes 2005-2010 | Themes 2011-2021 | Words | Inclusion Index |
|---|---|---|---|
| anxiety--2005-2010 | depression--2011-2021 | anxiety | 0.33 |
| cognitive therapy--2005-2010 | depression--2011-2021 | cognitive therapy | 0.33 |
| cognitive therapy--2005-2010 | stress reduction--2011-2021 | stress reduction | 0.33 |
| intervention--2005-2010 | depression--2011-2021 | intervention | 0.50 |
| inventory--2005-2010 | questionnaire--2011-2021 | inventory | 0.50 |
| meditation--2005-2010 | dispositional mindfulness--2011-2021 | dispositional mindfulness; responses | 0.10 |
| meditation--2005-2010 | meditation--2011-2021 | meditation; awareness; attention | 0.20 |
| meditation--2005-2010 | questionnaire--2011-2021 | self-report; metaanalysis; validation; scale | 0.09 |
| personality--2005-2010 | mindfulness--2011-2021 | individual-differences | 0.25 |
| personality--2005-2010 | questionnaire--2011-2021 | Model | 0.20 |
| personality--2005-2010 | stress reduction--2011-2021 | Personality | 0.20 |
| prevention--2005-2010 | depression--2011-2021 | prevention; depression; rumination; symptoms; therapy | 0.13 |
| reduction--2005-2010 | dispositional mindfulness--2011-2021 | reduction | 0.50 |

*Table 2.6 Top 20 cited papers by US and China-based first authors*

| US | | | | China | | | |
|---|---|---|---|---|---|---|---|
| Author | Year | DOI | Freq | Author | Year | DOI | Freq |
| BROWN KW | 2003 | 10.1037/0022-3514.84.4.822 | 325 | BROWN KW | 2003 | 10.1037/0022-3514.84.4.822 | 83 |
| BAER RA | 2006 | 10.1177/1073191105283504 | 286 | BISHOP SR | 2004 | 10.1093/CLIPSY/BPH077 | 38 |
| BISHOP SR | 2004 | 10.1093/CLIPSY/BPH077 | 177 | **DENG YQ** | 2012 | 10.1007/S12671-011-0074-1 | 34 |
| **KABAT-ZINN J.** | 1990 | | 121 | KABAT-ZINN J | 2003 | 10.1093/CLIPSY/BPG016 | 34 |
| BROWN KW | 2007 | 10.1080/10478400701598298 | 111 | BAER RA | 2006 | 10.1177/1073191105283504 | 32 |
| BAER RA | 2008 | 10.1177/1073191107313003 | 98 | BROWN KW | 2007 | 10.1080/10478400701598298 | 29 |
| KABAT-ZINN J | 2003 | 10.1093/CLIPSY/BPG016 | 98 | SHAPIRO SL | 2006 | 10.1002/JCLP.20237 | 26 |
| KABAT-ZINN J. | 1994 | | 98 | **PODSAKOFF PM** | 2003 | 10.1037/0021-9010.88.5.879 | 23 |
| BAER RA | 2003 | 10.1093/CLIPSY/BPG015 | 86 | KENG SL | 2011 | 10.1016/J.CPR.2011.04.006 | 19 |
| **BAER RA** | 2004 | 10.1177/1073191104268029 | 81 | KABAT-ZINN J. | 1994 | | 18 |
| KENG SL | 2011 | 10.1016/J.CPR.2011.04.006 | 78 | HOFMANN SG | 2010 | 10.1037/A0018555 | 17 |
| HOFMANN SG | 2010 | 10.1037/A0018555 | 71 | BAER RA | 2003 | 10.1093/CLIPSY/BPG015 | 16 |
| **CRESWELL JD** | 2007 | 10.1097/PSY.0B013E3180F6171F | 66 | BAER RA | 2008 | 10.1177/1073191107313003 | 16 |
| **HOLZEL BK** | 2011 | 10.1177/1745691611419671 | 66 | **HAYES A. F.** | 2013 | | 16 |
| SHAPIRO SL | 2006 | 10.1002/JCLP.20237 | 60 | **HU LT** | 1999 | 10.1080/10705519909540118 | 16 |
| **BARNES S** | 2007 | 10.1111/J.1752-0606.2007.00033.X | 59 | **WEINSTEIN N** | 2009 | 10.1016/J.JRP.2008.12.008 | 16 |
| **CARMODY J** | 2008 | 10.1007/S10865-007-9130-7 | 55 | **GARLAND EL** | 2015 | 10.1080/1047840X.2015.1092493 | 15 |
| **GROSSMAN P** | 2004 | 10.1016/S0022-3999(03)00573-7 | 55 | **COFFEY KA** | 2008 | 10.1177/1533210108316307 | 12 |
| **BROWN KW** | 2012 | 10.1016/J.PSYNEUEN.2012.04.003 | 51 | **GOLDIN PR** | 2010 | 10.1037/A0018441 | 12 |
| **KABATZINN J** | 1982 | 10.1016/0163-8343(82)90026-3 | 50 | **RASMUSSEN MK** | 2011 | 10.1080/10615806.2010.515681 | 12 |
| **SEGAL ZV.** | 2002 | | 50 | | | | |

*Note.* Documents are bolded if they do not appear in the Top-20 cited documents of the other country. In case of citation number ties all documents are retained.

*Table 2.7 Most cited documents from China*

| First Author | Year | DOI | WOS Citations | Topic |
|---|---|---|---|---|
| Liu QQ | 2017 | 10.1016/j.chb.2017.02.042 | 65 | Sleep Quality |
| Kong F | 2014 | 10.1016/j.paid.2013.09.002 | 61 | Life Satisfaction |
| Bao X | 2015 | 10.1016/j.paid.2015.01.007 | 55 | Stress |
| Zhang J | 2014 | 10.1016/j.aap.2014.03.006 | 35 | Safety Behavior |
| Wang X | 2014 | 10.1016/j.neuroscience.2014.08.006 | 32 | Default Mode Network |
| Zhang J | 2013 | 10.1016/j.paid.2013.04.004 | 31 | Safety Performance |
| Kong F | 2016 | 10.1080/17470919.2015.1092469 | 24 | Well-Being |
| Lu H | 2014 | 10.1016/j.neuroscience.2014.04.051 | 23 | Brain Structure |
| Chan KKS | 2017 | 10.1007/s12671-016-0675-9 | 21 | Stigma |
| Yang X | 2019 | 10.1007/s10826-018-01323-2 | 19 | Anxiety and Depression |

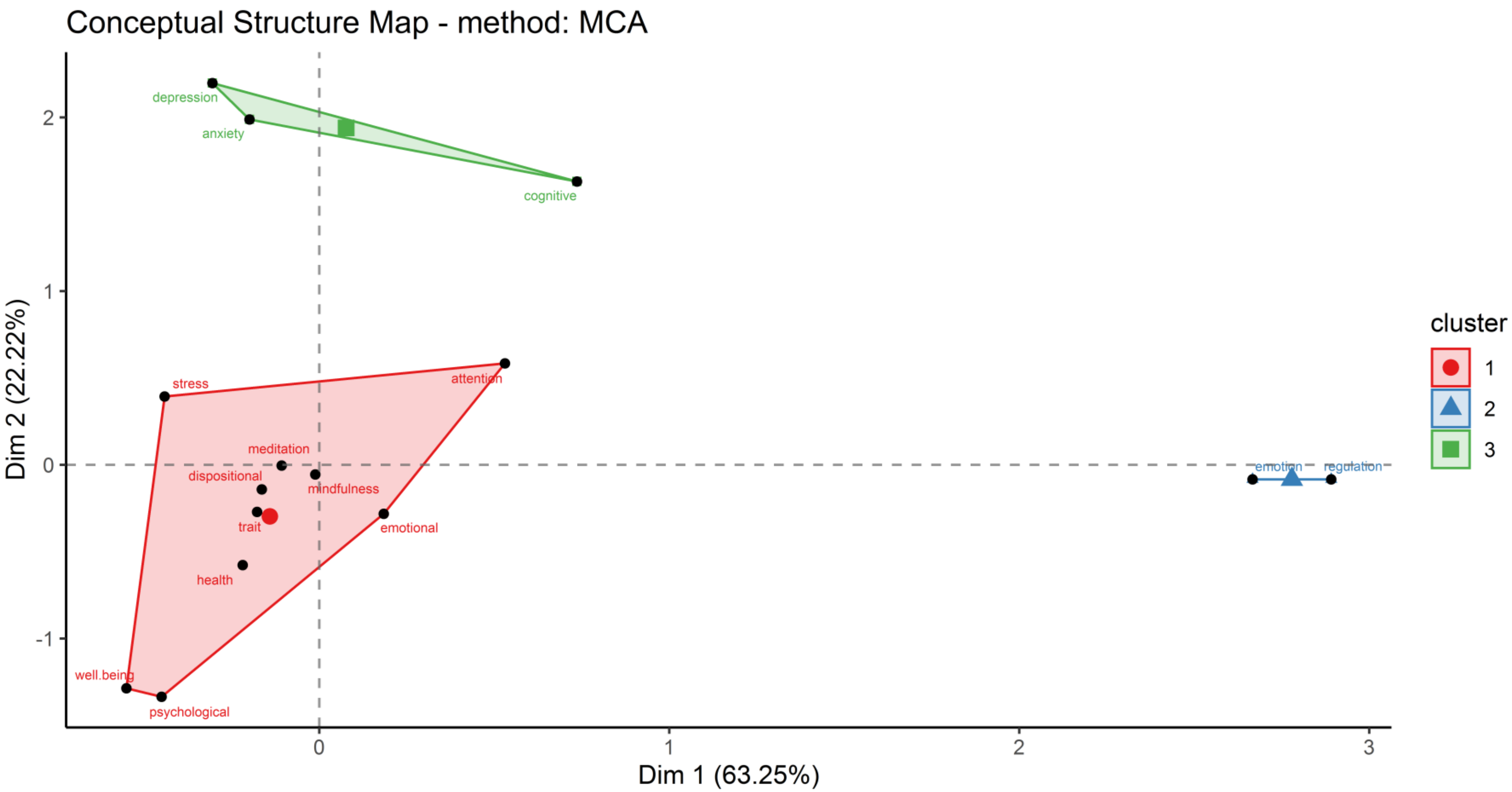



*Figure 2.1* MCA clustering of author keywords at 50 minimal degrees.

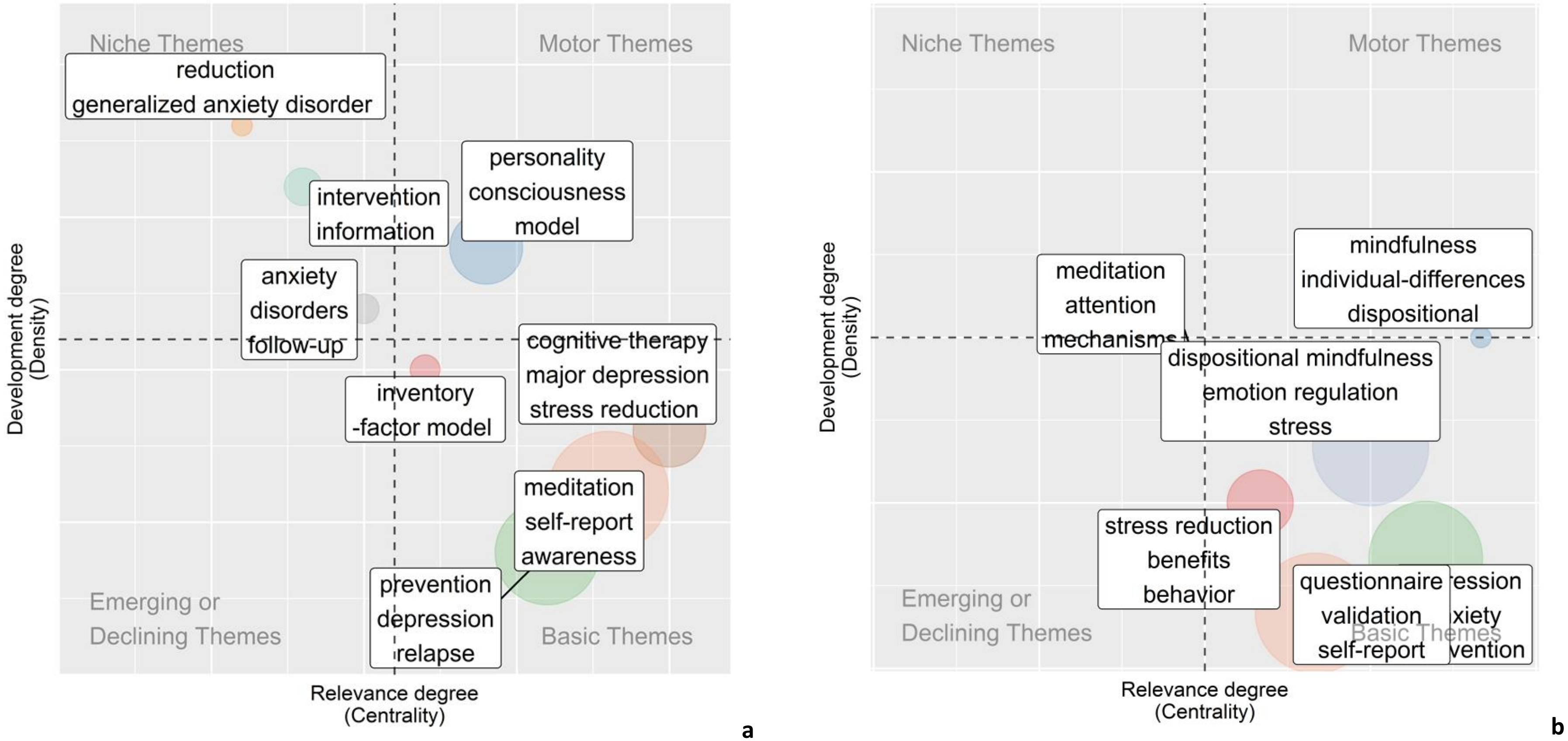



*Figure 2.2* Thematic maps from 2005 to 2010 (a) and from 2011 to 2021 (b)

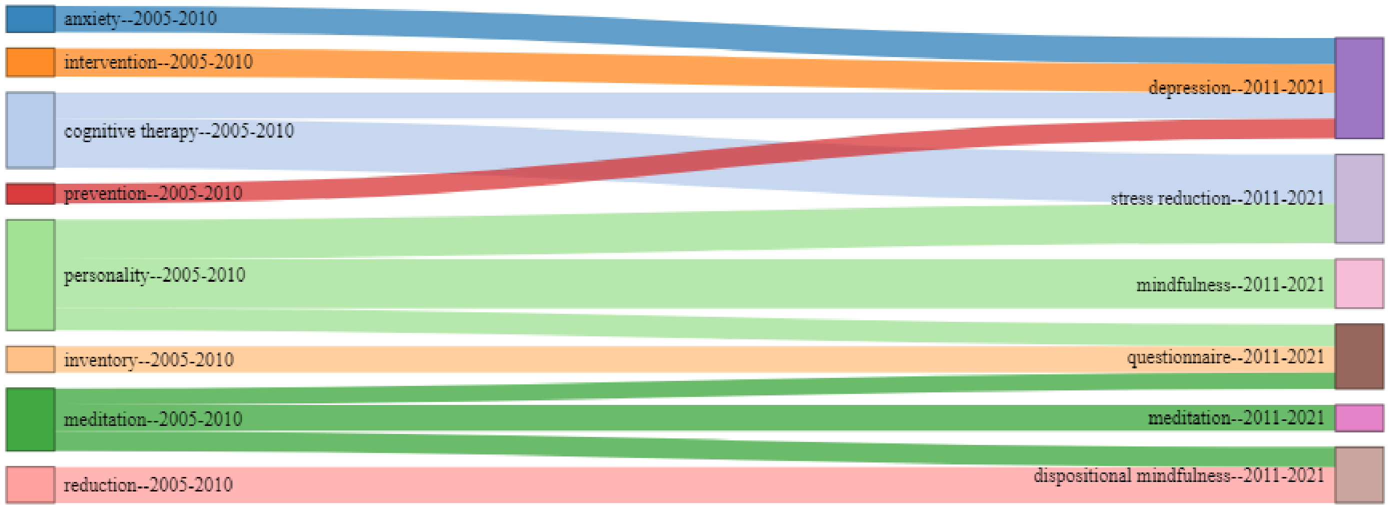



*Figure 2.3* Cluster change between thematic maps

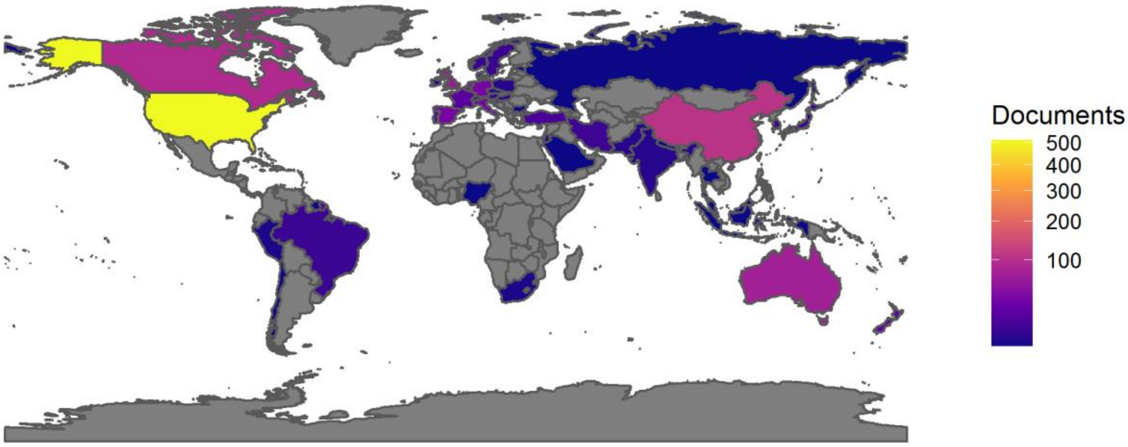



*Figure 2.4* Map of Documents by Country

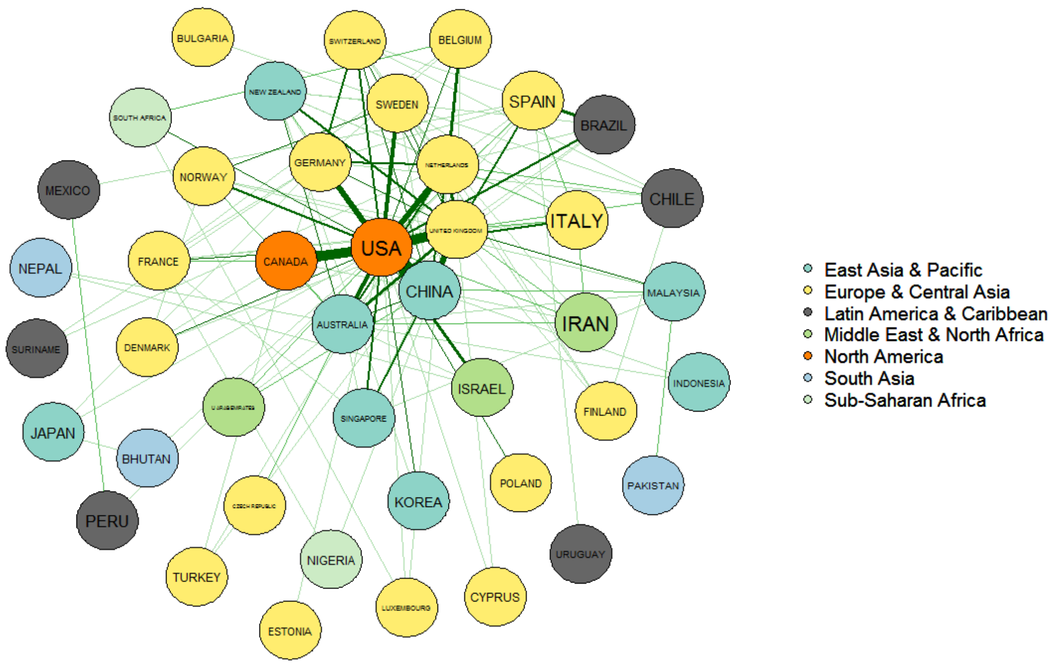



*Figure 2.5* Country-Level Collaboration Graph in the Trait Mindfulness Literature

## Study 2: Revisiting the Five-Facet Structure of Mindfulness[3]

### Preface

As noted in study one, a wide range of different approaches have been developed in the past 15 years to capture trait mindfulness. With the publication of the Five-Facet Mindfulness Questionnaire Baer et al. (2006) presented a first empirical summary of these measures. In the current study I present the first conceptual replication of the process underlying the development of the FFMQ, while also extending the original approach, including measures devised after the publication of the FFMQ. My aim for the current study was to provide empirical support for the replicability of the FFMQ, but also examine whether supposed Western and Eastern mindfulness measures can be conceptually distinguished. In the context of my thesis this study aimed to ensure the applicability of the five-facet conceptualization of mindfulness in the current context, increasing the confidence in the FFMQ for subsequent studies.

[3] This study has previously been published: Karl, J. A., & Fischer, R. (2020). Revisiting the five-facet structure of mindfulness. Measurement Instruments for the Social Sciences, 2(1), 7. https://doi.org/10.1186/s42409-020-00014-3
Minor revisions and stylistic changes have been made to the manuscript to establish coherence with the rest of the thesis.

How robust are our current conceptualizations of mindfulness? Should dispositional mindfulness be thought of as a one-dimensional construct or are there multiple facets and, if yes, how many? This question is important because different traditions of Eastern and Western mindfulness exist. Yet, it is unclear how sensitive current measures are to those distinctions or whether those approaches can be integrated. Dispositional mindfulness is defined as "paying attention in a particular way: on purpose, in the present moment, nonjudgmentally" (Kabat-Zinn, 1994, p. 4) and it has been measured with a number of instruments (Bergomi et al., 2013a, 2013b; Sauer et al., 2013). Trying to find a common structure, Baer et al. (2006) factor analyzed 112 items from the *Mindful Attention and Awareness* (MAAS) Scale, the *Kentucky Inventory of Mindfulness Skills* (KIMS), the *Freiburg Mindfulness Inventory* (FMI), the *Cognitive Affective Mindfulness Scale* (CAMS), and the *Southampton Mindfulness Questionnaire* (SMQ)and reported a five-factor solution of mindfulness when using principal axis factoring with an oblique rotation. Based on this emergent empirical structure, they developed the *Five Facet Mindfulness Questionnaire* (FFMQ) using the 39 highest loading items from the original pool of items. The identification of these five common dimensions across a number of widely used instruments has led to the implicit recognition and acceptance of a multidimensional model of mindfulness (i.e., the Five-Facet Model of Mindfulness, FFMM), with the FFMQ considered to be the prime measure of an underlying multidimensional model of mindfulness (which we call FFMM). Given the widespread use of the instrument and the theoretical implications of the conceptualization of mindfulness, it is important to verify and replicate the emergence of the FFMM even when using different mindfulness measures and with different samples to assess the appropriateness of the FFMQ to measure mindfulness and the validity of the FFMM as a conceptual model of mindfulness.

Since this seminal analysis by Baer et al. (2006), other scales measuring dispositional mindfulness, such as the *Philadelphia Mindfulness Scale* (Cardaciotto et al., 2008) and the *Langer Mindfulness Scale* (Pirson et al., 2018), have been developed. These scales were not included in the original analysis conducted by Baer et al. (2006), but a rigorous replication of the steps taken by Baer et al. (2006) including these scales may indicate the robustness of both the theoretical model of the FFMM and the empirical validity of the FFMQ. The aim of the current study is to examine the comprehensiveness and robustness of the five-factor structure by examining whether the similar five facets emerge if the factor analysis is extended to those new measures.

**History of Mindfulness Assessment**

To provide some historical context, the source scales of the FFMQ were supposed to capture a number of related but distinct dimensions, initially derived from an adaptation of Eastern philosophical thinking to Western audiences (Baer et al., 2006; Kucinskas, 2018). The MAAS (Brown & Ryan, 2003) assesses the lack of attention to one's emotions, thoughts, sensations, and behaviors in general and is proposed to measure present-awareness (Brown & Ryan, 2003; Carlson & Brown, 2005). The KIMS (Baer et al., 2004) conceptualizes mindfulness as a four-dimensional construct with Acting with Awareness, Accept Without Judgement, Describing, and Observing facets. The revised FMI (Walach et al., 2006) assesses a general factor of non-judgmental present-moment awareness, therefore adding the lack of self-evaluation as an important component of the construct. The CAMS-R (Feldman et al., 2007) assesses four facets of mindfulness: Self-Regulation of Attention, Orientation to Present-Moment Experience, Awareness of Experience, and Accepting or Non-Judging Attitude toward Experience. The SMQ (Chadwick et al., 2008) assesses mindfulness in response to distressing images, focusing on decentered awareness, staying open to difficult experience, non-judgmental acceptance, and seeing difficult cognitions as transient mental events without reacting to them to measure a single score of mindfulness.

Despite their differences, a factor analysis of these instruments using principal axis factoring with an oblique rotation based on all 112 items suggested five main facets were sufficient to represent the data (Baer et al., 2006). Of the original set of 112 items, 64 items loaded substantially on one of the five facets. The Observing facet measures the awareness of internal experiences (emotions, cognitions) and external experiences (sounds, sights, and smells). The Describing facet measures the tendency and ability to describe these internal and external experiences with words.

The Acting with Awareness facet measures the tendency to bring full awareness and undivided focus to actions and experiences. The Nonjudging facet measures the tendency to refrain from evaluating inner experiences. The Nonreactivity facet measures the tendency to accept emotions and states as transient and refrain from reacting to them. All these facets seem to capture elements that were central to the Eastern philosophical foundations of mindfulness, except that the spiritual and religious components have been excluded (Kabat-Zinn, 1994; Kucinskas, 2014).

**Mindfulness Assessment since the Development of the FFMQ**

Since the development of the FFMQ, a number of additional measures have been proposed. One such novel measure is the *Philadelphia Mindfulness Scale* (PHLMS: Cardaciotto et al., 2008) which measures present-moment awareness and acceptance as two related but empirically distinct concepts. These two dimensions maintain a Buddhist philosophical approach to mindfulness and previous research has shown this measure to be conceptually related to the FFMQ (Siegling & Petrides, 2016). Therefore, we expect that the *Philadelphia Mindfulness Scale* (PHLMS*)* items will emerge jointly with other related items and can be integrated in the five-facet theoretical model.

However, non-Buddhist measures of mindfulness have also been proposed more recently, most notably the *Langer Mindfulness Scale* (LMS: Pirson et al., 2018). Pirson et al. (2012, p. 3) defined mindfulness as: “a mindset of openness to novelty in which the individual actively constructs novel categories and distinctions”. This Western-approach to mindfulness is more focused on the socio-cognitive elements of mindfulness, highlighting that mindfulness is typically goal-oriented and involves problem-solving and other cognitive exercises. Instead of the more meditative-contemplative aspect of Eastern mindfulness conceptualizations, it explicitly draws on the external, material and social context of the individual. Their new measure is supposed to capture three-facets: novelty-production, novelty-seeking, and engagement. From our perspective, it is interesting to note that the philosophical orientation and the relevant motivational core of mindfulness is different, but the constituent cognitive and attentional elements might be similar. Not surprisingly, while theoretically and philosophically distinct, the LMS and the overall score of the *FFMQ* have been found to correlate moderately at *r* = .33 to .37 (Pirson et al., 2018; Siegling & Petrides, 2014). This raises the question whether these Western-based mindfulness components can be integrated in the existing structure of the FFMQ. Given the theoretical philosophical background of this Western mindfulness tradition, we expect the items of LMS would emerge on distinct factor(s) in a joint factor analysis of mindfulness constructs. One of the interesting questions is how distinct these Western-derived mindfulness dimensions are when analyzed together with instruments that have been inspired by Eastern philosophy.

**Current Research**

In summary, the FFMQ has emerged as the prime measure to capture the FFMM (Baer et al., 2006). The FFMQ has been derived in a bottom-up approach by factor analyzing pre-existing measures (Baer et al., 2006, p. 2). This empirically driven approach requires confirmation and replication to assess the theoretical appropriateness of the FFMQ as the principal measure of a multidimensional mindfulness construct (Magnusson, 1992; Tellis, 2017). While previous studies have employed a confirmatory strategy using only the final FFMQ (Gu et al., 2016; M. J. Williams et al., 2014), no study to date has undertaken a conceptual replication of the generation of the underlying FFMM. One reason this is important is to examine the potential presence of item wording effects in the current measurement of mindfulness (for studies reporting such method factors in the FFMQ see: Aguado et al., 2015; Van Dam et al., 2012). Further, these studies have shown that a bi-factor model of the FFMQ, in which all items load onto a general factor of mindfulness and their individual facets while including wording factors substantially improved the structure. This indicates that beyond their assignment to individual facets mindfulness items might share some common variance that could be explained by a general factor (for a discussion of this interpretation of a bi-factor model see: Bonifay et al., 2017). In the FFMQ this factor could represent Buddhist inspired mindfulness raising the question if a similar bi-factor model emerges when Western-oriented measures of mindfulness are included. Investigating the emergent structure of the mindfulness measures is also of interest, because both novel Buddhist inspired as well as Western-oriented measures of mindfulness have been developed since the publication of the FFMQ, raising important questions both about the comprehensiveness of the FFMM and the appropriateness of the FFMQ to measure such a multidimensional model of mindfulness. The current study aims to extend the current research on the dimensionality of mindfulness by re-examining the emergence of multi-dimensional mindfulness structures including recent measures of mindfulness.

# Methods

## Participants

We sampled 404 undergraduate students at Victoria University of Wellington. Five participants (1.24% of the total) started the questionnaire but did not finish it. Due to the low number of participants that did not answer the survey completely we removed those five individuals from the dataset, leaving an effective sample size of 399. The average age of the participants was 19.21(*SD* = 3.93) and 68.92% of the total sample were female.

**Previous mindfulness practice.** Of the total sample, 8.77% reported previous mindfulness experience, 9.52% reported yoga experience, and 10.03% reported meditation experience. This sample composition in terms of age and mindfulness experience is comparable to the original *FFMQ* study (Baer et al., 2006). Due to the low number of participants with previous meditation experience, we did not perform separate analysis comparing meditation practitioners and participants with no meditation experience.

**Procedure.** Participants filled out an online survey on Qualtrics (the Qualtrics survey file and a word version of the survey are available on the OSF: https://osf.io/k2m35/?view_only=14472f2f0bef4deb8f0097a4cd421414). The mindfulness scales were presented as part of a larger survey pack. The survey pack also contained measures of personality (Soto & John, 2017), reinforcement sensitivity (Corr & Cooper, 2016), values (Schwartz et al., 2012), impression management (Blasberg et al., 2014), self-deception (Paulhus & Reid, 1991), satisfaction with life (Diener et al., 1985), flourishing (Diener et al., 2010), and a number of behavioral tasks (pen choices) to assess group conformity. The complete data are available on the *OSF*. Individuals participated as part of an Introduction to Psychology course and received course credit.

## Open Science Statement

The current study reports an exploratory analysis into the structure of mindfulness. Recent studies (Silberzahn et al., 2018) demonstrate the impact of analytic freedom on reported outcomes. We aim to provide maximum transparency of the analysis by providing the full raw data set, the analytic code, and all materials associated with the study on the Open Science Framework (https://osf.io/k2m35/?view_only=14472f2f0bef4deb8f0097a4cd421414). The current study was part of a larger pack of surveys administered to the participants.

**Instruments**

**The Mindful Attention and Awareness Scale (MAAS).** The MAAS (Brown & Ryan, 2003) uses 15 items that a participant rates on a scale from 1- (*Almost always*) to 6 - (*Almost never*). Example items are: “I do jobs or tasks automatically, without being aware of what I'm doing.” and “I find myself listening to someone with one ear, doing something else at the same time.” Lower scores on these items indicate greater mindfulness.

**The Southampton Mindfulness Questionnaire (SMQ).** We used the 16-item SMQ (Chadwick et al., 2008), with a 7-point Likert scale ranging from 1- (*Strongly disagree*) to 7 - (*Strongly agree*). The questionnaire was preceded by the statement: “Usually when I experience distressing thoughts and images...”. Example items are: ”I am able just to notice them without reacting.” and “They take over my mind for quite a while afterwards.”

**The Cognitive and Affective Mindfulness Scale-Revised (CAMS-R).** The CAMS-R is a 12 item measure with four subcomponents (Feldman et al., 2007). Participants answered the items on a 4-point Likert scale ranging from 1- (*Rarely/Not at All*) to 4 - (*Almost Always*). Example items for the individual subcomponents are “It is easy for me to concentrate on what I am doing.” (Attention), “I am able to focus on the present moment.” (Present Focus), “It’s easy for me to keep track of my thoughts and feelings.” (Awareness), “I can tolerate emotional pain.” (Acceptance).

**The Freiburg Mindfulness Inventory (FMI).** We used the 14-items FMI (Walach et al., 2006) with the original 4-point Likert scale ranging from 1-(Rarely) to 4-(Almost always). Example items are

"I am open to the experience of the present moment." and "I sense my body, whether eating, cooking, cleaning or talking." In their original study Baer et al. (2006) used an earlier developmental version of the FMI which had 30 items.

**Kentucky Inventory of Mindfulness Skills (KIMS).** We used the 39-items KMI to assess a multi-dimensional conceptualization of mindfulness (Baer et al., 2004). The items are rated on a 5-point Likert scale ranging from 1-( *Never or very rarely true*) to 5-(*Very often or always true*). Example items are "I'm good at finding the words to describe my feelings." (Describing), "I notice changes in my body, such as whether my breathing slows down or speeds up." (Observing), "When I do things, my mind wanders off and I'm easily distracted." (Acting with Awareness), "I criticize myself for having irrational or inappropriate emotions." (Non-Judging).

**The Langer Mindfulness Scale.** We used the 14-item LMS (Pirson et al., 2018) to assess a multi-dimensional conceptualization of socio-cognitive mindfulness. The items are rated on a 7-point Likert scale ranging from 1-(*Strongly disagree*) to 7-(*Strongly agree*). Example items are: "I am rarely alert to new developments." (Engagement), "I make many novel contributions." (Novelty Producing), "I like to investigate things." (Novelty Seeking).

We report the reliabilities and scale descriptives of all measures in Table 3.1. We decided to evaluate reliability using ω, the Greatest Lower Bound (GLB), and coefficient H (H). These indicators have been shown in previous research to provide better estimations of reliability compared to α (McNeish, 2018; Trizano-Hermosilla & Alvarado, 2016). We nevertheless report α for comparison purposes. Both α and ω are reported with bootstrapped 95% confidence intervals. All reliability coefficients were obtained using the userfriendlyscience package (version 0.7.2) in R (Peters, 2018). The reliabilities were acceptable (values above .7), except for LMS Engagement, CAMS Awareness, CAMS Acceptance, and CAMS Present Focus.

**Analytical Approach**

We first examined the theoretically proposed fit for each mindfulness scale using separate CFAs. This analysis provides important information on the internal validity of each of these measures and therefore, offers important background information for understanding the replication study. For each scale, we fitted the structures which were proposed by the original authors of the measures. Specifically, we fitted a uni-dimensional model for the FMI, the SMQ, and the MAAS, respectively. For the PHLMS, we fitted a model with two correlated first-order factors (Acceptance, Awareness). For the LMS, we fitted a model with three correlated first-order factors (Novelty Producing, Novelty Seeking, Engagement). For the KIMS, we fitted a model with four first-order factors (Observing, Describing, Non-Judging, and Acting with Awareness) and a second-order factor representing mindfulness. For the CAMS-R, we fitted a model with four first-order factors (Attention, Present Focus, Awareness, and Acceptance) and a second-order factor representing overall mindfulness. Therefore, we have a number of single factor models (FMI, SMQ, MAAS); a two-factor model (PHLMS), a three-factor model (LMS), and two four-factor models with a second-order mindfulness factor (CAMS-R, KIMS).

Due to multi-variate non-normality of our data, all models were fitted using an WLMSV estimator rather than parceling items (Li, 2016; Maydeu-Olivares, 2017)[4]. We use the following fit indices : A $\chi^2$/degrees of freedom ratio of < 5 is considered acceptable (Wheaton et al., 1977), CFI and γ (with .90 defined as threshold for acceptable fit and .95 defined as threshold for good fit, Marsh et al., 2004), RMSEA (with less than 0.01, 0.05, and 0.08 to indicate excellent, good, and mediocre fit respectively, MacCallum et al., 1996), and SRMR (acceptable fit is indicated by values less than .08, Hu & Bentler, 1999). We further report $\chi^2$ and degrees of freedom for each model, but do not focus on these indicators due to the known dependency on sample size.

Second, we ran an exploratory factor analysis using all mindfulness items to investigate the structure across all items and all instruments. We started off with a parallel analysis using the

[4] The use of a WLMSV estimator was recommended by an anonymous reviewer of the published manuscript. We also fitted all model using an MLM estimator, which yielded substantively identical results.

complete pool of items from all the mindfulness scales to determine the optimal number of components while accounting for components occurring due to random chance. We used Glorfeld's (1995) conservative approach instead of Horn's (1965) parallel analysis. We retained components which had eigenvalues greater than one after adjusting the initial eigenvalues for the eigenvalues observed in a random data set.

To examine the unfolding of the factor structure (see Goldberg, 2006), we implemented an iterative process in which we ran a PCA with 1 up to the number of factors proposed by the parallel analysis. After extracting each set of components using a principal component analysis with a varimax rotation using the psych package (version 1.8.12) in R (Revelle, 2018), we correlated participants' scores on these components with the previously extracted component (Goldberg, 2006). This approach provides insight into the pattern of emergence of components (for examples see: De Raad et al., 2014; Raad & Oudenhoven, 2008, 2011).

## Results

### Confirmatory factor analysis

The CFA of the individual scales showed acceptable fit for the FMI, the MAAS, and the CAMS-R. Interestingly, the CAMS-R showed good fit while its individual scales had poor reliability. The other measures showed less than acceptable overall fit (see Table 3.2). Compared to previous studies using these measures we found that in our sample the FMI and CAMS-R showed better fit, whereas the PHLMS, KIMS, SMQ, and LMS showed worse fit compared to other studies (we include a table reporting fit statistics from previous studies which we used to compare our results against in Appendix B).

### Factorial Structure

The parallel analysis suggested 6 components (adjusted eigenvalues: 15.61, 8.27, 2.39, 1.55, 1.37, 1.27). We therefore extracted 1 to 6 components based on the parallel analysis explaining 34% of the total variance. For the 6-component structure, we report the highest negative and positive loading items for each component in Table 3.3 to allow for easier interpretation. Additionally, the full loading matrix for the 6-component solution can be found in Table 3.4. We only interpreted loadings > .40 when examining the loading matrices of the items. No items were deleted. Due to space constraints, we made the full rotated component matrices for all solutions available on the OSF: https://osf.io/k2m35/?view_only=14472f2f0bef4deb8f0097a4cd421414. The final six components were labelled: “Non-Judgement/Non-Reacting”, “Observing”, “Acting with Awareness”, “Reacting/Judgement”, “Describing”, “Openness/Western Mindfulness”.

When examining the single component extracted first, it was primarily defined by Non-Judgmental Awareness items. This single factor seems to support the interpretation of mindfulness in line with Kabat-Zinn’s definition of mindfulness as: “paying attention in a particular way: on purpose, in the present moment, nonjudgmentally” (1994, p.4), indicating that the core element of mindfulness is a quality of awareness rather than describing emotions or non-reactance. As can be seen in Figure 3.1, Observing then split off from this general component. In the third step, a component defined by describing and focus items emerged. This component was positively related to observing and negatively to judgment. In the fourth step the Describing/Focus component split into Self-Criticism and Describing/Openness. In the fifth step, Self-Criticism split into Acting with Awareness and Self-Criticism. In the sixth step Describing/Openness split into Describing and Openness.

Overall, the first distinct components within the larger structure to emerge were Observing and Non-Judgement in the three-component solution. These components remained uncorrelated to all other components (with the exception of Observing being correlated with Describing), highlighting the distinctiveness of these mindfulness components from the remainder of the mindfulness construct. A further empirically distinct component was Acting with Awareness which emerged in the five components solution, followed by describing and by Openness (LMS) in the 6-component solution.

Focusing on the origins of the individual components of the final six-component solution, Non-Judgment/Non-Reacting was defined by items of the SMQ and the FMI. Some unique items of the PHLMS, such as “I wish I could control my emotions more easily”, showed substantial negative loadings on this component. The items from the LMS did not load substantially on this component. Observing was mostly defined by PHLMS items, such as “When I walk outside, I am aware of smells or how the air feels against my face”. Some KIMS items measuring observing, such as “I pay attention to sensations, such as the wind in my hair or sun on my face” also loaded on this component. Overall, we did not find substantial negative loadings on this component. Acting with Awareness was positively defined by reverse keyed MAAS items, such as “I find myself doing things without paying attention”, and negatively defined by KIMS items, such as “When I do things, my mind wanders off and I’m easily distracted”. We found no substantial loadings of either LMS or PHLMS items. Reacting/Judgement was largely defined by PHLMS items, such as “If there is something I don't want to think about, I'll try many things to get it out of my mind". A number of KIMS items, such as "I tell myself that I shouldn't be feeling the way I'm feeling.", also loaded positively on the component. We did not find substantial negative loadings on this component. Describing was positively defined by KIMS items, such as "I'm good at finding the words to describe my feelings", and negatively by KIMS items, such as "It's hard for me to find the words to describe what I'm thinking". We did not find substantial loadings of the LMS and only one item ("When someone asks how I am feeling, I can identify my emotions easily") of the PHLMS loaded substantially. Last, Openness/Western Mindfulness was largely defined by LMS items, such as "I like to be challenged intellectually". The only two non-LMS item loading substantially positively on the component were from the KIMS ("I notice visual elements in art or nature, such as colors, shapes, textures, or patterns of light and shadow" and "I tend to evaluate whether my perceptions are right or wrong."). Substantial negative loading items were exclusively LMS items, such as "I am not an original thinker".

As suggested by an anonymous reviewer, to examine the possibility of a general response factor, we ran confirmatory factor analysis with lavaan using a WLSMV estimator (for further model specifications and analytical code, see the supplementary material on the OSF). We fitted a model in which each item loaded on the factor on which it showed the highest loading in the exploratory factor analysis reported above. Additionally, all items were loaded on a separate general response factor, which was uncorrelated with the substantive factors. Item loadings were freely estimated by standardizing the latent variable. The relative fit of the six-factor structure is significantly improved when including a general response factor ($\Delta$CFI from the model without response factor: .087). Unfortunately, we were unable to fully explore positively vs negatively wording factors because the emerging factors in our analysis were not well-balanced in their phrasing. In order to disentangle possible content and method-artefacts, future studies need to include balanced item sets using both positively and negatively phrased items across all domains. Our exploratory findings suggest that item wording effects need greater attention in the measurement of mindfulness.

## Discussion

The goal of the current research was to examine whether the commonly accepted multidimensional structure of mindfulness as exemplified in the FFMQ can be conceptually replicated using measures originally included in the development of the FFMQ while also including additional theoretically similar (PHLMS) and dissimilar (LMS) measures.

While we recovered three facets that expressed the same content as Observing, Describing, and Acting with Awareness in the FFMQ, we did not find separate Non-Judging and Non-Reacting components. This indicates that the distinction between those two components of mindfulness requiring distinct cognitive and behavioral reactions, while theoretically important, might not be sufficiently clear and distinct for participants in our sample. Both components require adaptation of cognitive and behavioral responses after noticing internal or external sensations, emotions and thoughts. These distinctions appear to be too subtle, as these two factors merged to a generic non-reactivity factor in our sample. Similar factors combining non-judgement and non-reaction have been reported in other instruments (for example the CAMS-R, PHLMS). At the same time, when including an additional measure of mindfulness with a distinct philosophical background, we identified an additional factor. Overall, this indicates that while some individual components can be recovered and are broadly in line with previous conceptualizations of mindfulness, we did not recover the complete structure of the FFMQ with all its nuances and it may miss additional components of interest to mindfulness researchers.

We found that an Openness/Western conceptualization of mindfulness emerged as a clearly defined separate component. This supports the theoretical separation of these measures because openness as a core component of a Western mindfulness definition can be empirically separated from items supposed to measure Eastern-philosophical perspectives on mindfulness (Pirson et al., 2018). Interestingly, the LMS is supposed to show a three-dimensional structure but in our sample the overall fit for the three factors was poor and in the item level analysis, a single distinct factor emerged. At the same time, our examination of the unfolding component structure provides important insight into the components that the Eastern and Western conceptualizations of mindfulness share, which helps to explain positive relationships between the LMS and the FFMQ reported in previous research (Siegling & Petrides, 2014, 2016). The positive relationship of the LMS noted in previous research might be due to the Describing facet of the FFMQ (the Describing facet showed the strongest positive correlations with the LMS during validation studies, see Pirson et al., 2018). In the current study the LMS/Openness components were most clearly associated with Describing during the unfolding of the facture structure and the LMS/Openness only split from this factor and emerged as a separate factor when six components were extracted. This suggests that the ability to describe one's feelings and experiences is an important correlate of being open for new experience as well as enjoying those experiences. Therefore, our analysis suggests that even though Western conceptualizations of mindfulness draw upon different philosophical traditions, the relevant social and cognitive components might still be shared with Eastern-based conceptualizations of mindfulness.

We found a component that expressed Judging/Reacting and was mostly defined by negatively worded items. This further highlights possible method artifacts in the measurement of mindfulness (Aguado et al., 2015). Studies using a person-centered approach to the FFMQ found a profile that was defined by judging, rather than non-judging (Bravo et al., 2016; Pearson et al., 2015). These patterns raise the possibility that a number of reversely worded items, possibly from the Non-Judging or Non-Reacting facets, do not measure the polar opposites of the positively worded items, but rather tap into a separate construct masked as a response style component. This is a finding consistent with previous studies that found that the fit of the FFMQ can be improved through a bi-factor model, indicating the potential presence of a g-factor of mindfulness explaining variance beyond the individual facets (Aguado et al., 2015; Van Dam et al., 2012). The interpretation of bi-factor models has been controversial (Bonifay et al., 2017) and further research is needed to understand the meaning of such a factor in the context of mindfulness

**Strengths**

Our current study brings together theoretically similar and distinct measures of mindfulness, highlighting the general robustness of the FFMM and appropriateness of the FFMQ to measure mindfulness. It also shows that it is possible to discriminate Western-based conceptualizations of mindfulness from Eastern mindfulness measures. At the same time, it appears that Western-based measures of mindfulness may tap into similar social and cognitive processes that are also fundamental to the traits and abilities captured by Eastern-based mindfulness measures. We used a shortened version of the FMI, therefore, our current study did not employ the exact measures of the study conducted by Baer et al. (2006). Nevertheless, we only recovered three facets (Observing, Describing, Acting with Awareness), and one facet that expressed a combination of non-reacting/non-judging. Together with our finding that some items form a negative wording factor this indicates that the current dimensional conceptualization of mindfulness might need revision.

**Limitations**

A limitation of our current study, while still closely resembling the sample used by Baer et al. (2006) during the development of the FFMQ, is the use of a sample of young adults in a Western educational context with a low percentage of active meditators. Previous research found that the observing facet is more strongly related to the general factor of mindfulness in samples with meditation experience (Lilja et al., 2013). However, our New Zealand-based sample is conceptually interesting because New Zealand has an official bi-cultural status, in which the national culture is actively co-constructed from both Western influences and traditional Maori culture (for a concise review of New Zealand history see: Mein Smith, 2011). This bi-cultural model undergirds the social and educational context which has led to more nuanced perceptions of the mind-body duality in a general population compared to North American or Western European settings. This interweaving of cultural practices is increasingly recognized and more explicit connections between specific Maori cultural practices and Eastern-based mindfulness practices are explored (Higgins & Eden, 2018).

Therefore, the insights from this sample are informative even in the absence of a larger number of meditators or mindfulness practitioners.

## Conclusions

Overall, we found that three of the five FFMM components (Observing, Describing, Acting with Awareness) emerged in a conceptual replication and two of the factors merged, which has been found in the structures of other mindfulness instruments. This indicates potentially simpler cognitive and behavioral mindfulness components in lay audiences than indicated by the FFMM. Furthermore, conceptually distinct LMS items emerged as a separate component, highlighting that a) at least three of the five dimensions of the FFMQ seem to reliably emerge even if new measures of mindfulness are included and b) that there might be additional components of mindfulness from a Western perspective that are not captured in the FFMM. A third important insight from the unfolding analysis is that the different facets capture distinct aspects of mindfulness with low intercorrelations across some of the facets across the different levels of unfolding, which implies that it is more relevant to use mindfulness scores at a facet level rather than as a general score. Finally, negative wording effects were also apparent, and a number of the negative items might not tap into the proposed concepts but rather capture response tendencies.

**Availability of data and materials.** All materials, data, and analytic code is available on the OSF (https://osf.io/k2m35/?view_only=14472f2f0bef4deb8f0097a4cd421414)

*Table 3.1 Reliability and scale descriptives of the mindfulness measures*

| | M | SD | α | α low | α high | ω | ω low | ω high | GLB | H |
|---|---|---|---|---|---|---|---|---|---|---|
| CAMS-R Attention | 2.20 | 0.62 | .731 | .679 | .775 | .731 | .671 | .772 | .726 | .735 |
| CAMS-R Present Focus | 2.48 | 0.57 | .491 | .396 | .573 | .509 | .385 | .595 | .518 | .528 |
| CAMS-R Awareness | 2.39 | 0.62 | .581 | .504 | .656 | .590 | .506 | .655 | .595 | .611 |
| CAMS-R Acceptance | 2.65 | 0.67 | .605 | .517 | .666 | .606 | .529 | .664 | .603 | .614 |
| Freiburg Mindfulness Inventory | 2.55 | 0.47 | .823 | .796 | .848 | .826 | .794 | .849 | .856 | .853 |
| Langer Mindfulness Scale Engagement | 4.97 | 0.97 | .598 | .514 | .672 | .605 | .518 | .668 | .616 | .626 |
| Langer Mindfulness Scale Novelty Producing | 4.30 | 0.93 | .653 | .591 | .709 | .689 | .637 | .736 | .742 | .790 |
| Langer Mindfulness Scale Novelty Seeking | 5.33 | 0.92 | .761 | .713 | .801 | .761 | .715 | .799 | .786 | .781 |
| MAAS | 3.46 | 0.73 | .837 | .810 | .862 | .839 | .811 | .863 | .839 | .865 |
| Philadelphia Mindfulness Scale Awareness | 3.52 | 0.62 | .811 | .779 | .839 | .812 | .777 | .839 | .865 | .830 |
| Philadelphia Mindfulness Scale Acceptance | 3.47 | 0.78 | .876 | .855 | .895 | .876 | .852 | .894 | .891 | .886 |
| Southampton Mindfulness Questionnaire | 3.79 | 0.84 | .865 | .844 | .885 | .867 | .843 | .886 | .902 | .878 |
| Kentucky Mindfulness Inventory Observing | 3.24 | 0.62 | .810 | .779 | .838 | .810 | .777 | .839 | .832 | .814 |
| Kentucky Mindfulness Inventory Describing | 3.05 | 0.77 | .870 | .847 | .888 | .875 | .854 | .894 | .882 | .895 |
| Kentucky Mindfulness Inventory Acting with Awareness | 2.71 | 0.49 | .657 | .600 | .710 | .645 | .572 | .705 | .758 | .787 |
| Kentucky Mindfulness Inventory Non-Judging | 2.86 | 0.77 | .863 | .840 | .884 | .869 | .843 | .887 | .908 | .892 |

*Notes:* α and ω are reported with 95% bias corrected confidence intervals, GLB = Greatest Lower Bound, H = Coefficient H

*Table 3.2 CFA fit of the individual mindfulness measures*

| Measure | $\chi^2$ | df | $\chi^2$/df | CFI | RMSEA | Gamma | Overall Fit |
|---|---|---|---|---|---|---|---|
| Freiburg Mindfulness Inventory | 168.288 | 77 | 2.186 | .922 | .055 | .968 | Good |
| Langer Mindfulness Scale | 19.636 | 74 | 2.576 | .881 | .063 | .96 | Poor |
| MAAS | 156.588 | 90 | 1.74 | .944 | .043 | .978 | Good |
| Philadelphia Mindfulness Scale | 379.246 | 169 | 2.244 | .884 | .056 | .95 | Poor |
| Southampton Mindfulness Questionnaire | 435.428 | 104 | 4.187 | .765 | .089 | .906 | Poor |
| Kentucky Mindfulness Inventory | 1407.766 | 696 | 2.023 | .766 | .051 | .916 | Poor |
| CAMS-R | 129.752 | 50 | 2.595 | .909 | .063 | .968 | Good |

Notes. All models were fitted with a WLMSV estimator. Overall fit is assessed as good if CFI > .90, RMSEA < .08, and SRMR < .08

*Table 3.3 Components extracted with their highest positive and negative loading items.*

| Component | Name | Positive | Negative |
|---|---|---|---|
| 1_1 | Non-Judgmental Awareness | I am able to accept the thoughts and feelings I have. | I think some of my emotions are bad or inappropriate and I shouldn't feel them.. |
| 2_1 | Judgmental Non-Awareness | I tell myself that I shouldn’t be feeling the way I’m feeling. | I am able to accept the thoughts and feelings I have. |
| 2_2 | Observing | I intentionally stay aware of my feelings. | I am not an original thinker. |
| 3_1 | Non-Judgment | I am friendly to myself when things go wrong. | I think some of my emotions are bad or inappropriate and I shouldn't feel them. |
| 3_2 | Observing | When talking with other people, I am aware of the emotions I am experiencing. | I am not an original thinker. |
| 3_3 | Describing/ Focus | When someone asks how I am feeling, I can identify my emotions easily. | It's hard for me to find the words to describe what I'm thinking. |
| 4_1 | Non-Judgment | I am friendly to myself when things go wrong. | Usually when I experience distressing thoughts and images... I get angry that this happens to me. |

| | | | |
|---|---|---|---|
| 4_2 | Observing | When I walk outside, I am aware of smells or how the air feels against my face. | I am rarely aware of changes. |
| 4_3 | Self-Criticism | I tell myself that I shouldn’t be feeling the way I’m feeling. | It seems I am “running on automatic” without much awareness of what I’m doing. |
| 4_4 | Describing/ Openness | I’m good at finding the words to describe my feelings. | It’s hard for me to find the words to describe what I’m thinking. |
| 5_1 | Non-Judgment | I am friendly to myself when things go wrong. | Usually when I experience distressing thoughts and images... I get angry that this happens to me. |
| 5_2 | Observing | When I walk outside, I am aware of smells or how the air feels against my face. | I am rarely aware of changes. |
| 5_5 | Self-Criticism | I tell myself that I shouldn’t have certain thoughts. | I accept myself the same whatever the thought/image is about in my mind. |
| 5_3 | Acting with Awareness | I find myself doing things without paying attention. | When I do things, my mind wanders off and I’m easily distracted. |

| | | | |
|---|---|---|---|
| 5_4 | Describing | I’m good at finding the words to describe my feelings. | It’s hard for me to find the words to describe what I’m thinking. |
| 6_1 | Non-Judgement / Non-Reacting | I am friendly to myself when things go wrong. | I wish I could control my emotions more easily. |
| 6_2 | Observing | When I walk outside, I am aware of smells or how the air feels against my face. | I am rarely aware of changes. |
| 6_3 | Acting with Awareness | I find myself doing things without paying attention. | When I do things, my mind wanders off and I’m easily distracted. |
| 6_5 | Reacting-Judgment | If there is something I don’t want to think about, I’ll try many things to get it out of my mind. | I accept myself the same whatever the thought/image is about in my mind. |
| 6_6 | Describing | I’m good at finding the words to describe my feelings. | It’s hard for me to find the words to describe what I’m thinking. |
| 6_4 | Openness / Western Mindfulness | I like to be challenged intellectually. | I am not an original thinker. |

*Table 3.4 Component loadings of the mindfulness items on the final six components.*

| C1 | C5 | C2 | C3 | C6 | C4 | Item | Measure |
|---|---|---|---|---|---|---|---|
| **.61** | -.14 | .09 | .12 | .03 | .06 | I am friendly to myself when things go wrong. | FRBRG |
| **.61** | -.12 | .17 | .12 | .06 | .04 | I am able to appreciate myself. | FRBRG |
| **.60** | -.21 | -.06 | -.02 | .19 | -.01 | I feel calm soon after. | SMQ |
| **.58** | -.23 | .04 | .06 | .08 | .11 | I am able to accept the experience. | SMQ |
| **.56** | -.14 | -.15 | .03 | .07 | .04 | I just notice them and let them go. | SMQ |
| **.55** | -.04 | .06 | .05 | -.08 | -.04 | I see my mistakes and difficulties without judging them. | FRBRG |
| **.54** | -.24 | .06 | -.01 | .16 | .05 | I 'step back' and am aware of the thought or image without getting taken over by it. | SMQ |
| **.53** | -.03 | -.08 | .11 | .07 | .16 | I watch my feelings without getting lost in them. | FRBRG |
| **.52** | -.29 | .17 | .15 | .21 | .09 | I am able to accept the thoughts and feelings I have. | CAMS |
| **.51** | -.30 | .09 | .09 | .09 | -.04 | I accept myself the same whatever the thought/image is about in my mind. | SMQ |
| **.51** | .11 | .18 | .10 | -.02 | .11 | I am open to the experience of the present moment. | FRBRG |
| **.49** | -.08 | .02 | -.01 | .12 | -.07 | I try just to experience the thoughts or images without judging them. | SMQ |
| **.49** | .07 | .15 | .29 | .16 | .08 | I am able to focus on the present moment. | CAMS |
| **.49** | .14 | .17 | .20 | .13 | .09 | I feel connected to my experience in the here-and-now. | FRBRG |
| **.49** | -.04 | .11 | .12 | .12 | .08 | I experience moments of inner peace and ease, even when things get hectic and stressful. | FRBRG |
| **.48** | -.15 | .25 | .03 | .09 | -.04 | I try to notice my thoughts without judging them. | CAMS |
| **.47** | .00 | .09 | .04 | -.06 | .12 | I am able to smile when I notice how I sometimes make life difficult. | FRBRG |
| **.46** | -.15 | -.21 | .03 | .11 | .09 | I am able just to notice them without reacting. | SMQ |
| **.46** | -.03 | .07 | -.01 | .08 | .10 | I notice how brief the thoughts and images really are. | SMQ |
| **.45** | -.14 | .10 | .12 | .04 | .16 | I can accept things I cannot change. | CAMS |
| **-.44** | **.44** | .18 | -.11 | -.11 | -.03 | I wish I could control my emotions more easily. | PHLMS |
| **.43** | .03 | .26 | .20 | .06 | .07 | When I notice an absence of mind, I gently return to the experience of the here and now. | FRBRG |
| .38 | -.19 | -.02 | -.02 | .12 | .26 | I can tolerate emotional pain. | CAMS |
| -.37 | .23 | .29 | -.31 | -.10 | -.05 | I lose myself in the thoughts/images. | SMQ |
| .37 | -.02 | .09 | .16 | -.08 | .21 | In difficult situations, I can pause without immediately reacting. | FRBRG |
| -.36 | .35 | .29 | -.13 | -.06 | -.04 | They take over my mind for quite a while afterwards. | SMQ |
| .31 | -.01 | .09 | .03 | .05 | .30 | I accept unpleasant experiences. | FRBRG |
| -.20 | .15 | .12 | -.15 | -.03 | .18 | I am preoccupied by the future. | CAMS |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| .18 | .05 | -.10 | -.09 | -.10 | -.09 | I seldom notice what other people are up to. | LMS |
| .07 | **.69** | .11 | -.04 | -.07 | -.21 | If there is something I don't want to think about, I'll try many things to get it out of my mind. | PHLMS |
| -.03 | **.68** | .04 | -.10 | -.07 | -.08 | I try to stay busy to keep thoughts or feelings from coming to mind. | PHLMS |
| .00 | **.66** | .07 | -.05 | -.06 | -.10 | I try to distract myself when I feel unpleasant emotions. | PHLMS |
| -.31 | **.65** | .01 | -.09 | -.11 | .08 | I tell myself that I shouldn't have certain thoughts. | PHLMS |
| -.17 | **.65** | .07 | -.07 | -.03 | -.08 | There are things I try not to think about. | PHLMS |
| .04 | **.64** | .10 | -.07 | -.04 | -.19 | When I have a bad memory, I try to distract myself to make it go away. | PHLMS |
| .17 | **.58** | .03 | -.20 | -.04 | -.19 | I try to put my problems out of mind. | PHLMS |
| -.35 | **.57** | -.02 | -.21 | -.14 | .15 | I tell myself that I shouldn't be feeling the way I'm feeling. | KIMS |
| -.30 | **.56** | .00 | -.10 | -.11 | .07 | There are aspects of myself I don't want to think about. | PHLMS |
| -.12 | **.55** | .04 | -.10 | -.06 | .03 | I tell myself that I shouldn't feel sad. | PHLMS |
| **-.40** | **.55** | -.06 | -.19 | -.13 | .16 | I think some of my emotions are bad or inappropriate and I shouldn't feel them. | KIMS |
| **-.41** | **.52** | -.04 | -.16 | -.13 | .19 | I tell myself that I shouldn't be thinking the way I'm thinking. | KIMS |
| **-.41** | **.49** | .04 | -.15 | -.09 | .20 | I criticize myself for having irrational or inappropriate emotions. | KIMS |
| -.32 | **.48** | .02 | -.13 | -.01 | .16 | I disapprove of myself when I have irrational ideas. | KIMS |
| -.22 | **.46** | .08 | -.17 | -.15 | -.07 | I find it so unpleasant I have to distract myself and not notice them. | SMQ |
| -.35 | **.46** | -.09 | -.18 | -.10 | .27 | I believe some of my thoughts are abnormal or bad and I shouldn't think that way. | KIMS |
| -.31 | **.45** | .06 | -.16 | .02 | .03 | I judge myself as good or bad, depending on what the thought/image is about. | SMQ |
| **-.43** | **.44** | .11 | -.15 | -.04 | -.08 | I get angry that this happens to me. | SMQ |
| -.14 | **.41** | .11 | -.10 | -.07 | -.15 | I try and push them away. | SMQ |
| -.22 | .38 | .22 | -.04 | .06 | .28 | I make judgments about whether my thoughts are good or bad. | KIMS |
| -.07 | .33 | .17 | -.22 | .00 | .21 | I tend to make judgments about how worthwhile or worthless my experiences are. | KIMS |
| -.26 | .30 | .15 | -.30 | .07 | .03 | I am preoccupied by the past. | CAMS |
| -.15 | .29 | .13 | -.03 | .05 | .03 | I judge the thought/image as good or bad. | SMQ |
| .05 | .20 | .05 | -.04 | .01 | .14 | When I do things, I get totally wrapped up in them and don't think about anything else. | KIMS |
| -.03 | -.13 | .05 | -.01 | -.09 | -.07 | I generate few novel ideas. | LMS |
| -.02 | .01 | **.67** | .06 | -.02 | .08 | When I walk outside, I am aware of smells or how the air feels against my face. | PHLMS |
| .03 | .00 | **.61** | .08 | -.04 | -.02 | When I shower, I am aware of how the water is running over my body. | PHLMS |
| .03 | .05 | **.59** | .12 | -.04 | .11 | I notice changes inside my body, like my heart beating faster or my muscles getting tense. | PHLMS |
| .03 | .00 | **.58** | .02 | .06 | .10 | I pay attention to sensations, such as the wind in my hair or sun on my face. | KIMS |
| .01 | .01 | **.57** | .01 | .05 | .08 | When I am startled, I notice what is going on inside my body. | PHLMS |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| .03 | .20 | **.57** | .14 | .28 | -.02 | When talking with other people, I am aware of the emotions I am experiencing. | PHLMS |
| -.03 | .03 | **.56** | .03 | .00 | .04 | When I take a shower or bath, I stay alert to the sensations of water on my body. | KIMS |
| .10 | -.02 | **.55** | .00 | -.02 | .04 | I notice changes in my body, such as whether my breathing slows down or speeds up. | KIMS |
| -.06 | .00 | **.52** | -.04 | -.08 | .22 | I pay attention to sounds, such as clocks ticking, birds chirping, or cars passing. | KIMS |
| .08 | -.02 | **.52** | .08 | -.01 | .05 | I notice the smells and aromas of things. | KIMS |
| .09 | .23 | **.51** | .09 | .27 | .01 | I am aware of thoughts I'm having when my mood changes. | PHLMS |
| .02 | .07 | **.50** | .05 | .33 | -.03 | Whenever my emotions change, I am conscious of them immediately. | PHLMS |
| -.02 | -.03 | **.49** | .04 | -.01 | .41 | I notice visual elements in art or nature, such as colors, shapes, textures, or patterns of light and shadow. | KIMS |
| .09 | .12 | **.46** | .11 | .22 | .11 | I am aware of what thoughts are passing through my mind. | PHLMS |
| .15 | .06 | **.46** | -.04 | -.01 | .10 | I pay attention to whether my muscles are tense or relaxed. | KIMS |
| .15 | .09 | **.44** | -.05 | .21 | .06 | I notice when my moods begin to change. | KIMS |
| .00 | .11 | **.43** | .14 | .19 | .07 | When talking with other people, I am aware of their facial and body expressions. | PHLMS |
| .05 | .03 | **.42** | -.14 | .05 | .21 | When I'm walking, I deliberately notice the sensations of my body moving. | KIMS |
| .10 | .16 | **.42** | .03 | .28 | .16 | I pay attention to how my emotions affect my thoughts and behavior. | KIMS |
| .28 | .09 | **.42** | .11 | .03 | .11 | I sense my body, whether eating, cooking, cleaning or talking. | FRBRG |
| .14 | .11 | **.41** | .08 | .30 | .28 | I intentionally stay aware of my feelings. | KIMS |
| .06 | .12 | **.40** | .07 | .10 | .07 | I notice how foods and drinks affect my thoughts, bodily sensations, and emotions. | KIMS |
| -.33 | .31 | .35 | -.13 | -.01 | .01 | I keep thinking about the thought or image after it's gone. | SMQ |
| .08 | .12 | -.28 | -.10 | -.04 | -.18 | I am rarely aware of changes | LMS |
| .17 | -.13 | .10 | **.67** | .12 | .12 | I find myself doing things without paying attention. | MAAS |
| .37 | -.12 | .03 | **.58** | .06 | .04 | I find it difficult to stay focused on what's happening in the present. | MAAS |
| -.10 | .09 | .19 | **-.58** | -.23 | -.07 | When I do things, my mind wanders off and I'm easily distracted. | KIMS |
| -.10 | .00 | .06 | **-.57** | -.11 | -.24 | I am easily distracted. | CAMS |
| .07 | -.20 | .06 | **.56** | .02 | .22 | I rush through activities without being really attentive to them. | MAAS |
| .24 | -.23 | .16 | **.56** | .15 | .02 | It seems I am "running on automatic,"� without much awareness of what I'm doing. | MAAS |
| -.02 | -.05 | .15 | **.54** | .03 | -.05 | I break or spill things because of carelessness, not paying attention, or thinking of something else. | MAAS |
| -.04 | -.10 | .04 | **.53** | .02 | .18 | I snack without being aware that I'm eating. | MAAS |
| .04 | -.17 | .03 | **.51** | .04 | -.11 | I find myself listening to someone with one ear, doing something else at the same time. | MAAS |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| -.27 | .23 | .02 | **-.50** | -.21 | -.06 | I don't pay attention to what I'm doing because I'm daydreaming, worrying, or otherwise distracted. | KIMS |
| .01 | -.12 | .19 | **.48** | .02 | -.01 | I drive places on 'automatic pilot' and then wonder why I went there. | MAAS |
| .30 | .07 | .04 | **.46** | .19 | .26 | It is easy for me to concentrate on what I am doing. | CAMS |
| .05 | -.11 | .16 | **.45** | -.01 | -.09 | I do jobs or tasks automatically, without being aware of what I'm doing. | MAAS |
| -.09 | .15 | -.15 | **-.42** | -.19 | .02 | I drive on "automatic pilot" without paying attention to what I'm doing. | KIMS |
| .00 | -.02 | .07 | **.40** | .09 | -.06 | I forget a person's name almost as soon as I've been told it for the first time. | MAAS |
| -.03 | .04 | .18 | -.39 | -.20 | .00 | I tend to do several things at once rather than focusing on one thing at a time. | KIMS |
| .09 | -.24 | .11 | .39 | .32 | -.11 | I could be experiencing some emotion and not be conscious of it until some time later. | MAAS |
| .15 | .05 | .05 | .38 | .09 | .30 | I am able to pay close attention to one thing for a long period of time. | CAMS |
| .35 | -.22 | -.12 | .36 | -.04 | -.07 | I find myself preoccupied with the future or the past. | MAAS |
| .19 | -.20 | .14 | .34 | .00 | .01 | I tend to walk quickly to get where I'm going without paying attention to what I experience along the way. | MAAS |
| .19 | -.16 | .07 | .33 | -.06 | .01 | I get so focused on the goal I want to achieve that I lose touch with what I'm doing right now to get there. | MAAS |
| .12 | .18 | -.09 | .30 | .05 | .11 | When I'm doing something, I'm only focused on what I'm doing, nothing else. | KIMS |
| -.05 | .16 | .17 | -.30 | -.15 | -.13 | When I'm working on something, part of my mind is occupied with other topics, such as what I'll be doing later, or things I'd rather be doing. | KIMS |
| -.05 | -.19 | .21 | .26 | .08 | -.03 | I tend not to notice feelings of physical tension or discomfort until they really grab my attention. | MAAS |
| .02 | .04 | .14 | .23 | .04 | .20 | When I'm reading, I focus all my attention on what I'm reading. | KIMS |
| -.17 | .13 | .00 | -.19 | -.01 | -.09 | I am impatient with myself and with others. | FRBRG |
| -.08 | .18 | .02 | -.16 | **-.76** | -.06 | It's hard for me to find the words to describe what I'm thinking. | KIMS |
| .13 | -.04 | .14 | .14 | **.75** | .21 | I'm good at finding the words to describe my feelings. | KIMS |
| -.08 | .19 | .04 | -.16 | **-.73** | -.08 | I have trouble thinking of the right words to express how I feel about things. | KIMS |
| .12 | -.09 | .17 | .10 | **.70** | .03 | Even when I'm feeling terribly upset, I can find a way to put it into words. | KIMS |
| .14 | -.07 | .26 | .16 | **.67** | .12 | I can usually describe how I feel at the moment in considerable detail. | CAMS |
| .14 | -.11 | .20 | .21 | **.66** | .09 | When someone asks how I am feeling, I can identify my emotions easily. | PHLMS |
| -.07 | .13 | .07 | -.17 | **-.59** | -.03 | When I have a sensation in my body, it's difficult for me to describe it because I can't find the right words. | KIMS |
| .01 | .00 | .24 | -.05 | **.56** | .29 | My natural tendency is to put my experiences into words. | KIMS |
| .24 | .00 | .12 | .08 | **.55** | .16 | I can easily put my beliefs, opinions, and expectations into words. | KIMS |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| .10 | -.05 | .26 | .22 | **.48** | .25 | I'm good at thinking of words to express my perceptions, such as how things taste, smell, or sound. | KIMS |
| .35 | -.06 | .28 | .21 | **.45** | .10 | It's easy for me to keep track of my thoughts and feelings. | CAMS |
| .07 | -.07 | .05 | .17 | .13 | **.59** | I like to be challenged intellectually. | LMS |
| .23 | -.02 | .03 | -.01 | .06 | **.57** | I try to think of new ways of doing things. | LMS |
| .16 | -.18 | .04 | .05 | .10 | **.56** | I find it easy to create new and effective ideas. | LMS |
| .19 | -.12 | -.03 | -.02 | .16 | **.51** | I make many novel contributions. | LMS |
| .17 | .09 | .20 | .02 | -.05 | **.51** | I like to figure out how things work. | LMS |
| .08 | -.12 | .26 | .01 | .07 | **.49** | I am very creative. | LMS |
| .12 | .13 | .25 | .02 | -.07 | **.46** | I like to investigate things. | LMS |
| .05 | .08 | -.09 | -.17 | -.18 | **-.45** | I am not an original thinker. | LMS |
| .07 | .08 | -.10 | -.06 | -.16 | **-.43** | I avoid thought provoking conversations. | LMS |
| .14 | .04 | .32 | -.03 | .04 | **.42** | I am very curious. | LMS |
| -.05 | .23 | .25 | -.04 | .12 | .39 | I tend to evaluate whether my perceptions are right or wrong. | KIMS |
| .00 | -.05 | -.09 | -.05 | -.07 | -.33 | I am rarely alert to new developments. | LMS |
| .17 | .03 | .29 | .21 | .14 | .32 | I pay attention to what's behind my actions. | FRBRG |
| .06 | .15 | .02 | .14 | .06 | .25 | I get completely absorbed in what I'm doing, so that all my attention is focused on it. | KIMS |
| -.03 | .16 | .13 | -.15 | -.12 | .22 | When I'm doing chores, such as cleaning or laundry, I tend to daydream or think of other things. | KIMS |

*Note.* C1 Non-Judgement / Non-Reacting, C2 Observing, C3 Acting with Awareness, C4 Reacting/ Judgement, C5 Describing, C6 Openness / Western Mindfulness.

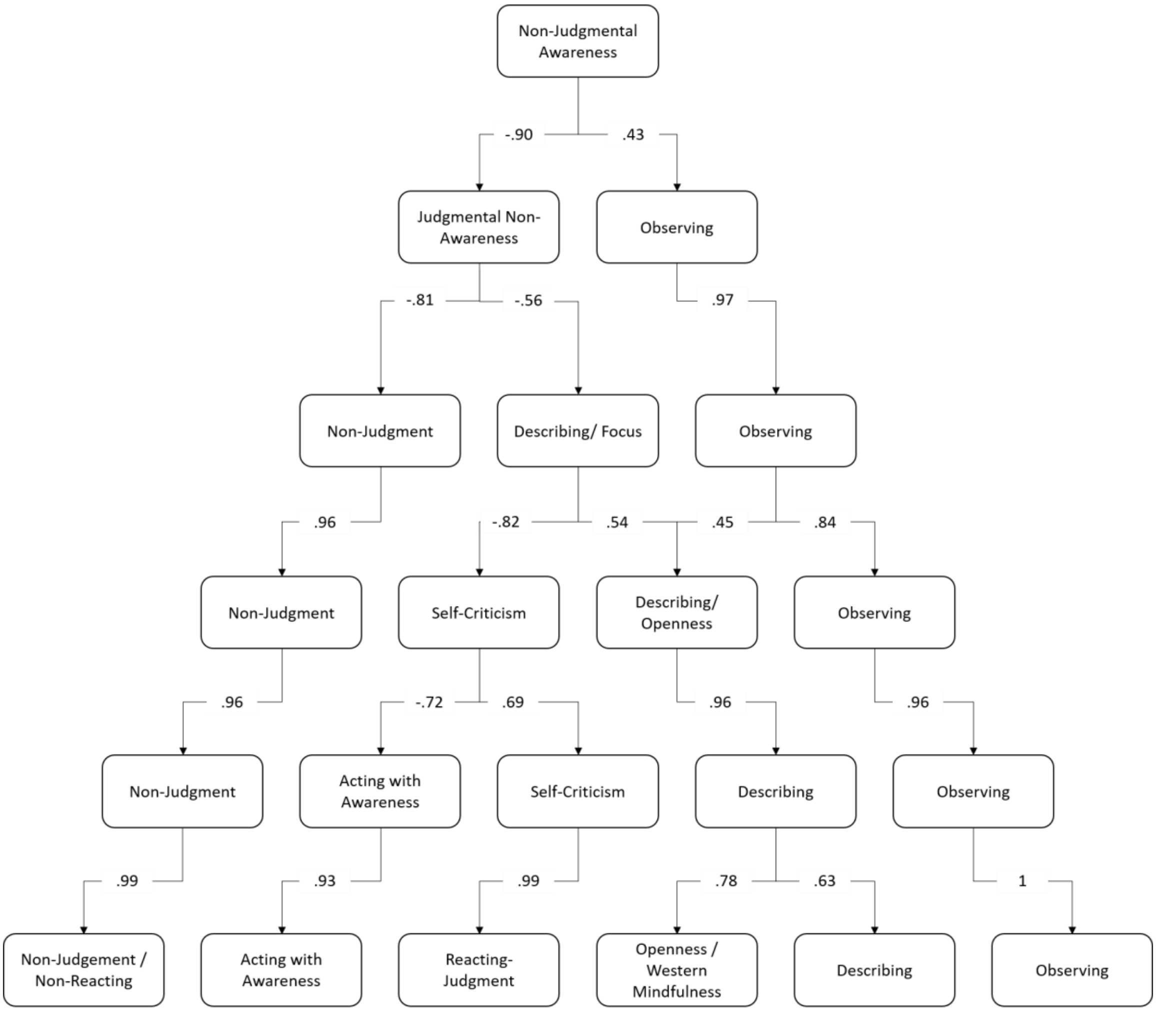



*Figure 3.1 Unfolding structure of the six-component solution.*

Each components scores are correlated with components scores at the prior level. We show correlations $r \geq .40$.

# Chapter 2 Individual Differences and Mindfulness Reconsidered

# Study 3: Individual Differences and Mindfulness[5]

## Preface

In study one I found that current research is strongly focused on outcomes, but less research is focused on potential predictors. In the current study I try to offer a first insight into the question: What might make some people more mindful than others? Previous research has indicated that dispositional mindfulness is related to both the five-factor model of personality (FFM) and the reinforcement sensitivity theory (RST). While these two approaches are both theoretically and empirically related, previous research has examined their associations with trait mindfulness in isolation. This leaves open the question about their unique contributions while accounting for shared effects. In this study I try to join these lines of research to support the theoretical and empirical integration of trait mindfulness into the wider field of individual differences.

[5] This study has been published as a pre-print: Karl, J. A., & Fischer, R. (2019). Individual Differences and Mindfulness. PsyArXiv. https://doi.org/10.31234/OSF.IO/Z2CX6
Minor revisions and stylistic changes have been made to the manuscript to establish coherence with the rest of the thesis

Mindfulness has emerged as an important construct linked to health and well-being (Garland, Beck, et al., 2015). Yet, it is currently unclear how such a broad construct relates to stable individual differences and personality dimensions. We ask: How can we understand mindfulness? Why are some people seemingly more mindful than others? Observations of variability between individuals (Baer et al., 2006) suggest stable individual differences and raise the question of whether we can link mindfulness to a larger network of personality structures to help us understand individual differences in mindfulness. Such an individual difference approach will also open up avenues for addressing larger theoretical questions about mindfulness, including a better understanding the underlying personality characteristics of the mindful person. We examine correlations of the Five-Factor Mindfulness Questionnaire (Baer et al., 2006) with the five-factor model (FFM) of personality (Soto & John, 2017) and reinforcement sensitivity theory (RST; Corr & Cooper, 2016).

**Mindfulness**

Kabat-Zinn defined mindfulness as: “paying attention in a particular way; on purpose, in the present moment, and nonjudgmentally” (Kabat-Zinn, 1994, p. 4). While this definition has been initially applied to states during mindfulness interventions, it has also been widely influential in studies of trait mindfulness (for examples of trait mindfulness studies incorporating this definition see: Medvedev et al., 2016; Tomlinson et al., 2018). Broadly, trait mindfulness has conceptualized as the general tendency of individuals to exhibit mindfulness states (Krägeloh, 2020) similar to comparable distinctions drawn in the whole-trait theory of personality (Fleeson & Jayawickreme, 2015) which views personality traits as descriptors of density distributions of personality states. In line with this conceptualization, measures of trait mindfulness have been found to capture state variance over and above the majority of variance explained by trait differences (Truong et al., 2020). Similarly, change in the distribution of mindfulness states has been shown to result in trait differences over time (Kiken et al., 2015).

Initial measures such as the Mindful Attention and Awareness Scale (Brown & Ryan, 2003) conceptualized trait mindfulness as a uni-dimensional construct. Baer et al. (2006) empirically found five distinct facets of mindfulness: Non-Judging (the ability to interact with one's emotions and thoughts in a non-judgmental way), Non-Reacting (the ability to abstain from immediately reacting to negative stimuli), Acting with Awareness (awareness of one's own moment to moment actions and thoughts), Observing (awareness of sensory perceptions), and Describing (recognizing and labeling one's own emotions). The facets were subsumed under a higher-order factor of mindfulness, although recent research has questioned the utility of a singular higher order structure (Aguado et al., 2015; Tran et al., 2013). These observations raise questions about the conceptual meaningfulness of an overall mindfulness score and point out the need to investigate possibly underlying personality structures to better understand the individual difference component of mindfulness at the facet level.

**Personality traits and mindfulness**

The most commonly used personality theory is the Five Factor model (McCrae & Costa, 1987, 1997) differentiating between five domains: Neuroticism, Conscientiousness, Openness, Agreeableness, and Extraversion. Individuals high in Neuroticism are more likely to experience anxiety and engage in rumination. This may tax processing resources and as a consequence reduce the ability of individuals to engage in present moment awareness and emotion regulation facets of mindfulness. At the same time, this negative emotionality may not interfere with an awareness of internal or external stimuli (Hanley & Garland, 2017; Spinhoven et al., 2017). Individuals high in Conscientiousness are motivated to obtain future rewards through self-discipline and ordered behavior and are better at self-regulating spontaneous impulses, which in turn may facilitate sustained present-moment attention aspects of mindfulness (e.g., acting with awareness, Hanley & Garland, 2017). Openness is associated with curiosity and open-mindedness to new ideas, as well as an interest in philosophical thoughts, which may facilitate both the observation and description of internal mental and emotional states (Hanley & Garland, 2017). The positive emotion-focus and high interpersonal energy component of Extraversion might also increase the ability to both monitor and communicate emotional states in oneself and others. Finally, high Agreeableness might be beneficial for developing dispositional mindfulness due to increased empathy. Agreeableness was associated with emotion regulation (Haas et al., 2007) and empathy (Graziano et al., 2007), and has also been found to be related to all mindfulness facets and strongest to Describing (Hanley & Garland, 2017).

A more biologically oriented personality theory is reinforcement sensitivity theory (RST), which describes plausible underlying motivational systems that are thought to systematically relate to the Five Factors of personality (Corr & Cooper, 2016). There are at least three major systems: the Behavioral Approach System (BAS), which is activated to obtain incentives; the Fight-Flight-Freeze System (FFFS), which is activated in response to immediate aversive stimuli; and the Behavioral Inhibition System (BIS), which is activated to avoid aversive stimuli and to process conflict between different positively and negatively valued goals (e.g., deciding on two equally attractive job offers). Despite an overall similarity, the BIS could be understood as anxiety sensitivity, whereas FFFS as fear sensitivity. Importantly, while both BIS and Neuroticism are related to anxiety, they are thought to work at different temporal points. BIS is responsible for screening for stimuli, judging them, and engaging in fast reactions. If this fails to resolve the perceived threat, slower processes tied to neuroticism are activated to find a resolution for the conflict (Gray, 2004). Both BIS and FFFS might reduce the ability to be non-judgmental and non-reactive to stimuli due to increased monitoring for danger and experienced negative affect. Indeed, BIS and FFFS have been found to be negatively associated with mindfulness (Harnett et al., 2016; Keune et al., 2012; Sauer et al., 2011), specifically with facets that measure emotion and attention regulation (Acting with Awareness, Non-Judging, and Non-Reactivity; Reese et al., 2015).

Individuals with a high level of BAS are better at detecting and acting on reward signals. We could expect that they are also better able to observe and describe both internal and external stimuli (Hamill et al., 2015). There are a number of subsystems within BAS (Reward-Reactivity, Reward-Interest, Goal-Drive Persistence, Impulsivity), but past research has only focused on the overall BAS score. We explore the individual relationships between the subcomponents and mindfulness facets. Using the RST approach, which separates specific motivational systems, together with consistent behavioral predispositions as captured by the FFM measures will allow a better understanding of both motivational systems and regular behavioral consistencies in their relationship with mindfulness as a core emotional regulation process.

## The current study

The aim of the current study is to provide further insight into the relationship of mindfulness with behavioral activation/inhibition and the FFM. We aim to extend previous research that looked at the effect of the FFM and RST on mindfulness separately (Giluk, 2009; Hanley, 2016; Hanley & Garland, 2017), and extend studies which did not directly compare the combined effect of personality and RST (Reese et al., 2015). Using both approaches simultaneously will allow us to differentiate between the unique effects of RST and FFM on mindfulness and achieve a more nuanced understanding of which aspects of mindfulness are linked to global systems of the RST and which are connected to specific effects of the FFM.

# Method

## Participants

Our sample consisted of 404 undergraduate students at a New Zealand university[6]. Five participants (1.24% of the total) started the questionnaire but did not finish it. We removed those participants from the dataset as not enough information was available to impute their data. This left us with an effective sample size of 399 with an average age of 19.21 years (SD = 3.93). Most of our sample was female (*N* =275), and the rest identified as male (*N* =121) or did not specify their gender

[6] This sample overlaps with study 2, and uses data on the FFMQ, personality, and reinforcement sensitivity that were administered in the same survey pack. No overlapping scales beyond demographics are used.

(*N* = 3). Of the total sample 8.77% reported previous mindfulness experience, 9.52% reported yoga experience, and 10.03% reported meditation experience. Participants filled out an online survey on Qualtrics (we provide a text version of the survey on the OSF: https://osf.io/c8bqa/?view_only=96789ad5993e4980987d9014288c94d3). Participants could either come into a lab to answer this survey or complete it online. Participants took part in the current research for course credit as part of an Introduction to Psychology course. Ethical approval was given by the School of Psychology Human Ethics Committee under delegated authority of Victoria University of Wellington's Human Ethics Committee.

**Measures**

**Mindfulness.** The FFMQ (Baer et al., 2006) consists of 39 items which measure five facets of mindfulness. Participants rated their agreement on a 5-point Likert-type scale from 1-(*Never or very rarely true*) to 5-(*Very often or always true*). Example items are "When I'm walking, I deliberately notice the sensations of my body moving" and "I'm good at finding words to describe my feelings".

**Personality.** We used the BFI-2 to assess personality (Soto & John, 2017). The overall scale had 60 items and participants reported their agreement with each item on a 1-(*Disagree strongly*) to 5-(*Agree strongly*) Likert-scale. Example items were "I am someone who is outgoing, sociable" and "I am someone who is compassionate, has a soft heart".

**Behavioral Approach vs Avoidance Motivation.** We used the Reinforcement Sensitivity Theory of Personality Questionnaire (RST-PQ; Corr & Cooper, 2016). The RST-PQ assesses a general BIS factor, a factor measuring Fight, Flight, Freezing Systems (FFFS), and four factors of BAS (Goal-Drive Persistence, Impulsivity, Reward Reactivity, and Reward Interest). All 42 items were measured on a 7-point Likert scale ranging from 1-(*Not at All*), to 7-(*Highly*). Example items are "I am an avoidant sort of person" and "I often find myself not wanting to touch certain objects". Reliabilities for all measures were satisfactory (see Table 4.1). Table 4.2 reports the means and correlation table of the personality variables with mindfulness. A full correlation table is available via OSF: https://osf.io/c8bqa/?view_only=96789ad5993e4980987d9014288c94d3)

**Results**

To examine the combined effects of the FFM and RST on the facets of the FFMQ, we fitted a SEM path model with an MLM estimator (rather than an WMSV estimator as our data was not categorical) to account for potential multi-variate non-normality in the continuous scale scores using lavaan (Rosseel, 2012) in R (R Core Team, 2018). We then bootstrapped the standardized solution of the fitted model using 1000 bootstraps, to obtain bootstrapped standardized effects with 95% bootstrapped confidence intervals and bootstrapped p-values.

Overall, we found a nuanced picture of the relationship between RST, personality, and the facets of mindfulness. Non-Judgement was only predicted by BIS ($\beta$ = -.481[-.614, -.332], $p$ < .001) and BAS-Reward Reactivity ($\beta$ = .168[.048, .278], $p$ =.035). Observing was only associated with Openness ($\beta$ = .366[.260, .456], $p$ <.001). Similar, Non-reacting was only related to Neuroticism ($\beta$ = -.604[-.706, -.491], $p$ <.001). In contrast, Acting with Awareness was significantly predicted by BIS ($\beta$ = -.374[-.524, -.226], $p$ < .001) and Conscientiousness ($\beta$ = .233[.122, .354], $p$ = .001). Describing was predicted by BIS ($\beta$ = -.273[-.423, -.138], $p$ =.011), Extraversion ($\beta$ = .298[.184, .397], $p$ < .001), and Openness ($\beta$ = .237[.146, .331], $p$ < .001). We show the full results of the SEM in Table 4.3. RST and the FFM predicted 36.40% of variance in Non-Judging, 35.30% of variance in Acting with Awareness, 22.00% of variance in Observing, 41.60% of variance in Non-Reacting, and 32.61% of variance in Describing. This pattern suggests that Non-Reacting is the mindfulness facet most closely associated with broad personality traits as measured by the FFM and the RST, whereas Observing is less strongly associated with these personality traits. This indicates that the relative association of RST and the FFM with the FFMQ is highly dependent on the facet under study.

## Discussion

Our study found that different facets of mindfulness seem to be related to different and not necessarily compatible trait dimensions. This draws the unity of the mindfulness construct into question, highlighting the need to investigate the relationship between mindfulness and personality dimensions at a facet level. Our results also show that at the facet level, specific components of mindfulness might be understood as stable personality-like dimensions. The mindfulness facets were largely unconnected to BAS, FFFS. In contrast, the facets were substantially related to BIS, Neuroticism, Extraversion, Conscientiousness, and Openness. In particular, Non-Reacting shows a clear highly negative relationship with Neuroticism, and BIS a highly negative relationship with Non-Judging. Considering how mindfulness might be part of personality can provide a way forward to develop more explicit theories about the origins of individual differences in mindfulness. Below we outline possible pathways of how biologically-rooted individual differences might contribute to individual differences in dispositional mindfulness. We identified three major patterns:

First, one of the major differences we found was the differential effect of RST and FFM on Non-Reacting and Non-Judging. Non-Reacting was predicted by Neuroticism but not BIS, and Non-Judging was predicted by BIS but not by Neuroticism. This finding is in line with the RST that views BIS as responsible for threat assessment and Neuroticism as shaping reactions to non-immediate threats (Corr & McNaughton, 2012; Gray, 2004). Interestingly, this also suggests a potential temporal order of the mindfulness processes with Non-Judging preceding Non-Reacting. This would explain recent findings that show that Non-Judging, but not Non-Reacting, is negatively related to false alarm rates in attention trials (Cosme & Wiens, 2015; Rosenstreich & Ruderman, 2017). Overall, the finding that BIS and Neuroticism correlated with different facets of mindfulness reflects their unique roles according to RST and offers potential new insight into how mindfulness facets can be integrated into biologically-mediated models.

Second, we found that Openness predicted both Observing and Describing. Describing was positively related to Extraversion. One possible reason for this relationship might be that extroverts tend to talk more about abstract concepts (Beukeboom et al., 2013) and use more positive and less negative emotion words (Pennebaker & King, 1999; Yarkoni, 2010). Simply put, highly extravert individuals might be better at verbalizing and describing their emotions, because they have more daily practice doing so in their social environment. Additionally, Extraversion and Openness form a plasticity meta-factor which is thought to allow for malleable situation-appropriate behavior (DeYoung, 2015). Plasticity has also been found to be related to Describing and Observing (Hanley et al., 2018).

Our findings lend further support to the hypothesis that plasticity as a meta-factor is an important predictor of Observing and Describing. By then including the RST we can make some finer distinctions between Observing and Describing. Similar to other studies that included only BIS (Reese et al., 2015), we found that BIS predicts Describing but not Observing. One reason for this might be that both Describing and Observing are relevant for monitoring situational cues (internal and external, respectively). High BIS might negatively impact Describing (internal monitoring & labeling of experiencing) because attention is shifted towards monitoring external cues that might be relevant to avoid threat. A bias in the locus of attention would also explain why some studies have found that BIS is positively related to Observing (Hamill et al., 2015). High BIS individuals are motivated to detect threats (Klackl et al., 2018), and are more likely to pay attention to external stimuli such as sounds, smells, and tactile sensations. The greater awareness of such stimuli might result in greater Observing scores. To summarize, our findings indicate that while both Observing and Describing express monitoring of situational cues, high BIS results in a bias towards external monitoring, reducing the ability to monitor and describe internal states.

This pattern is also congruent with our finding that BIS negatively predicts Acting with Awareness. Acting with Awareness reflects “meta-awareness” (the ability to be aware that one is aware; Hargus et al., 2010; Seli et al., 2017). High BIS might impede such meta-awareness, similar to Describing, by shifting attentional resources to external stimuli. This would also explain the positive relationship between Conscientiousness and Acting with Awareness. The self-regulatory aspect of Conscientiousness necessitates attention to be allocated to monitor one's behavior and cognitions for them to be effectively regulated (Koestner et al., 1992; McCrae & Löckenhoff, 2010).

### Limitations

Our current results are based on students with low meditation experience. Our study also uses cross-sectional data which precludes any claims to causality. Nevertheless, our trait predictors are relatively stable individual difference variables that are partially rooted in genetical and neuroanatomical processes (Corr & Cooper, 2016; Gray, 1970), suggesting a path from personality to mindfulness. Future studies could employ a longitudinal approach to explicitly test directions of causality.

## Conclusion

At the beginning of the current article we asked:" Why are some people seemingly more mindful than others?" Our results indicate two main conclusions. First, the why is dependent on the facet under study as can be seen with the differential effect of BIS and Neuroticism on Non-Reacting versus Non-Judging. Focusing on those finer grained differences between the facets allows us to draw more nuanced conclusions about the place of mindfulness in the network of individual differences. Second, we found a pattern of relationships between RST, FFM, and mindfulness that suggests that locus of attention plays a major role. Dimensions of RST and FFM that draw attention to external stimuli substantially and negatively correlated with all mindfulness facets that deal with internal monitoring. In contrast, dimensions such as Openness and Conscientiousness might predispose individuals to allocate attention to internal stimuli, either to explore them (Openness) or to regulate them (Conscientiousness).

*Table 4.1 Reliability of the individual measures of the current study.*

| | $\alpha$ | $\alpha_{low}$ | $\alpha_{high}$ | $\omega$ | $\omega_{low}$ | $\omega_{high}$ | GLB | H |
|---|---|---|---|---|---|---|---|---|
| **Mindfulness** | | | | | | | | |
| Observing | .79 | .76 | .82 | .79 | .76 | .82 | .84 | .81 |
| Non-Reacting | .84 | .82 | .87 | .85 | .82 | .87 | .87 | .85 |
| Acting with Awareness | .88 | .86 | .90 | .88 | .86 | .90 | .91 | .89 |
| Non-Judging | .93 | .92 | .94 | .93 | .92 | .94 | .93 | .94 |
| Describing | .91 | .90 | .93 | .91 | .90 | .93 | .91 | .93 |
| **Personality** | | | | | | | | |
| Extraversion | .86 | .84 | .88 | .86 | .84 | .88 | .89 | .88 |
| Agreeableness | .77 | .74 | .80 | .76 | .73 | .80 | .86 | .82 |
| Conscientiousness | .84 | .81 | .86 | .84 | .82 | .86 | .90 | .86 |
| Neuroticism | .90 | .89 | .92 | .91 | .89 | .92 | .93 | .91 |
| Openness | .84 | .81 | .86 | .84 | .82 | .86 | .90 | .85 |
| **Reinforcement Sensitivity** | | | | | | | | |
| Behavioral Inhibition | .94 | .94 | .95 | .95 | .94 | .95 | .95 | .95 |
| Fight-Flight-Freeze | .78 | .75 | .82 | .79 | .75 | .82 | .86 | .83 |
| Goal-Drive Persistence | .87 | .86 | .89 | .88 | .86 | .90 | .92 | .89 |
| Impulsivity | .75 | .71 | .78 | .74 | .70 | .78 | .80 | .84 |
| Reward Interest | .79 | .76 | .82 | .79 | .76 | .82 | .86 | .82 |
| Reward Reactivity | .86 | .84 | .88 | .86 | .84 | .88 | .91 | .87 |

*Note*: In addition to α we report ω, which does not assume tau-equivalence of item-loadings, and the Greatest Lower Bound, which aims to maximize the error components to obtain the lowest possible reliability, (GLB; for a recent study on the advantages of using ω and GLB over α see: Trizano-Hermosilla & Alvarado, 2016) Low and high values for α and ω represent the 95% confidence interval, GLB = Greatest Lowest Bound, H = Coefficient H

*Table 4.2 Correlation of the FFMQ with individual difference variables.*

| | M | SD | Acting with Awareness | Non-Judging | Describing | Observing | Non-Reacting |
|---|---|---|---|---|---|---|---|
| | | | **Mindfulness** | | | | |
| Acting with Awareness | 2.82 | 0.70 | | | | | |
| Non-Judging | 3.05 | 0.92 | .30** | | | | |
| Describing | 3.07 | 0.87 | .40** | .27** | | | |
| Observing | 3.22 | 0.70 | .11* | -0.03 | .23** | | |
| Non-Reacting | 2.78 | 0.69 | .22** | .31** | .27** | .12* | |
| | | | **Reinforcement Sensitivity** | | | | |
| Fight-Flight-Freeze | 3.76 | 1.08 | -.24** | -.26** | -.21** | -0.08 | -.25** |
| BIS | 4.45 | 1.15 | -.42** | -.56** | -.39** | 0.07 | -.48** |
| Reward Interest | 4.55 | 1 | .12* | 0.08 | .20** | .24** | .17** |
| Goal-Drive Persistence | 5.03 | 1.04 | .28** | .17** | .28** | .17** | .13** |
| Reward Reactivity | 5.11 | 0.96 | 0.02 | .16** | .22** | .24** | .10* |
| Impulsivity | 4.44 | 1.02 | -.29** | -0.03 | -0.04 | 0.06 | 0.05 |
| | | | **Personality** | | | | |
| Extraversion | 3.27 | 0.72 | .14** | .23** | .43** | 0.08 | .20** |
| Agreeableness | 3.6 | 0.58 | .15** | .14** | .09+ | .13** | -0.01 |
| Openness | 3.73 | 0.65 | .21** | 0.05 | .30** | .41** | 0.05 |
| Neuroticism | 3.2 | 0.86 | -.32** | -.48** | -.30** | 0.06 | -.63** |
| Conscientiousness | 3.2 | 0.68 | .39** | .18** | .20** | .14** | 0 |

*Note:* $^{*}p < .05$, $^{**}p < .01$, $^{***}p<.001$, Due to space constraints we only report the correlation of the predictors with the FFMQ, a full correlation table is available from the OSF.

*Table 4.3 FFM-Facets and RST predictors of the FFMQ*

| | Observing | Describing | Non-Judging | Non-Reacting | Acting with Awareness |
|---|---|---|---|---|---|
| BIS | .034[-.129, .191] | -.273[-.413, -.138]* | -.481[-.614, -.332]*** | -.085[-.233, .062] | -.374[-.524, -.226]*** |
| FFFS | -.079[-.190, .033] | .011[-.089, .111] | -.014[-.109, .078] | .007[-.081, .103] | -.007[-.110, .095] |
| **BAS** | | | | | |
| Goal-Drive Persistence | -.138[-.281, .003] | .102[-.053, .245] | .012[-.115, .141] | .087[-.028, .210] | .126[-.012, .257] |
| Impulsiveness | .086[-.032, .205] | -.073[-.186, .048] | .056[-.056, .168] | .055[-.055, .158] | -.150[-.264, -.027] |
| Reward Interest | .050[-.099, .192] | -.176[-.300, -.046] | -.176[-.301, -.045] | .028[-.108, .159] | .024[-.096, .152] |
| Reward Reactivity | .187[.043, .330] | .112[-.023, .251] | .168[.048, .278]* | -.060[-.179, .063] | -.077[-.209, .058] |
| **FFM Factors** | | | | | |
| Agreeableness | -.010[-.109, .093] | -.039[-.147, .066] | .085[-.009, .177] | -.048[-.137, .046] | -.003[-.101, .086] |
| Conscientiousness | .164[.051, .274] | .054[-.062, .170] | .109[-.014, .231] | -.084[-.184, .030] | .233[.122, .354]** |
| Extraversion | -.060[-.201, .071] | .298[.184, .397]*** | -.051[-.169, .070] | -.098[-.204, .009] | -.069[-.173, .042] |
| Neuroticism | .091[-.034, .231] | .016[-.117, .143] | -.121[-.252, .009] | -.604[-.706, -.491]*** | -.024[-.155, .103] |
| Openness | .366[.260, .456]*** | .237[.146, .331]*** | .032[-.055, .117] | .061[-.023, .144] | .127[.031, .230] |

*Notes*. All values are given are based on the standardized solution computed with 1000 bootstraps. Confidence intervals are based on the empirically observed distribution across the 1000 bootstraps. Significance is based on *p* values, which are computed independently from the confidence intervals of the estimates. *** *p* < .001, ** *p* <.01, * *p* <.05

## Study 4: The Development of Mindfulness in Young Adults: The Relationship of Personality, Reinforcement Sensitivity, and Mindfulness[7]

### Preface

This study builds directly on study three. While in study three I presented an initial investigation into the combined effects of reinforcement sensitivity and big five personality traits on trait mindfulness, all directional claims were based on theoretical assumptions. The following study aims to address this limitation by using a longitudinal approach and to explicitly examine directional relationships between the constructs over time. This study extends the contribution of study three to the literature by providing the first longitudinal data on the relationship between reinforcement sensitivity, big five personality and trait mindfulness.

[7] This study has been published: Karl, J. A., Fischer, R., & Jose, P. E. (2021). The Development of Mindfulness in Young Adults: The Relationship of Personality, Reinforcement Sensitivity, and Mindfulness. Mindfulness. https://doi.org/10.1007/s12671-020-01576-3 Minor revisions and stylistic changes have been made to the manuscript to establish coherence with the rest of the thesis.

Mindfulness practitioners argue that mindfulness needs to be cultivated through careful and prolonged practice and may not change without focused interventions (C. Kang & Whittingham, 2010). At the same time, the ability to pay attention to the present, in a non-judgmental way may be part of a larger personality trait complex. Indeed, mindfulness could be thought of as being systematically related to personality dimensions focused on emotion regulation, attention to detail and openness to sensual experiences (emotional stability, conscientiousness & openness), associations that have been reported in previous studies

The five factor model (FFM) of personality (McCrae & Costa, 1997) proposes that behavioral differences between individuals are organized along five higher order dimensions: Openness (a desire for stimulating external and internal experiences), Conscientiousness (the ability to delay rewards & follow rules), Extraversion (the desire for social interaction), Agreeableness (the ability to maintain social relationships through empathy, respectfulness, and trust), and Neuroticism (the tendency to experience anxious or depressive moods, as well as emotional volatility).

These five dimensions are thought to be rooted in global neurobehavioral systems that vary between individuals, such as sensitivity to reward and punishment proposed in the revised Reinforcement Sensitivity Theory (Corr et al., 2013). This theory proposes a number of basic systems, including the Behavioral Approach System (BAS), the Behavioral Inhibition System (BIS) and a Fight-Flight-Freeze System. These systems are thought to be driven by genetically derived predisposition which are in turn being calibrated and updated in response to environmental stimuli during crucial phases of a person's development (Corr & Matthews, 2020). The BAS orients individuals towards rewards and aids in obtaining those rewards through four main pathways: Reward Interest, Reward Reactivity, Goal-Drive Persistence, and Impulsivity. Individuals can be habitually motivated to direct their attention to potential rewards by either increasing the physical pleasure felt from rewards (Reward Reactivity) or by expressing an increased interest in rewards, which is associated with greater exploration of possibly rewarding stimuli (Reward Interest). There are also individual differences in the tendency to seek to obtain temporally or spatially immediate rewards, labelled

Impulsivity, or delay obtaining immediate rewards to obtain larger rewards in the future, labelled Goal-Drive Persistence (Corr & Matthews, 2020; Corr & McNaughton, 2012). In contrast, the BIS is characterized by greater punishment sensitivity and monitoring of possible goal conflict, commonly expressed as worry about future threats and anxiety (Corr & Cooper, 2016; Gray, 1970). Although there are currently no direct measures of BIS/BAS at physiological or neurobiological levels, there are validated instruments that allow self-report measurement of the psychological and behavioral expressions of these two systems.

The interplay between these two major systems are thought to give rise to the behavioral traits as captured by the Big Five. Specifically, Neuroticism is thought to originate from differences in the sensitivity of the BIS. In contrast, Extraversion is thought to originate from individual differences in the sensitivity of the different BAS components (Gray, 1970). These two personality traits have been most consistently associated with BIS/BAS, but the other three Big Five traits have also been linked to BIS/BAS in theory (Corr et al., 2013; Fischer, 2017) and through empirical investigations (Fischer & Karl, 2020). Openness summarizes aspects of personality aimed at obtaining external rewards (e.g., money) and internal rewards (e.g., positive affect, excitement) and has been linked to BAS Reward Interest/ Reactivity (Corr et al., 2013). Similarly, Conscientiousness, aimed at obtaining temporally distant rewards, is thought to be related to high BAS Goal-Drive Persistence (Corr & Cooper, 2016). Finally, Agreeableness is more complex in its relationship to BIS/BAS (Corr et al., 2013) and might be linked to both BIS and BAS through an interplay of maintaining rewarding social connections (related to BAS), but also the need for restricting one's behavior to fit social rules (a core component of BIS, see Fischer, 2017). Evidence seems to suggest relative consistent positive relations of Agreeableness with BAS, but both positive and negative correlations with BIS, which could be due to instrument and sample effects (Corr & Cooper, 2016; Smits & Boeck, 2006). Importantly, while BIS/BAS are thought to underly individual differences in the Big Five, they do not completely overlap (Corr et al., 2013; Fischer & Karl, 2020). For example, while BIS is thought to underly Neuroticism, it is more closely linked to the anxiety and rumination components of Neuroticism, rather than to the emotional volatility aspects (Slobodskaya & Kuznetsova, 2013). Therefore, including both models of personality simultaneously allows us to examine their unique temporal associations with dispositional mindfulness, which can also open opportunities for a better targeting of future mindfulness interventions.

To date, the relationship between mindfulness, personality, and reinforcement sensitivity have been examined in cross-sectional studies (e.g. Giluk, 2009; Hanley & Garland, 2017; Karl & Fischer, 2019), which unfortunately do not allow any inferences about temporal relationships and possible causality. Currently, personality researchers tend to make claims about directionality (e.g., BIS influencing Non-Judging) based on theoretical assumptions that responses to BIS/BAS scales represent underlying neuroanatomical differences in emotion and cognition (Dolatyar & Walker, 2020). Consequently, correlations of BIS or BAS with mindfulness are interpreted as the influence of more basic neurobiological systems on mindfulness-based emotional experience and response (Corr & Cooper, 2016). Given the presumed biological determination of the Big Five (McAdams & Pals, 2006), similar claims could be made about correlations of mindfulness with Big Five measures. Seen from this perspective, personality changes are thought to be driving changes in mindfulness over time.

Recent theories and empirical evidence suggest that personality is dynamic and responds to situational and developmental cues, with major transitions taking place during early adulthood (Bleidorn et al., 2013; Kandler, 2012; Roberts et al., 2006). Early adulthood is a particularly interesting period because it is marked by a convergence of biological changes, normative expectations, and social maturation effects related to identity and self-perceptions (Fischer, 2017; Roberts & Davis, 2016). First, early adulthood is a time of substantial neurobiological development involving a wide net of neurological systems assumed to underlie personality dynamics (Costa & McCrae, 2006). Second, during early adulthood individuals typically assume new social and work roles in their life, moving away from their parents to further their education (study) or secure paid employment, seek to find a stable life partner, and start forming a family. These new roles are accompanied by a range of expectations about appropriate behavior. Conforming to these normative expectations tends to be rewarded (e.g. academic achievement), whereas failure to adhere role-related expectations might be punished (e.g. social exclusion). Conforming to these social norms and associated expectations is typically linked to changes in personality traits in a socially-normative fashion (Bleidorn et al., 2013; Lodi-Smith & Roberts, 2016). Finally, changes in self-perceptions and identity development during this period might play a substantial role in trait changes. Transitory periods with challenges and associated success or failure to adjust to those challenges may create feedback loops that shape personality. For example, if a university student views themselves as studious and their studies are central to their identity, they are likely to dedicate an increased time to study. In case of academic achievements, this view might be validated and then reinforces an emerging trait-like behavioral predisposition of working hard (Göllner et al., 2017). Alternatively, failure may lead to a need to re-evaluate self-perceptions and associated behaviors (Roberts & Davis, 2016).

Wrzus and Roberts (2017) developed the TESSERA framework to explain the mechanisms of personality change. Their model distinguishes short, medium and long-term developmental processes. The period when individuals are leaving their parental home and assume new roles and responsibilities as university students is exceptionally rich in processes that are likely to induce personality processes that will lead to personality changes in the short-term, which can then crystallize and become stabilized in the following years. Crucial processes in this period include increased personal reflections in reaction to environmental stimulation, role modelling by peers, feedback and reinforcement of novel behaviors in response to changed roles and responsibilities, self-regulation, accommodation and assimilation to new behavioral demands.

In line with these theoretical processes, Roberts et al. (2006) reported a meta-analysis of 92 longitudinal studies (total N = 50,120) which indicated that our focal age of 18 to 22 years was characterized by some of the most profound personality trait changes observable during the life span. The largest developmental changes were observed for aspects of Extraversion and Openness, and the weakest and non-significant changes for this age group were observed for Agreeableness and Conscientiousness. Personality may even systematically change over a period of two weeks provided individuals are motivated to change their personality and are provided with daily reminders (Stieger et al., 2020). In summary, personality during early adulthood is relatively malleable and likely to change depending on the individual's social environment and normative role expectations. Given these personality dynamics during this age period, the current research aims to examine how personality traits and mindfulness may relate to each other during a formative period of young people's lives.

In contrast to personality traits, dispositional mindfulness has been conceptualized as a relatively stable individual difference variable, but with some malleability due to practice and life events (Baer et al., 2008). Young adults might experiment with alternative life experiences and may take up new habits while transition from home to a more independent adult life. The opportunities and challenges during the first year of university create multiple situations in which individuals need to regulate their emotional and behavioral impulses and may become aware of their thoughts, feelings, and perceptions to successfully navigate novel environments. Given the above-noted processes at biological, social, and subjective levels that have a likely impact on personality development, one might expect that similar changes could occur for individual differences in dispositional mindfulness. Research on the developmental trajectories of mindfulness has largely focused on the role of attachment styles (Stevenson et al., 2017). Less is known about other developmental influences on mindfulness. Given the well-documented patterns of change in personality in young adults (Roberts & Davis, 2016) and the consistent cross-sectional link of personality and reinforcement sensitivity traits to mindfulness (Hanley et al., 2018; Reese et al., 2015), it is especially important to examine the role of personality and reinforcement sensitivity in the development of mindfulness during early adulthood.

The BIS system is geared towards the evaluation of risk and the rapid activation of (emotional) reaction without necessarily passing through conscious awareness. Hence this automatic evaluation component would interfere with the various facets of mindfulness that require conscious awareness, acceptance of negative thoughts, and downregulating strong emotional responses. In line with these expectations, previous cross-sectional studies (Karl & Fischer, 2019; Reese et al., 2015) reported negative correlations between BIS and mindfulness facets expressing present-moment awareness of one's behavior and emotions, as well as judgment of one's emotional reactions. This suggests that individuals with higher activation of BIS tend to engage in fast, automatic behavior aimed at avoiding negative stimuli due to increased punishment sensitivity (Gray, 2004; Keune et al., 2012), which conflicts with the conscious processing necessary to be aware of one's emotions and actions and, crucially, to non-critically consider them.

In contrast, highly Reward-Reactive individuals (part of BAS) are generally oriented towards experiencing positive stimuli and emotions (Corr & Cooper, 2016), which in turn should decrease their focus on negative stimuli and making it easier for them to be non-judgmental towards their own experiences. Patterns in line with these theoretical predictions have been observed in previous cross-sectional studies (Dolatyar & Walker, 2020; Karl & Fischer, 2019).

Focusing on the Big Five, individuals high in Conscientiousness are typically motivated to obtain future rewards through self-discipline and restraining rash impulses that may interfere with longer term-goals. These behavioral tendencies necessitate sustained present-moment attention. In line with this reasoning, Acting with Awareness has been found to be positively related to Conscientiousness in cross-sectional observations (Haliwa et al., 2020). Openness represents a drive towards curiosity and being open-minded about new ideas, as well as an interest in deep philosophical thoughts (Soto & John, 2017). Therefore, individuals higher in openness might be more likely to engage with and pay attention to both internal and external stimuli, as this is consistent with the greater sensitivity towards sensory and intellectual stimulation. Conforming these patterns, mindfulness dimensions expressing openness and awareness towards both internal (Describing) and external stimuli (Observing) have been consistently found to be positively related to Openness (Spinhoven et al., 2017). Neuroticism expresses emotional fragility and is characterized by emotional volatility, depression, and anxiety (Soto & John, 2017) and overlaps strongly with BIS (Fischer & Karl, 2020). Individuals higher on Neuroticism might find it difficult to remain non-reactive towards negative events and regulate negative emotions because they are more sensitive to emotions and more likely to experience emotional disturbances. In line with this, Non-Reacting has shown negatively correlations with Neuroticism (Hanley et al., 2018; Karl & Fischer, 2019).

Extraversion captures both a high energy state in social contexts and greater orientation towards positive emotionality. The high sociability of Extraversion is expressed in the tendency of extraverts to talk more about abstract concepts and emotions (Beukeboom et al., 2013) and use more positive and less negative emotion words (Pennebaker & King, 1999; Yarkoni, 2010). In order to be able to express emotion terms in social interactions in both abstract and situation appropriate ways, it is necessary for extraverts to accurately perceive and describe their emotions. In line with this, past cross-sectional research reported positive correlations between Extraversion and Describing (Haliwa et al., 2020; Karl & Fischer, 2019).

Our current study extends previous studies by: a) explicitly testing the links among mindfulness, Big Five, and BIS/BAS using b) longitudinal data during early adulthood and c) accounting for inter-individual differences in personality and mindfulness, meaning that our patterns can be interpreted as intra-individual changes over time. Our hypothesis is that changes in personality will be associated with changes in mindfulness over time. Specifically, we hypothesized that over time BIS is associated with decreased mindfulness, especially Non-Judgement, Acting with Awareness and Describing (Hypothesis 1), BAS is associated with increasing Non-Judgement (Hypothesis 2), Consciousness is associated with increasing Acting with Awareness (Hypothesis 3), Openness is associated with increasing Describing and Observing (Hypothesis 4), Neuroticism is associated with decreasing Non-Reacting (Hypothesis 5) and Extraversion is associated with increasing Describing (Hypothesis 6).

## Methods

### Participants

The sample were first year students enrolled in a New Zealand University. They voluntarily took part in the study in exchange for course credits as partial fulfilment of course requirements in a two-part introductory course to psychology. Our final sample was representative of this group in age (*M* = 18.57, *SD* = 2.39) and gender (75.33 % female).

**Procedure**

The current study was offered at four time-points throughout the academic year to all enrolled psychology students, but participation in each wave was voluntary and not contingent on previous participation. Because of the way that the degree is structured, this procedure resulted in different participant pools at each time-point, with changing populations across the four-month period depending on enrolment patterns. The current study only includes participants that filled out the three-initial time-points. Time-point one (*N* = 715) was collected at the start of the academic year. Time-point two (*N* = 604) was collected about two months later (*M* = 57 days, *SD* = 5). Finally, time-point three (*N* = 617) was collected at the start of the second trimester (*M* = 70 days, *SD* = 5). Across these three time-points 227 participants were matched using their unique student identification numbers, representing the current sample of the study. We also collected an additional fourth time-point (*N* = 151) at the end of the second trimester, but only 50 participants were matched across all time-points due to overall low response rate at this final time-point (which coincided with exam periods and reduced participation rates). Therefore, the fourth time-point was omitted from our analysis, but the data are available on the OSF for interested readers together with the data of all participants who were not matched across the initial three time-points. The time period (Time-point 1 to Time-point 3) captures the crucial first four months of university life of a cohort of young adults.

**Measures**

**Mindfulness.** Dispositional mindfulness was measures using the FFMQ-SF (Bohlmeijer et al., 2011). This scale measures the five facets of mindfulness using 24 items on a 1 (*Never or very rarely* *true*) to 5 (*Very often or always true*) Likert scale. Example items are "I'm good at finding words to describe my feelings." or "I can easily put my beliefs, opinions, and expectations into words".

**Personality.** To assess the Big Five personality structure the BFI-S was used (Soto & John, 2017). This scale measures the five personality dimensions with three subscales each. The overall scale is composed of 30 items and participants reported their agreement with each item on a 1 (*Disagree strongly*) to 5 (*Agree strongly*) Likert scale. Example items are "I am someone who is outgoing, sociable" (Extraversion) and "I am someone who is compassionate, has a soft heart" (Agreeableness).

**Behavioral Approach vs. Avoidance Motivation.** We used the Reinforcement Sensitivity Theory Personality Questionnaire (RST-PQ: Corr & Cooper, 2016). The RST-PQ assesses a general BIS factor, a factor measuring Fight, Flight, Freezing Systems (FFFS), and four factors of BAS (Goal-Drive Persistence, Impulsivity, Reward Reactivity, and Reward Interest). All 42 items are measured on a 7-point Likert scale ranging from 1 (*Not at All*) to 4 (*Highly*). Example items are "I am an avoidant sort of person" (BIS) and "I often find myself not wanting to touch certain objects" (BIS).

All our measures yielded acceptable α and ω reliability across the three waves and showed at least metric invariance across waves (Full details of the analysis of temporal invariance are given in Appendix C). The internal reliabilities of the measures can be found in Table 5.1.

**Data Analysis**

We investigated the longitudinal relationships among mindfulness, personality, and behavioral inhibition/activation using a Random Intercept Cross-Lagged Panel Model (RI-CLPM, for an introduction to this model see: Hamaker et al., 2015). This model has been shown to allow for more accurate estimations of the random effects by allowing random intercepts to be estimated for subjects, which parcels out between-subjects variance. Our analysis therefore controls for individual differences between individuals and only models within-person temporal changes, independent of stable trait-like individual differences, a characteristic particularly relevant for a longitudinal study of personality traits (Mund & Nestler, 2019). For comparison purposes, we also ran this model as regular CLPM by forcing the variances and covariances of the random intercepts to 0. This model showed substantially worse fit compared to the RI-CLPM ($F(136) = 440.68$, $p < .001$), and this result supports our choice of the RI-CLPM analytical model. Nevertheless, we report the CLPM on the OSF for the interested reader. For each wave, the five mindfulness facets (Observing, Describing, Acting with Awareness, Non-Judging, and Non-Reacting), the five dimensions of personality (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism), the BIS, Fight Flight Freeze Sensitivity, and the four sub-facets of BAS according to the revised RST were included. A simplified conceptual representation of the model adopted from Hamaker et al. (2015) can be found in Figure 5.1. Due to the option for participants to not respond to individual items at waves 1 and 3 each model was fitted with full information maximum likelihood (FIML) to replace missing data and a maximum likelihood estimator with robust (Huber-White) standard errors (MLR) to adjust for multi-variate non-normality. As specified in our pre-registration, all auto-regressive and cross-lagged effects were included. This resulted in a very conservative model which might lack the power to detect smaller effects (Masselink et al., 2018). Therefore, we also ran separate models for mindfulness and personality, mindfulness and reinforcement sensitivity. We report these models on the OSF.

In our model all variables at one timepoint predicted all variables at the next time point. This model allows us to covary out between-subject variation and only examine the within-subjects effects of the variables over time (Hamaker et al., 2015). Overall, all regression coefficients reported by this model represent change over time, rather than stability or rank order. Prior to testing the hypotheses, we investigated whether the instrument properties were invariant across time points (see: Fischer & Karl, 2019). Because our data were found to be metrically equivalent over time (identical factor loadings of items on latent variables across time), we decided to constrain all paths (auto-regressive and cross-lagged) to be equal across timepoints to reduce the interpretative complexity.

## Results

We first examined the rank-order stability over time as well as the relative changes from wave to wave, which are reported in Table 5.2. As can be seen there, the average stability for the Big Five traits is quite high, ranging on average from .67 for Agreeableness to .83 for Extraversion. For RST, the rank-order stability is somewhat lower, ranging from an average of .60 for BAS-Reward Reactivity to .70 for the Fight-Flight-Freeze-System. Focusing on the important results related to mindfulness facets, the overall rank-order stability was somewhat lower, ranging from .50 for Non-Judging to .64 for Describing. Therefore, mindfulness appears to be more flexible and dynamic compared to personality traits, especially within the Big Five tradition. The cross-sectional correlation between the measures in the full samples at each wave, together with descriptive information on mean and standard deviation, is shown in Tables 1 to 3 in Appendix C.

We then examined the wave-to-wave mean changes within our sample. For mindfulness, Non-Reacting increased from time 2 ($M_{T2}$ = 2.896) to time 3 ($M_{T3}$ = 2.982, $p < .05$; $d$ = -.121). Non-Judgement also changed significantly from time 2 to time 3 ($M_{T2}$ = 2.813; $M_{T3}$ = 2.978, $p < .01$, $d$ = -.209). For the Big Five, the means of Agreeableness ($M_{T1}$ = 3.726, $M_{T2}$ = 3.810, $M_{T3}$ = 3.704; $p < .05$; $d_{T1\text{-}T2}$= -.134; $d_{T2\text{-}T3}$ = .158) and Extraversion ($M_{T1}$ =3.102, $M_{T2}$ = 3.167, $M_{T3}$ = 3.099; $p < .05$; $d_{T1\text{-}T2}$= -.083; $d_{T2\text{-}T3}$ = .089) changed over time. For RST, the means of BIS ($M_{T1}$ = 2.707, $M_{T2}$ = 2.648; $p < .05$; $d_{T1\text{-}T2}$= .108), BAS-Impulsivity ($M_{T1}$ = 2.589, $M_{T2}$ = 2.510; $p < .01$; $d_{T1\text{-}T2}$= .139), BAS-Reward Reactivity ($M_{T1}$ = 2.848, $M_{T2}$ = 2.785; $p < .05$; $d_{T1\text{-}T2}$= .132), BAS-Reward Interest ($M_{T1}$ = 2.649, $M_{T2}$ = 2.543; $p < .001$; $d_{T1\text{-}T2}$= .184), BAS-Goal Drive Persistence ($M_{T1}$ = 2.941, $M_{T2}$ = 2.870; $p < .05$; $d_{T1\text{-}T2}$= .132), and FFFS ($M_{T1}$ = 2.499, $M_{T2}$ = 2.367; $p < .001$; $d_{T1\text{-}T2}$= .228) all changed over time. The full results are reported in Table 5.3.

To contextualize our change data, the meta-analysis by Roberts et al. (2006) reported mean d values of .41 for social dominance facet of Extraversion, .06 for the social vitality facet of Extraversion, .05 for Agreeableness, .04 for Conscientiousness, .12 for Emotional Stability, and .37 for Openness during the 18 to 22 year bracket. Hence, our changes are of smaller magnitude compared to the meta-analytic changes reported in other studies. The temporal changes in mindfulness facets, specifically for the Non-reacting and Non-judgement facets of mindfulness, were larger than those for any of the personality traits in our study.

To test the longitudinal within-subjects relationship of RST-R, BFI, and the FFMQ, we fitted a random intercept cross-lagged panel model with an MLR estimator to adjust for multivariate non-normal data. The model showed excellent fit to the data: CFI = .977, RMSEA = .040 [.033, .047], SRMR = .041. We examined the fitted model for support of our pre-registered hypotheses. Overall, none of the predicted effects that were observed in cross-sectional studies were replicated when examining change dynamics over time. Specifically, BIS did not significantly and negatively predict Non-Judgement ($B = -.127$ [-.609, .356], $p = .607$), Acting with Awareness ($B = -.038$ [-.458, .382], $p = .858$) or Describing ($B = -.145$ [-.661, .370], $p = .580$); BAS-Reward-Reactivity did not positively predict Non-Judgement ($B = -.053$ [-.441, .335], $p = .790$); Conscientiousness was not positively associated with greater Acting with Awareness over time ($B = .155$ [-.177, .486], $p = .361$); Openness was not positively predicting Describing ($B = -.005$ [-.305, .296], $p = .976$) and Observing ($B = -.156$ [-.504, .191], $p = .378$) over time; Neuroticism was not negatively predicting Non-Reacting ($B = -.008$ [-.299, .284], $p = .960$); and finally, Extraversion was not positively related to Describing over time ($B = -.135$ [-.501, .231], $p = .469$).

Instead, higher BAS-Goal-Drive Persistence positively predicted increased Acting with Awareness over time ($B = .369$ [.044, .693], $p = .026$). Greater persistence in pursuing distant goals was associated with positive changes in this facet of Mindfulness. We additionally found significant effects in the opposite direction: mindfulness predicted within-person changes in personality traits over time. Higher Non-Reacting predicted a reduction of BIS over time ($B = -.132$ [-.247, -.017], $p = .025$), with individuals with more skills to stay calm in emotional situations decreasing their reported levels of anxiety and emotional lability. Also, higher Acting with Awareness predicted an increase in BAS-Goal-Drive Persistence ($B = .173$ [.041, .305], $p = .010$). Taken together with our finding that the personality-to-mindfulness relationship was also positive, this bi-directional pattern suggests that Acting with Awareness and BAS-Goal-Drive Persistence mutually reinforce each other over time.

Additionally, higher Non-Judging predicted an increase in Conscientiousness (B = .147 [.014, .280], $p = .031$): individuals who reported abstaining from judging their emotions and cognitions reported greater increases in Conscientiousness over time. Finally, higher levels of Describing predicted an increase in Neuroticism over time (B = .200 [.034, .366], $p = .018$). Individuals who were better able to accurately describe their emotions reported increases in emotional instability over time. A simplified figure with all significant effects is presented in Figure 5.2 and full results of all effects on mindfulness are reported in Table 5.5 in Appendix C.

**Discussion**

In the current study we report on the temporal relationship between mindfulness and personality traits during a crucial period during early adulthood. Contrary to our hypotheses, mindfulness facets predicted personality trait changes within the person over time. Two possible reasons why the predicted changes from personality to mindfulness were not supported by our data might be due a) the predicted development processes may occur over longer time-frames or b) changes in mindfulness might occur as a result of important life- events that are influenced by personality, but that are not directly related to personality as such (Bleidorn et al., 2018; Lüdtke et al., 2011).

Focusing on the positive relationships first, in our sample we found a positive feedback loop between Acting with Awareness and Goal-Drive Persistence. This indicates that the attainment of long-term rewards (the Goal-Drive Persistence component of the BAS) is enabled and supported by higher level conscious processes (Gray, 2004), including those captured by the Acting with Awareness Mindfulness component, which in turn then further increases activation of Goal-Drive Persistence Personality components. Hence, goal pursuit and conscious awareness of one's pursuit of goals mutually reinforce each other. For example, to lose weight (while we acknowledge that the success of these efforts is additionally dependent on age, SES, and a range of other home-life factors) both a long-term strategy (e.g., diet plan) as well as conscious situational awareness (mindful eating) are necessary and mutually reinforcing. The diet plan necessitates that individuals make conscious eating decision, in turn the awareness of one's behavior allows for effective long-term strategies to be put in place and the monitoring of progress towards reaching the desired goal. This finding provides a potential explanation why Acting with Awareness might be negatively related to behaviors such as smoking frequency (Adams et al., 2014) and eating disorders (Adams et al., 2012; Lavender et al., 2011) via both allowing for conscious decision making in contrast to automatic behavior and by fostering the ability to delay rewards. Our patterns imply that Acting with Awareness and Goal-Drive Persistence formed a feedback loop over time. This loop provides a potential explanation why Acting with Awareness within the mindfulness network is repeatedly found to impact behavioral change that requires both moment to moment awareness and long-term planning. Additionally, our finding that higher Non-Judging predicts greater Conscientiousness support the possibility that mindfulness facets expressing effective emotional regulation might buffer against perceived set-backs and facilitate greater goal pursuit and self-regulation (Hanley, 2016).

Focusing on potentially negative patterns: The Describing facet of mindfulness predicted an increase in Neuroticism over time. This finding might seem counter-intuitive given the cross-sectional findings in some studies that these constructs are negative related (Barnhofer et al., 2011; Iani et al., 2017) as well as general decreases of Neuroticism over time (Roberts et al., 2006). Nevertheless, theories such as the Monitor and Acceptance Theory (Lindsay & Creswell, 2017) provide a potential explanation for this relationship over time. The Describing component does not contain a non-judgmental component, but rather represents momentary awareness of one's emotions. Increased awareness of one's emotions without the necessary skills to manage them might contribute to greater emotional volatility, depression, and anxiety, with the awareness of one's emotions thereby contributing to greater anxiety, depression, and emotional volatility over time rather than reducing them. This pattern fits in with current discussions that describe mindfulness as a complex set of independent processes which can be differentially aligned in a general population (Lindsay & Creswell, 2017) and may lead to these within-person increases in Neuroticism that run counter to age-normative processes. Our finding therefore has potentially important implications for mindfulness practice. For example, when individuals describe their emotions through writing may induce negative affect for some individuals that are vulnerable or may not have the necessary emotion-regulation capacities, in the case that they recollect negative events (Pennebaker & Beall, 1986). Engaging in activities that align with the Describing facet of mindfulness on its own might foster Neuroticism and potentially lead to increased rumination and depression, which requires attention to and interactions with other mindfulness skills that help with emotion regulation.

Overall, the patterns found highlight the need for further research to examine the disconnect between the relatively consistent cross-sectional relationships between mindfulness and personality traits reported across multiple studies (Hanley, 2016; Karl & Fischer, 2019), which were not replicated over time in our sample. One potential research direction would be to adopt a network perspective similar to those adopted in personality research to understand the core variables that affect within-person dynamics over time (Cramer et al., 2010; Epskamp, 2020). It is possible that while mindfulness and individual difference domains are cross-sectionally related due to consistency effects or measurement artefacts, the dynamic changes within-individuals may be governed by a different and potentially smaller set of more specific behaviors and emotion-regulation processes. Such research may also provide new insights for attempts to develop possible patient-centered or personalized mental health interventions, which may take into account the needs, preferences and capacities of each individual.

**Limitations and Future Research Directions**

One clear limitation of our current study is the relatively short interval between measurement points (i.e., 2.5 months) and having data on only three time points. It is possible that some personality and mindfulness changes manifest and unfold via different temporal dynamics. Clearly more research using finer time sampling over longer periods is needed. One prediction based on our current patterns is that we expect stronger effects of mindfulness on personality and possibly vice versa over longer time frames. These longer-term within-person effects may approximate cross-sectional relationships. This is a hypothesis worth exploring in further research. Studies using longer timeframes would also allow testing the differential influence of general developmental trends vs significant idiosyncratic live events on the development of personality and mindfulness. One standard limitation of survey designs is that it relies on ratings, which raise common method bias issues (Podsakoff et al., 2003). We included method effects in our psychometric assessment in order to account for some of these effects (see Appendix C), but future research using other non-rating based methods or multimethod approaches to more explicitly examine method effects is needed. Furthermore, our current sample was made up from mostly female university students in a relative restricted age range, limiting generalizability of our findings. The current results cannot inform on temporal dynamics in different age cohorts or whether there are gender differences across different developmental periods. Nevertheless, our sample is similar in composition to previous studies that examined the cross-sectional relationships between mindfulness and personality (Hanley, 2016). This clearly highlights the necessity of dispositional mindfulness research to diverse its sampling base. Future research, using a more diverse sample, could also address the question what role mindfulness practice plays in the longitudinal development of trait mindfulness and personality. We would predict that active mindfulness practice would amplify individual differences and hence, strengthen mindfulness effects on personality over time.

**Open Science Statement**

This study was pre-registered prior to the end of the data collection and analysis and our time-stamped predictions made available on the OSF. The pre-registered code and data to reproduce the analyses, a table describing the proposed mechanisms and references, a full model table with all results including longitudinal relationships between all the personality trait variables and mindfulness variables on all time points is available on OSF (https://osf.io/8kufq/?view_only=b43589a0bb974a8788b415f0b7ac2e40).

We also provided an additional data set containing a subsample of participants on the OSF that we used to validate the short version of the FFMQ against the long version. The data set contains several further scales that might be of interest to researchers on mindfulness or personality, and we invite interested readers to use this data in their own analysis.

*Table 5.1 Reliability of the measures in the study.*

| | α | ω | GLB | H |
|---|---|---|---|---|
| | Time 1 | | | |
| Non-Reacting | .746[.716, .776] | .751[.722, .780] | .757 | .772 |
| Non-Judging | .757[.729, .786] | .768[.741, .795] | .803 | .822 |
| Observing | .675[.637, .714] | .676[.637, .715] | .692 | .682 |
| Describing | .821[.800, .842] | .824[.804, .845] | .866 | .85 |
| Acting with Awareness | .729[.698, .761] | .727[.695, .759] | .802 | .802 |
| BIS | .926[.918, .934] | .927[.919, .935] | .932 | .935 |
| Fight Flight Freeze Sensitivity | .769[.743, .795] | .771[.746, .797] | .817 | .781 |
| BAS-Impulsiveness | .759[.732, .786] | .763[.736, .789] | .816 | .786 |
| BAS-Reward Reactivity | .802[.780, .824] | .804[.783, .826] | .843 | .826 |
| BAS-Goal Drive Persistence | .858[.842, .875] | .863[.847, .878] | .861 | .877 |
| BAS-Reward Interest | .821[.801, .842] | .824[.804, .844] | .872 | .84 |
| Extraversion | .760[.732, .787] | .771[.745, .797] | .813 | .808 |
| Agreeableness | .721[.689, .753] | .736[.705, .766] | .789 | .747 |
| Conscientiousness | .738[.709, .768] | .750[.722, .779] | .742 | .767 |
| Neuroticism | .825[.805, .845] | .826[.806, .846] | .830 | .843 |
| Openness | .677[.641, .714] | .686[.651, .722] | .719 | .727 |
| | Time 2 | | | |
| Non-Reacting | .757[.726, .788] | .761[.731, .791] | .769 | .782 |
| Non-Judging | .785[.758, .813] | .793[.767, .820] | .807 | .833 |
| Observing | .708[.670, .746] | .709[.671, .747] | .723 | .711 |
| Describing | .821[.798, .844] | .824[.802, .846] | .869 | .846 |
| Acting with Awareness | .784[.757, .811] | .787[.760, .814] | .807 | .836 |
| BIS | .925[.917, .934] | .926[.918, .935] | .917 | .935 |
| Fight Flight Freeze Sensitivity | .800[.776, .824] | .803[.780, .827] | .860 | .821 |
| BAS-Impulsiveness | .751[.721, .780] | .754[.725, .784] | .812 | .778 |
| BAS-Reward Reactivity | .789[.764, .814] | .793[.768, .817] | .857 | .818 |
| BAS-Goal Drive Persistence | .831[.810, .851] | .837[.817, .857] | .830 | .860 |

| | | | | |
|---|---|---|---|---|
| BAS-Reward Interest | .786[.760, .812] | .791[.765, .816] | .846 | .813 |
| Extraversion | .744[.713, .776] | .754[.723, .784] | .819 | .791 |
| Agreeableness | .741[.710, .771] | .745[.713, .777] | .831 | .761 |
| Conscientiousness | .715[.679, .750] | .725[.691, .759] | .781 | .755 |
| Neuroticism | .841[.821, .860] | .842[.823, .862] | .892 | .850 |
| Openness | .710[.674, .746] | .720[.685, .754] | .777 | .786 |
| | | Time 3 | | |
| Non-Reacting | .755[.722, .787] | .761[.730, .792] | .784 | .798 |
| Non-Judging | .797[.769, .824] | .812[.787, .836] | .851 | .871 |
| Observing | .759[.726, .792] | .761[.728, .793] | .785 | .764 |
| Describing | .822[.798, .845] | .823[.800, .846] | .847 | .867 |
| Acting with Awareness | .755[.723, .788] | .761[.730, .792] | .807 | .839 |
| BIS | .923[.914, .933] | .924[.915, .934] | .958 | .932 |
| Fight Flight Freeze Sensitivity | .807[.783, .831] | .809[.785, .833] | .848 | .820 |
| BAS-Impulsiveness | .740[.708, .773] | .743[.711, .775] | .780 | .767 |
| BAS-Reward Reactivity | .808[.784, .831] | .810[.786, .834] | .874 | .838 |
| BAS-Goal Drive Persistence | .858[.840, .876] | .862[.845, .880] | .875 | .875 |
| BAS-Reward Interest | .819[.796, .842] | .822[.800, .845] | .871 | .845 |
| Extraversion | .749[.717, .782] | .757[.726, .788] | .834 | .779 |
| Agreeableness | .763[.732, .794] | .775[.746, .805] | .812 | .781 |
| Conscientiousness | .734[.699, .768] | .744[.711, .777] | .793 | .762 |
| Neuroticism | .844[.824, .865] | .846[.826, .866] | .895 | .865 |
| Openness | .732[.697, .766] | .736[.702, .770] | .736 | .755 |

*Table 5.2 Temporal Stability of the facets in the analysis.*

| | Stability from T1 to T2 | Stability from T1 to T3 | Stability from T2 to T3 | Average |
|---|---|---|---|---|
| Non-Reacting | .530 | .504 | .593 | .542 |
| Non-Judging | .459 | .507 | .544 | .503 |
| Observing | .656 | .596 | .630 | .627 |
| Describing | .654 | .613 | .639 | .635 |
| Acting with Awareness | .538 | .625 | .601 | .588 |
| FFFS | .771 | .695 | .647 | .704 |
| BIS | .710 | .663 | .628 | .667 |
| BAS-Impulsivity | .747 | .640 | .653 | .680 |
| BAS-Reward Reactivity | .647 | .562 | .597 | .602 |
| BAS-Goal Drive Persistence | .703 | .687 | .653 | .681 |
| BAS-Reward Interest | .723 | .704 | .599 | .675 |
| Agreeableness | .694 | .634 | .679 | .669 |
| Conscientiousness | .728 | .726 | .763 | .739 |
| Neuroticism | .790 | .759 | .777 | .775 |
| Openness | .798 | .738 | .735 | .757 |
| Extraversion | .836 | .821 | .821 | .826 |

*Table 5.3 Means and mean differences in the longitudinal samples at each time-point*

| | $M_{t1}$ | $M_{t2}$ | $M_{t3}$ | $p_{t1\_t2}$ | $p_{t2\_t3}$ | $d_{t1\_t2}$ | $d_{t2\_t3}$ |
|---|---|---|---|---|---|---|---|
| Acting with Awareness | 3.014 | 2.967 | 2.981 | .350 | .687 | .060 | -.024 |
| Non-Reacting | 2.958 | 2.896 | 2.982 | .182 | .046 | .086 | -.121 |
| Non-Judging | 2.878 | 2.813 | 2.978 | .213 | .001 | .086 | -.209 |
| Observing | 3.559 | 3.543 | 3.543 | .778 | .955 | .016 | .003 |
| Describing | 3.091 | 3.129 | 3.117 | .347 | .716 | -.052 | .021 |
| BIS | 2.707 | 2.648 | 2.610 | .033 | .190 | .108 | .075 |
| BAS-Impulsivity | 2.589 | 2.510 | 2.557 | .004 | .121 | .139 | -.086 |
| BAS-Reward Reactivity | 2.848 | 2.785 | 2.827 | .019 | .185 | .132 | -.079 |
| BAS-Reward Interest | 2.649 | 2.543 | 2.560 | .000 | .629 | .184 | -.029 |
| BAS-Goal Drive Persistence | 2.941 | 2.870 | 2.855 | .010 | .594 | .132 | .030 |
| FFFS | 2.499 | 2.367 | 2.407 | .000 | .273 | .228 | -.061 |
| Neuroticism | 3.047 | 3.040 | 2.984 | .904 | .138 | .005 | .066 |
| Agreeableness | 3.726 | 3.810 | 3.704 | .011 | .003 | -.134 | .158 |
| Conscientiousness | 3.169 | 3.128 | 3.132 | .267 | .968 | .055 | -.002 |
| Openness | 3.611 | 3.602 | 3.630 | .842 | .496 | .008 | -.033 |
| Extraversion | 3.102 | 3.167 | 3.099 | .029 | .028 | -.083 | .088 |

*Notes.* All p-values are based on paired sample t-tests between adjacent time-points.

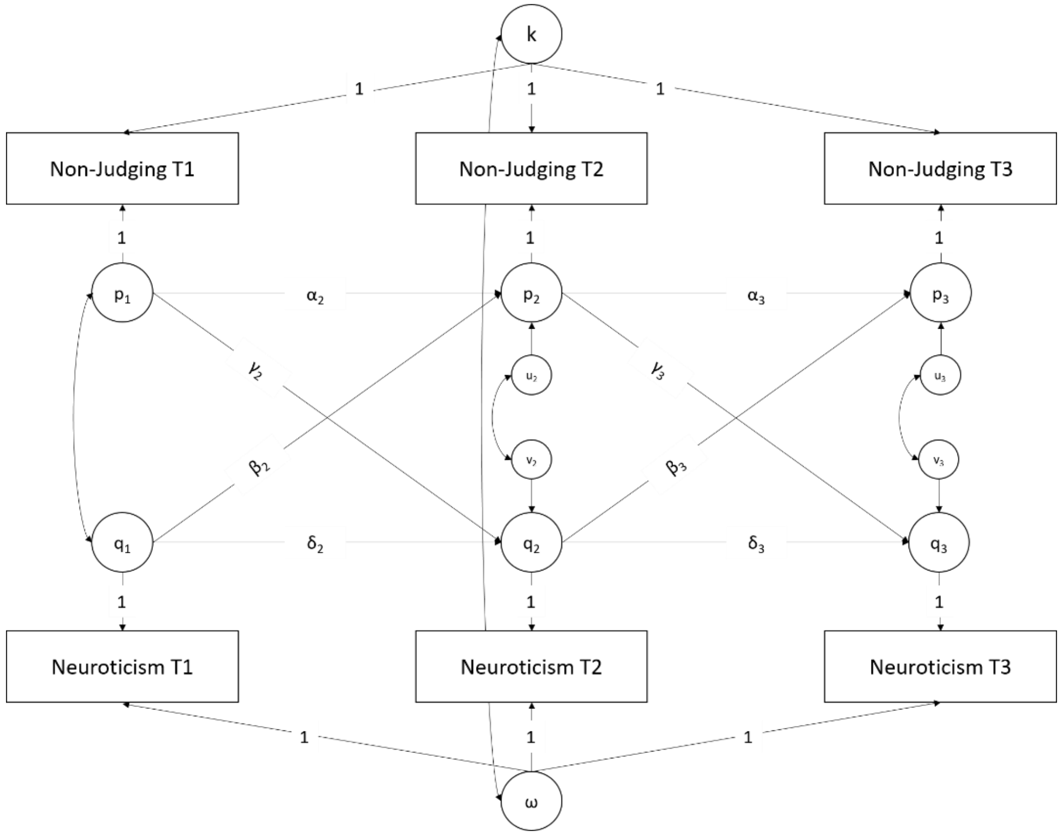



*Figure 5.1 Simplified conceptual model of the analytical model.*

In the model ω and κ account for between subject stability, allowing for the estimation of the within-subjects carry-over effects (α, δ) and the crossed effects (β, γ). Our model included the five factors of personality (Openness, Agreeableness, Conscientiousness, Extraversion, Neuroticism), the components of the Reinforcement sensitivity theory (BIS, FFFS, BAS-Impulsiveness, BAS-Goal Drive Persistence, BAS-Reward Interest, BAS-Reward Reactivity) and the five facets of mindfulness (Observing, Describing, Non-Reacting, Non-Judging, Acting with Awareness).

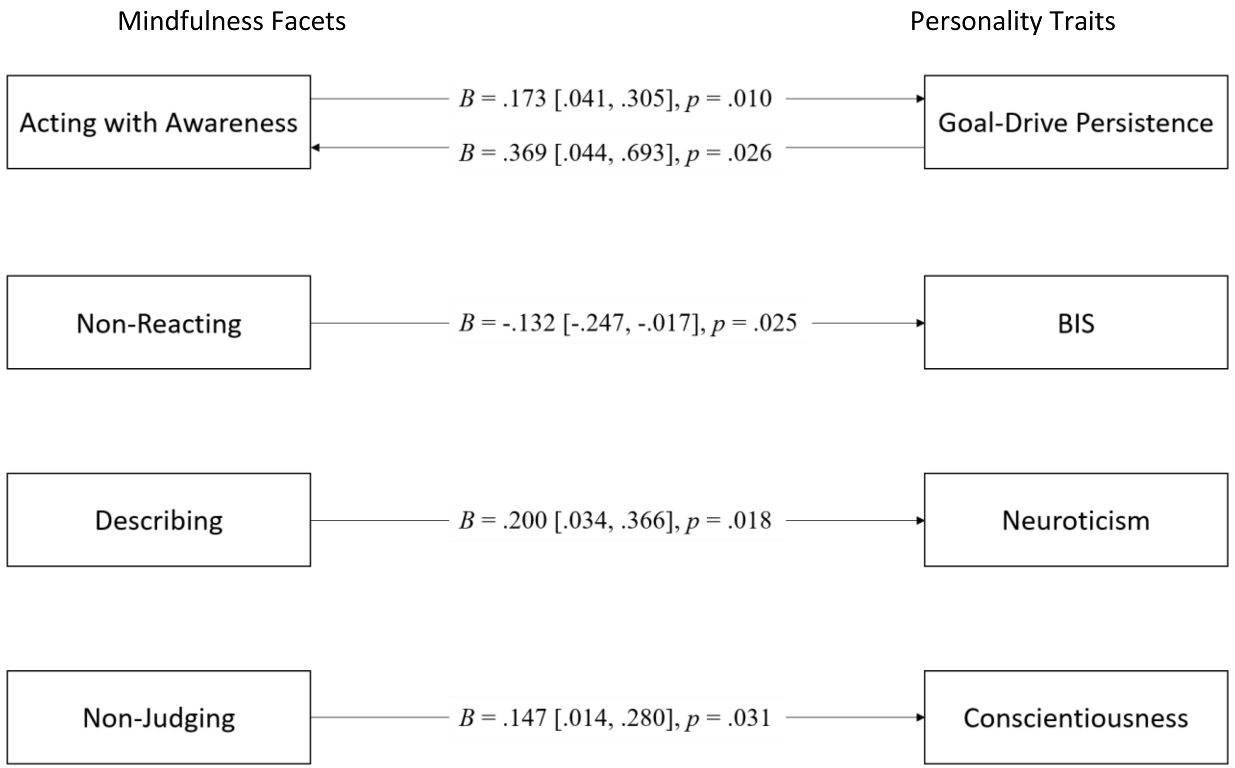



*Figure 5.2 Significant focal effects across time in the RI-CLPM*

# Chapter 3: Mindfulness across Cultures

## Study 5: A Primer to (Cross-Cultural) Multi-Group Invariance Testing Possibilities in R[8]

### Preface

Psychology has become less WEIRD in recent years, marking progress towards becoming a truly global psychology. However, this increase in cultural diversity is not matched by greater attention to cultural biases in research. A significant challenge in culture-comparative research in psychology is that any comparisons are open to possible item bias and non-invariance. Unfortunately, many psychologists are not aware of problems and their implications, and do not know how to best test for invariance in their data. In this study I provide a general introduction to invariance testing and a tutorial of three major classes of techniques :1) confirmatory and multi-group confirmatory factor analysis, with extension to exploratory structural equation modelling, and multi-group alignment; 2) iterative hybrid logistic regression as well as 3) exploratory factor analysis and principal component analysis with Procrustes rotation. This study lays the conceptual groundwork for the last study of this thesis by highlighting explicitly both conceptual and practical issues relevant for cross-cultural comparisons. Beyond the current thesis, this study has made substantial contributions to the literature and has been cited more than 50 times at the time of writing the thesis, highlighting the need in the field for a unified guide to theory and methods of cultural comparisons.

[8] This study has been published as: Fischer, R., & Karl, J. A. (2019). A primer to (cross-cultural) multi-group invariance testing possibilities in R. *Frontiers in Psychology*, *10*, 1507. https://doi.org/10.3389/fpsyg.2019.01507
Minor revisions and stylistic changes have been made to the manuscript to establish coherence with the rest of the thesis. The author of the thesis has contributed 50% of work on this manuscript. The programming of the proto-package described in this chapter is 100% the work of the author of the thesis.

The current study outlines best practices of multi-group comparisons across cultures. Within the mindfulness field, research on cross-cultural comparability of the construct have been limited (Christopher, Charoensuk, et al., 2009; Christopher, Christopher, et al., 2009). One potential reason for this paucity is the lack of awareness around issues of measurement equivalence in the wider mindfulness literature. This study aims to provide mindfulness researchers with the most current theoretical and practical foundations to support the proliferation of these techniques in the field. To increase the ease of use, we provide an initial proto-package for the R statistical programming environment which is aimed at aiding researchers in performing multi-group comparisons.

We live in an ever increasingly connected world and today it is easier than ever before to administer surveys and interviews to diverse populations around the world. This ease of data gathering with instruments often developed and validated in a single region of the world is matched by the problem that it is often difficult to interpret any emerging differences (for a discussion see: F. F. Chen, 2008; Fischer & Poortinga, 2018). For example, if a researcher is interested in measuring depression or well-being, it is important to determine whether the instrument scores can be compared across cultural groups. Is one group experiencing greater depression or psychological distress compared to another group? Hence, before we can interpret results in theoretical or substantive terms, we need to rule out methodological and measurement explanations. Fortunately, the methods have advanced significantly over the last couple of years, with both relatively simple and increasingly complex procedures being available to researchers (Boer et al., 2018; Vandenberg & Lance, 2000). Some of the more advanced methods are implemented in proprietary software, which may not be available to students and researchers, especially in lower income societies. There are excellent free and online resources available, most notable using the programming language R (R Core Team, 2020). Unfortunately, many psychological researchers are not aware of the interpretational problems in cross-cultural comparative research and fail to adequately test for measurement invariance (see: Boer et al., 2018). Our tutorial aims to demonstrate how three different and powerful classes of analytical techniques can be implemented in a free and easy to use statistical environment available to student and staff alike which requires little computer literacy skills. We provide the code and example data online (*see https://osf.io/agr5e/).* We strongly encourage readers to download the data and follow the code to gain some experience with these analyses.

We aim to provide a basic introduction that allows novices to understand and run these techniques. The three most common approaches are exploratory and confirmatory methods within the classic test theory paradigm as well as item response theory approaches. We also include recent extension such as exploratory structural equation modelling and multi-group alignment. Although these approaches often differ at the philosophical and theoretical level, at the computational level and in their practical implementation, they are typically converging (Fontaine, 2005). We provide a basic introduction and discuss them together here. We encourage readers interested in more technical discussions and their conceptual and computational distinctions to consult more technical overviews and extensions (e.g. Boer et al., 2018; Borsboom, 2006; Fontaine, 2005; Hambleton & Jones, 1993; J. Long, 1983; J. S. Long, 1983; Meredith, 1993; Tabachnick & Fidell, 2007).

Throughout the tutorial, we use a two-group comparison. Unfortunately, results from two sample comparisons are open to a host of alternative interpretations, even if method issues can be ruled out. Therefore, we strongly encourage researchers to include more than two samples in their research design. Multiple-sample comparisons can pose some additional analytical choices for researchers (especially for the EFA component) and we discuss easily available options for expanding the analyses to more than two samples. In the final section, we directly compare the different methods and their relative advantages and disadvantages.

**The basic principle of measurement invariance testing**

With invariance testing, researchers are trying to assess whether an instrument has the same measurement properties in two or more populations. We need to distinguish a number of different properties of measurement instruments. In order to provide a common terminology, we use the item response theory approach (we will be ignoring the person parameters) and note equivalent parameters in classic test theory terms, where necessary. Because in psychology we often do not have access to objective indicators, our best estimate about the psychological expression of interest when evaluating a test is the overall score on a test. This overall score is taken as an estimate of the underlying ability parameter of the person or the level of latent variable (the psychological trait we would like to measure). Invariance testing of instruments focuses on the relationship between each individual item and the overall score of the instrument. It is important to highlight that cross-cultural researchers use different types of data for invariance testing and that the interpretation of the overall score differs depending on the type of test being examined. For example, an intelligence test will capture the extent to which individuals answer questions correctly, which then leads to clear interpretations of the parameters in terms of item difficulty and item discrimination. For researchers using rating scales, these same parameters are often interpreted in terms of factor loadings (how well an item relates to a presumed underlying factor) and intercepts (is there some guessing or response bias involved, that is not related to the latent variable). The interpretation therefore differs somewhat, but the statistical properties are similar. For example, if an individual has a higher score on the underlying ability as either a true ability or a preference or trait, then they should report a higher mean (the person is more likely to answer an item 'correctly'). When dissecting the relationship between an item and the overall score, there are three main parameters: 1) the item difficulty or item location, 2) the discrimination or slope parameter and 3) a parameter for pseudo-guessing, chance or the intercept (see Figure 6.1). The item difficulty describes how easy or difficult an item is, in other words, the amount of a latent trait that is needed for an individual to endorse an item with a 50% probability (for rating scales) or answer it correctly (for ability tests). Item discrimination or the slope describes how well an item discriminates between individuals (both for ability tests and rating scales). In factor analytic terms it can also be thought of as the item loading – how strongly the item is related to the latent variable. The guessing parameter refers to the point where individuals with a low level of ability (for ability tests) or expression of a psychological trait (for rating scales) may still able to guess the correct answer (on a test) or responds with a higher score than would be indicated by their latent trait score. In factor analytic terms, this is conceptually equivalent to the intercept. More parameters can be estimated and tested (in particular within a multivariate latent variable framework), but these three parameters have been identified as most important for establishing cross-cultural measurement invariance (e.g. Fontaine, 2005; Meredith, 1993; Vandenberg & Lance, 2000). Of these three parameters, item discrimination and guessing parameters are the most central and have been widely discussed in terms of how they produce differential item functioning across groups.

**Levels of measurement invariance and differential item bias**

In cross-cultural comparisons, it is important to identify whether these parameters are equivalent across populations, to rule out the possibility that individuals with the same underlying ability have a different probability to give a certain response to a specific item depending on the group that they belong to (see Figure 6.2).

There are at least three different levels of invariance or equivalence that are often differentiated in the literature (see: Fontaine, 2005; Meredith, 1993; Milfont & Fischer, 2010; van de Vijver & Leung, 1997; Vandenberg & Lance, 2000). The first issue is whether the same items can be used to measure the theoretical variable in each group. For example, is the item “I feel blue” a good indicator of depression[9]? If the answer is yes, we are dealing with configural invariance. The loadings (the extent to which each item taps into the underlying construct of depression) are all in the same direction in the different groups (this is why this sometimes called form invariance), nevertheless, the specific factor loadings or item discrimination parameters may still differ across samples.

[9] Colour connotations are often language specific. For example, feeling blue might indicate intoxication in German, but not depression per se.

If the item discrimination or factor loadings are identical across the samples, then we are dealing with metric invariance. The item discriminates similarly well between individuals with the same underlying trait. Equally, the item is related to the same extent to the underlying latent variable in all samples. This implies that an increase in a survey response to an item (e.g., answering with a 3 on 1-7 Likert scale instead of a 2) is associated with the same increase in depression (the latent variable that is thought to cause the responses to the survey item) in all groups sampled. If this condition is met for all items and all groups, we can compare correlations and patterns of means (e.g., profiles) across cultural samples, but we cannot make claims about any latent underlying construct differences (see Fontaine, 2005). See Panels b and c in Figure 6.2 for an example where an increase in the underlying ability of trait is associated with equal changes in responses to an individual item, but there are still other parameters that differ between samples.

If we want to compare instrument scores across groups and make inferences about the underlying trait or ability levels, we need to also at least constrain guessing or intercept parameters (and also item difficulty in IRT). Metric invariance only means that the slopes between items and latent variables are identical, but the items may still be easier or difficult overall or individuals might be able to guess answers. Therefore, we have to constrain intercepts to be equal. If this condition is met, we have scalar or full score invariance. The advantage of full score equivalence is that we can directly compare means and interpret any differences in terms of the assumed underlying psychological construct.

These levels of invariance are challenged by two major item biases. Uniform item bias describes a situation where the item equally well discriminates between individuals with the same underlying true ability. In this case the curves are parallel and the items do not differ in discrimination (slopes). People of one group have an unfair advantage over the other group, but the relative order of individuals within each group is preserved (see panels b and c in Figure 6.2). Non-uniform item bias occurs when the order of individuals along the true underlying trait is not reflected in the item responses (see panels a and d in Figure 6.2). The item responses differ across groups and true levels of the underlying ability. The most important parameter here is item discrimination, but other parameters may also change. Together, these item biases are often examined in the context of differential item functioning (DIF).

The methods discussed below differ in the extent to which they allow researchers to identify item bias and invariance in these parameters. Exploratory factor analysis with Procrustes rotation is the least rigorous method, because it only allows an overall investigation of the similarity of factor loadings, but it does not typically allow analysis at the item level. Multi-group confirmatory factor analysis and differential item functioning analysis with logistic regression allow an estimation of both the similarity in factor loadings and intercepts/guessing parameters. We briefly describe the theoretical frameworks, crucial analysis steps and how to interpret the outputs in a two-group comparison. We then compare the relative advantages and disadvantages of each method and their sensitivity to pick up biases and violations of cross-cultural invariance.

**What to do if invariance is rejected?**

All the techniques that we describe below rely on some form of fit statistic – how much does the observed data deviate from the assumption that the statistical parameters are equal across groups? The different techniques use different parameters and ways to test this misfit, but essentially it always comes down to an estimation of the deviation from an assumed equality of parameters. Individual items or parameters are flagged for misfit. The most common immediate strategy is to conduct exploratory analyses to identify a) the origin of the misfit or DIF and to then b) examine whether excluding specific items, specific factors or specific samples may result in improved invariance indicators. For example, it might be that one item shows some translation problems in one sample and it is possible to exclude this item or to run so-called partial invariance models (see below). Or there might be problems with a specific factor (e.g., translation problems, conceptual issues with the psychological meaning of factor content – often called cultural construct bias). It might be possible to remove the factor from the analyses and proceed with the remaining items and factors. Or it may also happen that one sample is found to be quite different (e.g., different demographics or other features that distinguish the sample from the other samples including differences in reading ability, education, economic opportunities). In this case, it is possible to exclude the individual sample and proceed with the remaining cultural samples. The important point here is that the researcher needs to carefully analyze the problem and decide whether it is a problem with an individual item, scale or sample, or whether it points to some significant cultural biases at the conceptual level.

We would like to emphasize that it is perfectly justified to conduct an invariance analysis and to conclude that it is not meaningful to compare results across groups. In fact, we wish more researchers would take this stance and call out test results that should not be compared across ethnic or cultural groups. For example, if the factor structures of an instrument are not comparable across two or more groups, a comparison of means and correlations are invalid. There is no clear interpretation of any mean differences if there is no common structure. Hence, invariance analysis can be a powerful tool for applied psychologists to counter discrimination and bias as well as cultural psychologists interested in demonstrating cultural relativism. Unfortunately, too often the insights from invariance analyses are ignored and researchers proceed with cross-cultural comparisons, which are then inherently meaningless (see: Boer et al., 2018).

## Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) is probably the most widespread measurement model approach in psychology. Most constructs in psychological research cannot be directly observed but need to be inferred from several observed indicators (Gorsuch, 1983; Horn, 1965). These indicators can be recorded behaviors or responses to Likert type scales, returning to our example of depression, we may infer levels of an underlying depression variable through observations of sleeping problems, changes in mood, or weight gain. The general advantage and appeal of CFA is that explicitly tests the theoretical structure that a researcher has about the instrument. CFA using a theory-driven approach for modeling the covariance between items, meaning it is a measurement model that treats items as indicators of a theoretically assumed underlying latent constructs (Bollen, 1989; J. Long, 1983). The researcher needs to decide a priori which items are expected to load (are indicators of the latent variable) on which latent variable. Typically, researchers are interested in the simple structure, in which each item is expected to load on only one latent factor (see Figure 6.3). Figure 6.3 shows the main parts of a CFA model. Observed indicators (e.g., item responses to a survey) are represented by squares, whereas estimated parameters are symbolized by ovals or circles. Each item in our example is allowed to load on one latent variable. The resulting factor loadings represent the relationship of the observed indicator to each of the extracted latent factors. The strength of the loadings can range from 0 (no relationship) to either -1 or 1 (identical; if the latent variables are standardized, in unstandardized situations the loadings are dependent on the measurement scale). In our example, the first four items only load on factor 1, whereas the last three items only load on factor 2. In multi-group analyses, we also estimate the item intercept (which is conceptually similar to the pseudo guessing parameter discussed above

For technical (identification) purposes, one of the factor loadings is typically set to 1 to provide identification and a meaningful scale. It also important to have at least three items per latent factor (although this rule can be relaxed, see Bollen, 1989). CFA is demanding in terms of data quality, assuming at least interval data that is multivariate normally distributed, an assumption that is unfortunately often violated. Some procedures have been developed to correct for a violation of multivariate normality (see for example, Satorra & Bentler, 1988), which are implemented and can be requested in the R package that we describe below.

CFA is confirmatory: the theoretically proposed structure of implied covariances among items is statistically tested and compared to the observed covariances based on the sample specific item responses. One of the most important questions is how to evaluate whether the model fits the data. Various different fit indices are available. The deviation of the theoretically predicted to the empirically observed covariances is captured in using the chi-square statistic. This is the oldest and probably most important diagnostic tool for deciding whether the theoretical prediction was plausible or not. The smaller the chi-square value, the less the theoretical model deviates from the observed sample covariance matrix. The exact fit of the theory to the data can be evaluated with a significance test, therefore this is often called an exact fit test (see Barrett, 2007). Ideally, we want a non-significant chi-square value. Unfortunately, there are both conceptual and statistical drawbacks for the chi-square. First, any theoretical model is only an approximation of reality, therefore any chi-square is a priori known to be incorrect and bound to fail because reality is more complex than implied in simple models (Browne & Cudeck, 1992). Statistically, the test is sample size dependent. Any model will be rejected with a sufficiently large sample size (Bentler & Bonett, 1980; Bollen, 1989; or an example of cross-cultural study demonstrating this dependence, see Fischer et al., 2011).

To overcome these problems, a number of alternative fit measures have been proposed (even though most of them still are derived from the $\chi^2$ statistic). Here we focus on the most commonly reported fit statistics (Hu & Bentler, 1999), which can be differentiated into (a) incremental or comparative and (b) lack-of-fit indices. Incremental or comparative fit models compare the fit of the theoretical model against an alternative model. This is (typically) an independence model in which no relationships between variables are expected. Higher values are indicating better fit with values above .95 indicating good fit (Hu & Bentler, 1998). The Tucker–Lewis Index (TLI) or nonnormed fit index (NNFI) and the comparative fit index (CFI; Bentler, 1990) are the most commonly reported and more robust indicators (Hu & Bentler, 1999). Lack of fit indices in contrast indicate better fit, if the value is lower. The standardized root mean square residual (SRMR; Bollen, 1989) compares the discrepancy between the observed correlation matrix and the implied theoretical matrix. Smaller values indicate that there is less deviation. Hu and Bentler (1999) suggested that values less than .08 are acceptable. The root mean square error of approximation (RMSEA; Browne & Cudeck, 1992) uses a similar logic, but also takes into account model complexity and rewards more parsimonious models. Historically, values ranging between 0.06 and 0.08 were deemed acceptable, but simulations by Hu and Bentler (1998, 1999) suggested that a cut-off of .06 might be more appropriate.

However, it is important to note that the selection of fit indices and their cutoff criteria are contentious. Marsh, Hau, and Wen (2004) warned researchers against blindly adopting cutoff values suggested by specific simulations such as the famous Hu and Bentler study (1998; 1999). One specific issue is that models with higher factor loadings (indicating more reliable models) might be penalized by these fit indicators (Y. Kang et al., 2016; McNeish et al., 2018), which creates a paradoxical situation in that theoretically better and more reliable models are showing worse fit. They suggested to also examine other fit indices such as McDonald's Non-Centrality Index (NCI: McDonald, 1989). We urge researchers to take a cautious approach and to evaluate model fit as well as examining the overall factor loadings and residuals when determining model fit. If your model is fitting well, but has poor factor loadings and shows large residuals, it is probably not the best model. A good strategy is to compare a number of theoretically plausible models and then select the model that makes most theoretical sense and has the best fit (MacCallum & Austin, 2000; Marsh et al., 2004).

Often, researchers would first test the model separately in each cultural group. This can provide valuable insights into the structure in each group. However, the individual analyses in each sample do not provide information about whether the structure is identical or not across groups. For this, we need to conduct a multi-group analysis. This is the real strength of CFA, because we can constrain relevant parameters across groups and test whether the fit becomes increasingly worse. If there is overall misfit, it then becomes possible to test whether individual items or groups cause misfit. Therefore, multi-group CFA provides information at both the scale and item level, making it a powerful tool for cross-cultural researchers.

To proceed with the examination of invariance, a number of parameters can be constrained across groups or samples in a hierarchical fashion which allow a test of the invariance levels that we described at the beginning of this article. The first step is form invariance (Cheung & Rensvold, 2000; Meredith, 1993) or configural invariance (Byrne et al., 1989). All items are expected to load on the same latent factor. The second level is factorial invariance (Cheung & Rensvold, 2000) or metric invariance (Byrne et al., 1989), in which the factor loadings are forced to be equal across groups. This tests whether there is non-uniform item bias (see above). The third level that is necessary to test is scalar invariance (Vandenberg & Lance, 2000) or intercept invariance (Cheung & Rensvold, 2000), which constrains the item intercepts to be equal across groups. It tests whether there is uniform item bias present in an item. It is desirable to obtain scalar invariance because then means can be directly compared across groups. Unfortunately, few cross-cultural studies do test this level of invariance (Boer et al., 2018).

At each step, researchers have to decide whether their more constrained model still fits the data or not. In addition to the fit indicates that we have discussed above, it is common to examine change in fit statistics. The traditional change statistic is the chi-square difference test, in which the chi-square of the more restricted model is compared to the chi-square of the more lenient model. A significant chi-square difference indicates that model fit is significantly worse in the more restricted model (Anderson & Gerbing, 1988). However, as before, the chi-square is sample size dependent and therefore, other fit indices have been introduced. Little (1997) was the first to suggest that differences in the NNFI/TLI and CFI are informative. Similarly, it is possible to examine changes in RMSEA (Little et al., 2007). For these change in fit indices, current standards are to accept models that show differences equal to or less than 0.01. Some authors also suggested examining other fit indices, including ΔMcDonald's NCI (Y. Kang et al., 2016). All these fit indices are judged in relation to deterioration in fit between more and less restricted models, with cut-offs based on either experience or simulations. Unfortunately, there is no universal agreement on acceptable standards (see: F. F. Chen, 2007; Milfont & Fischer, 2010; Putnick & Bornstein, 2016). For example, Rutkowski & Svetina (2014) ran simulation models focusing specifically on conditions where researchers have more than 10 samples in their measurement invariance analysis and suggested that in these multi-group conditions criteria for metric invariance tests could be relaxed, but that the criteria for judging scalar invariance should remain at traditional cut-offs of less than 0.01.

What do you need to do if factorial invariance is rejected at any of these steps? First, it is advisable to investigate the models in each group separately and to also check modification indices and residuals from the constrained model. Modification indices provide information of how much the $\chi^2$ would change if the parameter was freed up. There are no statistical guidelines of how big a change has to be in order to be considered meaningful. Theoretical considerations of these modification indices are again important: There might be both meaningful theoretical (conceptual differences in item meaning) or methodological reasons (item bias such as translation issues, culture specificity of item content, etc.) why either factor loadings or intercepts are different across groups. The appropriate course of action depends on the assumed reasons for misfit. For example, a researcher may decide to remove biased items (if there are only few items and if this does not threaten the validity of the overall scale). Alternatively, it is possible to use partial invariance, in which the constrains on specific items are relaxed (Byrne et al., 1989 see below)

**How to run a multi-group CFA in R**

We describe the steps using the lavaan (Rosseel, 2012) and semTools (Contributors, 2016) package, which need to be loaded (see the online supplementary materials). For illustration purposes, we use data from Fischer et al. (2019) in which they asked employees in a number of countries, including Brazil and New Zealand (total N = 2,090, we only included a subset of the larger data set here), to report whether they typically help other employees (Helping behaviour, 7 items) and whether they make suggestions to improve work conditions and products (Voice behaviour, 5 items). Individuals responded to these items on a 1-7 Likert type scale.

**Running the CFA.** The first CFA relevant step after reading in the data and specifying missing data (see online supplementary materials) is to specify the theoretical model. We need to create an object that contains the relevant information, e.g., what item loads on what factor and whether factors and/or item errors are correlated. The way this is done is through regression-like equations. Factor loadings in R are indicated by =~ and covariances (between factors or error terms for items) are indicated by ~~. The model is specified similar to writing regression equations.

In our case, the model is:

```
cfa_model<- ‘
help =~ help1+help2+help3+help4+help5+help6+help7
voice=~ voice1+voice2+voice3+voice4+voice5 ‘
```

We have seven items that measure helping behavior and five items that measure voice behaviors. Now, we need to run the model and test whether the theoretical model fits to our data. The basic command is:

```
fit_cfa <- cfa(cfa_model, data = example)
```

**Running a Multi-Group CFA.** This creates an object that has the statistical results. The current command does not specify separate CFAs in the individual groups, but tests the model in the total sample. To separate the models by group, we need to specify the group (important note: in lavaan, we will not get the separate fit indices per group, but only an overall fit index for all groups combined; if you want to run separate CFAs in each group, it is useful to subset the data first, see the online supplementary materials for data handling):

```
fit_cfa_country <- cfa(cfa_model, data = example, group = "country")
```

To get the statistical output and relevant fit indices, we can now call the object that we just created in the summary() function:

```
summary(fit_cfa_country, fit.measures = TRUE, standardized = TRUE, rsquare = TRUE)
```

The fit.measures argument requests the commonly described fit indices that we described above. The standardized command provides a standardized solution for the loadings and variances that is more easily interpreted. In our case, the fit is mixed overall: $\chi^2$ (106) = 928.06, $p < .001$, CFI = .94, TLI = .93, RMSEA = .086, SRMR = .041. For illustration purposes, we continue with this model, but caution that it is probably not demonstrating sufficient fit to be interpretable.

**Invariance Testing – Omnibus Test.** To run the invariance analysis, we have two major options. One is to use a single command which runs the nested analyses in a single run:

```
measurementInvariance(model = cfa_model, data = example, group = "country")
```

We specify the theoretical model to test, our data file and the grouping variable (country). In the output, Model 1 is the most lenient model, no constraints are imposed on the model and separate CFA's are estimated in each group. The fit indices mirror those reported above. Constraining the loadings to be equal, the difference in $X^2$ between Model 1 and 2 is not significant: $\Delta\chi^2(df = 10) = 16.20$, $p = .09$, and the change in both CFI (.00) and RMSEA (.003) are negligible. Since $X^2$ is sensitive to sample size, the CFI and RMSEA parameters might be preferable in this case (see Cheung & Lau, 2012; Milfont & Fischer, 2010; Putnick & Bornstein, 2016). When further constraining the intercepts to be equal, we have a significant $X^2$ difference again: $\Delta\chi^2(df = 10) = 137.03$, $p < .001$. The difference in CFI (.009) and RMSEA (.003) are also below commonly accepted thresholds, therefore, we could accept our more restricted model. However, as we discussed above, the overall fit of the baseline model was not very good and some of the fit indices have conceptual problems. In the ccpsyc package, we included a number of additional fit indices that have been argued to be more robust (see for example: Y. Kang et al., 2016). Briefly, to load the ccpsyc package (the devtools package is required for installation), call this command:

```
devtools::install_github("Jo-Karl/ccpsyc")

library(ccpsyc)
```

The function via the ccpsyc package is called equival and we need to specify the CFA model that we want to use, then the relevant data file (dat = example) and the relevant grouping variable (group = "country"). For this function, the group variable needs to be a factor (e.g., the country variable is not numerical variable). It is important to note that the equivalent function fits all models using a robust MLM estimator rather than an ML estimator.

An example of the function is:

```
equival(cfa_model, dat = example, group = "country", orthog = F)
```

In our previous example, the fit indices were not acceptable even for less restricted models. Therefore, the more restricted invariance tests should not be trusted. This is a common problem with CFA. If there is misfit, we can either trim the parameter (drop parameters, variables or groups from the model that are creating problems) or we can add parameters. If we decide to remove items from the model, the overall model needs to be rewritten, with the specific items removed from the revised model (see the steps above). One question that you as a researcher needs to consider is whether removing items may change the meaning of the overall scale (e.g., underrepresentation of the construct; see Fontaine, 2005). It might also be informative for a cross-cultural researcher to consider why a particular item may not work as well in a given cultural context (e.g., through qualitative interviews with respondents or cultural experts to identify possible sources for misfit).

To see which parameters would be useful to add, we can request modification indices.

This can be done using this command in R:

```
mi <- modificationIndices(fit_cfa)
```

We could now simply call the output mi to show the modification indices. It gives you the expected drop in $X^2$ as well as what the parameter estimates would be like if they were freed up. Often, there are many possible modifications that can be done and it is cumbersome sifting through a large output file. It can be useful to print only those modification indices above a certain threshold. For example, if we want to only see changes in $\chi^2$ above 10, we could add the following argument:

```
mi <-modificationIndices(fit_cfa, minimum.value = 10, sort = T)
```

We also added a command to have the results sorted by size of change in $\chi^2$ for easier examination. If we now call the object as usual (just write *mi* into your command window), this will give us modification indices for the overall model, that is modification indices for every parameter that was not estimated in the overall model. For example, if there is an item that may show some cross-loadings, we now see how high that possible cross-loading might be and what improvement in fit we would achieve if we were to add that parameter to our model. The function also gives us a bit more information, including the expected parameter change values (column epc) and information about standardized values (sepc.lv: only standardizing the latent variables; sepc.all: standardizing all variables; sepc.nox: standardizing all but exogenous observed variables).

**Invariance Testing – Individual Restrictions & Partial Invariance.** This leads us to the alternative option that we can use for testing invariance. Here, we manually construct increasingly restricted models. This option will also give us opportunities for partial invariance. We first constrain loading parameters in the overall cfa command that we described above:

```
metric_test <-cfa(cfa_model, data = example, group = "country", group.equal =
c("loadings"))
```

As can be seen here, we added an extra command *group.equal* which now allows us to specify that the loadings are constrained to be equal. If we wanted to constraint the intercepts at the same time, we need to use: group.equal = c(“loadings”, “intercepts”). We can get the usual output using the summary function as described above.

We could now request modification indices for this constrained model to identify which loadings may vary across groups:

```
mi_metric <- modificationIndices(metric_test, minimum.value = 10, sort. = T)
```

As before, it is possible to restrict the modification indices that are printed. We could also investigate how much better our model would be if we freed up some parameters to vary across groups. In other words, this would tell us if there are some parameters that vary substantively across groups and if it is theoretically plausible, we could free them up to be group specific. This then would become a partial invariance model (see Meredith, 1993). We provide the *lavTestScore.clean* function in the ccpsyc package which has as single input a metrically constrained CFA model. The relevant command is:

```
lavTestScore.clean(metric.test)
```

If we wanted to relax some of the parameters (that is running a partial invariance model), we can use the *group.partial* command. Based on the results from the example above, we allowed the third help item to load freely on the help latent factor in each sample:

```
fit_partial <- cfa(cfa_model, data = example, group = "country", group.equal = c("loadings"),
group.partial = c("help =~ help3 "))
```

**Estimating effect sizes in item bias in CFA: dmacs**

The classic approach to multi-group CFA does not allow an estimation of the effect size of item bias. As we did above, when running a CFA to determine equivalence between groups, researchers rely on differences in fit measures such as ΔCFI and $\Delta\chi^2$. These cut-off criteria inform researchers whether a structure is equivalent across groups or not, but they do not provide an estimate of the magnitude of misfit. To address this shortcoming Nye and Drasgow (2011) proposed an effect size measure for differences in mean and covariance structures ($d_{MACS}$). This measure is estimating the degree of non-equivalence between two groups on an item level. It can be interpreted similar to established effect sizes (Cohen, 1988) with values of greater than 0.20 being considered small, 0.50 are medium, and 0.80 or greater are large. It is important that these values are based on conventions and do not have any direct practical meaning or implication. In some contexts (e.g., high stakes employment testing), even much smaller values might be important and meaningful in order to avoid discrimination against individuals for specific groups. In other contexts, these criteria might be sufficient.

**How to do the analysis in R**

To ease the implementation of $d_{\mathrm{MACS}}$, we created a function in R as part of our ccpsyc package that allows easy computation (see the online supplementary materials for how to install this package and function). The function dMACS in the ccpsyc package has three arguments: *fit.cfa* which takes a lavaan object with two groups and a single factor as input, as well as a *group1* and *group2* argument in which the name of each group has to be specified as string. The function returns effect size estimates of item bias (dMACS) for each item of the factor. In our case, we could specify first a CFA model with only the helping factor, then run the lavaan multi-group analysis.

```
help_model <- 'help =~ help1 + help2 + help3 + help4 + help5 + help6 + help7'

help_cfa <- cfa(help_model, data = example, group = "country")
```

We now can call:

```
dMACS(help_cfa, group1 = "NZ", group2 = "BRA")
```

to get the relevant bias effect size estimates. One of the items (item 3) shows a reasonably large dMACS value (.399). As you will remember, this item also showed problematic loading patterns in the CFA reported above, suggesting that this item might be problematic. Hence, even when the groups may show overall invariance, we may still find item biases in individual items.

**Limitations of dMACS**

A limitation of the current implementation of $d_{\mathrm{MACS}}$ is that the comparison is limited to a unifactorial construct between two groups. After running the overall model, researchers need to respecify their models and test each dimension individually.

**Strengths and weaknesses of CFA**

CFA is a theory-driven measurement approach. It is ideal for testing instruments that have a well-established structure and we can identify which items are expected to load on what latent variables. This technique provides an elegant and simple test for all important measurement questions about item properties with multi-dimensional instruments. At the same time, CFA is not without drawbacks. First, it requires interval and multivariate normally distributed data. This can be an issue with the ordinal data produced by Likert-type scales if the data are heavily skewed. Nevertheless, a number of studies have shown that potential issues can be overcome by the choice of estimator (for example Flora & Curran, 2004; Holgado–Tello et al., 2008; Li, 2016). Second, establishing adequate model fit and what counts as adequate are tricky questions and this is continuously debated in the measurement literature. Third, CFA ideally requires moderately large sample (N > 200; e.g., Barrett, 2007). Fourth, nonnormality and missing data within and across cultural groups can create problems for model estimation and identifying the problems can become quite technical. However, the technique is becoming increasingly popular and has many appealing features for cultural psychologists.

**Exploratory Structural Equation Modelling**

CFA is a powerful tool, but it has limitations. One of the biggest challenges is that a simple structure in which items only load on one factor is often empirically problematic. Exploratory factor analysis (see below) presupposes no structure, therefore any number of cross-loadings are being permitted and estimated, making it a more exploratory technique. To provide a theory-driven test while allowing for the possibility of cross-loadings, Exploratory Structural Equation Modeling (ESEM, Asparouhov & Muthén, 2009) has been proposed. ESEM combines an exploratory factor analysis approach that allows an unrestricted estimation of all factor loadings which can then be further compared with a standard structural equation approach. Technically, an EFA is conducted with specific factor rotations and loading constraints. The resulting loading matrix is then transformed into structural equations which can be further tested and invariance indices across groups can be estimated. ESEM also allows a better estimation of the correlated factor-structures than EFA as well as provides more unbiased estimates of factor covariances than CFA (because of the restrictive assumption of a simple structure with no cross-loadings for CFA). The ESEM approach has been proposed within Mplus (Muthén & Muthén, 2018), but it is possible to run compatible models within R (see Guàrdia-Olmos et al., 2013).

We use the approach described by Kaiser (2018). The first step is to run an exploratory factor analysis.

```
beh_efa <- fa(example[-1], nfact = 2, rotate = "geominQ", fm = "ml")
```

As before, we are creating an output object (beh_efa) that contains the results of the factor analysis (fa). We specify the data set 'example' and the square brackets indicates that we want to run the analysis only for the survey data excluding the country column (example[-1]). We specify 2 factors (nfact = 2) and ask for a specific type of factor rotation that is used by Mplus (rotate = "geominQ"). Finally, we specify a Maximum Likelihood estimator (fm = "ml").

We now will prepare the output of this EFA to create structural equations that can be further analyzed within a CFA context.

```
beh_loadmat <- zapsmall(matrix(round(beh_efa$loadings, 2), nrow = 12, ncol = 2))

rownames(beh_loadmat) <- colnames(example[-1])
```

We use the function *zapsmall* to get the rounded factor loadings from the two factors in the previous EFA (this is the (round(beh_efa$loadings,2) component). The $ sign specifies that we only use the factor loadings from the factor analysis output. We have 12 variables in our analysis, therefore we specify nrow = 12. We have two factors, therefore we specify two columns (ncol = 2). To grab the right variable names, we include a command that assigns the row names in our loading matrix from the respective column (variable) names in our raw data set. Since we have the country variable still in our data set, we need to specify that this column should be omitted: example[-1]. All the remaining column names are taken as the row names for the factor analysis output.

To create the structural equations, we need to create the following function:

```
new_model <- vector()

for (i in 1:2) {

  new_model[i] <- paste0("F",i,"=~ ", paste0(c(beh_loadmat[,i]), " * ",
names(beh_loadmat[,1]), collapse = " + "))

}
```

The term i specifies the number of factors to be used. In our case, we have 2 factors. We then need to specify the relevant loading matrix that we created above (beh_loadmat). If we now call:

```
new_model
```

we should see the relevant equations that have been computed based on the EFA and which can be read as a model to be estimated within a CFA approach. Different from our CFA model above, all items are now listed and the loading of each item on the two different factors is now specified in the model.

```
[1] "F1=~ 0.55 * help1 + 0.58 * help2 + 0.69 * help3 + 0.91 * help4 + 0.78 * help5 + 0.77 *
help6 + 0.51 * help7 + 0.15 * voice1 + -0.02 * voice2 + -0.01 * voice3 + 0.13 * voice4 + -0.01
* voice5"

[2] "F2=~ 0.12 * help1 + 0.1 * help2 + 0.01 * help3 + -0.1 * help4 + 0.04 * help5 + 0.02 *
help6 + 0.24 * help7 + 0.65 * voice1 + 0.85 * voice2 + 0.77 * voice3 + 0.65 * voice4 + 0.8 *
voice5"
```

We now run a classic CFA, similar to what we did before. We specify that the estimator is Maximum Likelihood (estimator = “ML”). For simplicity, we only want to call some of the fit measures using the fit measures function within lavaan.

```
beh_cfa_esem <- cfa(new_model, data = example, estimator = "ML")
fitmeasures(beh_cfa_esem, c("cfi", "tli", "rmsea", "srmr"))
```

This analysis was done on the full data set including the Brazilian and NZ data simultaneously, but we are obviously interested in whether the data are equivalent across groups or not (when using this specific model). We can set up a configural invariance test model by specifying the grouping variable and calling the relevant fit indices:

```
fitmeasures(cfa(
  model = new_model,
  data = example,
  group = "country",
  estimator = "ML"),
  c("cfi","tli","rmsea","srmr"))
```

If we want to now constrain the factor loadings or intercepts to be equal across groups, we can add the same restrictions as described above. For example, for testing scalar invariance in which constrain both the loadings and intercepts to be equal, we can call this function:

```
fitmeasures(cfa(
  model = new_model,
  data = example,
  group = "country",
  estimator = "ML",
  group.equal = c("loadings", "intercepts")),
  c("cfi","tli","rmsea","srmr"))
```

If we compare the results from the ESEM approach with the invariance test reported above, we can see that the fit indices are somewhat better. Above, our CFA model did not show the best fit. Both the CFI and RMSEA showed somewhat less than desirable fit. Using ESEM, we see that the fit of the configural model is better (CFI=.947; RMSEA=.076) than the original fit (CFI=.943, RMSEA=.086). Further restrictions to both loadings and intercepts show that the data fits better using the ESEM approach, even when using more restrictive models.

**Limitations of ESEM**

ESEM is a relatively novel approach which has been used by some cross-cultural researchers already (e.g. Marsh et al., 2009; Vazsonyi et al., 2015). However, given the relative novelty of the method and small number of studies that have used it, some caution has to be taken. A recent computational simulation (Mai et al., 2018) suggests that ESEM has problems with convergence (e.g., the algorithm does not run), especially if the sample sizes are smaller (less than 200 or the ratio of variables to cases may be too small). Mai and colleagues recommended ESEM when there are considerable cross-loadings of items. In cases where cross-loadings are close to zero and the factor structure is clear (high loadings of items on the relevant factors), ESEM may not be necessary. Hence, ESEM might be an appealing method if a researcher has large samples and there are substantive cross-loadings in the model that cannot be ignored.

## Invariance Testing using Alignment

As yet another extension of CFA approaches, recently Multi-Group Factor Analysis Alignment (from here on: alignment) has been proposed as a new method to test scalar invariance (Asparouhov & Muthén, 2014). This method aims to address issues in MGCFA invariance testing, such as difficulties in establishing exact scalar invariance with many groups. The main difference between MGCFA and alignment is that alignment does not require equality restrictions on factor loadings and intercepts across groups.

Alignment's base assumption is that the number of noninvariant measurement parameters and the extent of measurement non-invariance between groups can be held to a minimum for each given scale through producing a solution that features many approximately invariant parameters and few parameters with large non-invariances. The ultimate goal is to compare latent factor means, therefore the alignment method estimates factor loadings, item intercepts, factor means, and factor variances. The alignment method proceeds in two steps (Asparouhov & Muthén, 2014). In the first step an unconstrained configural model is fitted across all groups. To allow the estimation of all item loadings in the configural model, the factor means are fixed to 0 and the factor variances fixed to 1. In the second step, the configural model is optimized using a component loss function with the goal to minimize the non-invariance in factor means and factor variances for each group (for a detailed mathematical description see: Asparouhov & Muthén, 2014). This optimization process terminates at a point at which "there are few large noninvariant measurement parameters and many approximately noninvariant parameters rather than many medium-sized noninvariant measurement parameters." (Asparouhov & Muthén 2014, p. 497). Overall, the alignment process allows for the estimation of reliable means despite the presence of some measurement non-invariance. Asparouhov and Muthén (2014) suggest a threshold of 20% non-invariance as acceptable. The resulting model exhibits the same model fit as the original configural model but is substantially less non-invariant across all parameters considered. Alignment was developed in Mplus (Muthén & Muthén, 2018), but adaptations are also becoming available in R. Here, we show an example of an alignment analysis using the sirt package (Robitzsch, 2019).

**How to run a Multi-Group Factor Analysis Alignment in R**

The *sirt* package provides three useful functions *invariance_alignment_cfa_config*, *invariance.alignment,* and *invariance_alignment_constraints*. These functions build upon each other to provide an easy implementation of the alignment procedure. We use again the example of the helping scale.

We initially fit a configural model across all countries. The *invariance_alignment_cfa_config* makes this straightforward. The function has two main arguments *dat* and *group*; dat takes a data frame as input that only contains the relevant variables in the model. It is important to stress that alignment can currently only fit uni-dimensional models. In our case we select all help variables (help1,…, help7) from the example data set (dat=example[paste0("help", 1:7) – the use of the paste0 command selects only the help items from 1 to 7 from the example data set). The group argument takes a grouping variable with the same number of rows as the data provided to the dat argument. In the current case we provide the country column from our data set.

```
par <- invariance_alignment_cfa_config(dat = example[paste0("help", 1:7)], group =
example$country)
```

The *invariance_alignment_cfa_config* function returns a list (in the current case named par) with λ (loadings) and ν (intercepts) for each country and item in addition to sample size in each country and the model fitted. The output of this function can be directly processed in the *invariance.alignment* function. Prior to that the invariance tolerance needs to be defined. Asparouhov and Muthén (2014) suggested 1 for λ and 1 for ν. Robitzsch (2019) utilises a stricter criterion of λ = .40 and ν = .20. These tolerances can be varied using the align.scale argument of the *invariance.alignment* function. The first value in a vector provided in this argument represents the tolerance for ν, the second the tolerance for lambda λ. Further, alignment power needs to be set in the align.pow argument. This is routinely defined as .25 for λ and ν, respectively. Last, we need to extract λ and ν from the output of the *invariance_alignment_cfa_config* function and provide them to the lambda and nu argument of the *invariance.alignment* function.

```
mod1 <- invariance.alignment(lambda = par$lambda, nu = par$nu, align.scale = c(.2, .4),
align.pow = c(.25, .25))
```

The resulting object can be printed to obtain a number of results such as aligned factor loadings in each group and aligned means in each group. We are focusing on the relevant indicators of invariance. $R^2$ values of 1 indicate a greater degree of invariance, whereas values close to 0 indicate non-invariance (Asparouhov & Muthén, 2014).

```
mod1$es.invariance["R2",]
```

In our current analysis we obtain an $R^2$ of .998 for loadings and 1 for intercepts. This indicates that essentially all non-invariance is absorbed by group-varying factor means and variances.

Alignment can also be used to assess the percentage of non-invariant λ and ν parameters using the *invariance_alignment_constraints* function. This function takes the output object of the *invariance.alignment* function as input. Additionally, ν and λ tolerances can be specified.

```
cmod1 <- invariance_alignment_constraints(mod1, lambda_parm_tol = .4, nu_parm_tol = .2)

summary(cmod1)
```

We found that for both factor loadings and factor intercepts none of items exhibited substantial non-invariance (indicated by 0% for the Percentage of non-invariant item parameters). Asparouhov and Muthén (2014) suggested a cut-off of 25% non-invariance to consider a scale non-invariant.

**Limitations of Alignment**

While alignment is a useful tool for researchers interested in comparisons with many groups, it also has limitations. First, convergence again can be an issue, especially for two group comparisons (Asparouhov & Muthén, 2014). Second, the alignment technique is currently limited to uni-factorial constructs precluding the equivalence test of higher order constructs or more complex theoretical structures. Finally, it is a new method and more work may be necessary to understand practically and theoretically meaningful thresholds and cut-offs in a cross-cultural context.

## Differential item functioning using ordinal regression (Item response theory)

One of the most common techniques for detecting differential item functioning (DIF) within the IRT family are logistic regression methods, originally developed for binary response items. It is now possible to use Likert-type scale response options (so-called polytomous items) as ordinal response options. The central principle of DIF testing via logistic regression is to test the probability of answering a specific item based on the overall score of the instrument (as a stand-in for the true trait level, as discussed above). DIF testing via logistic regression assumes that the instrument tested is uni-dimensional. The crucial tests evaluated are whether a) there are also significant group effects (e.g., does belonging to a specific group make answering an item easier or more difficult, over and above the true trait level) and b) there are group by ability interactions (e.g., trait effects depend on the group a person belongs to). The first test estimates uniform item bias and the second test estimates non-uniform item bias. Hence, the procedure uses a nested model comparison (similar to CFA invariance testing). A baseline model only includes the intercept. Model 1 includes the estimated true trait level, model 2 adds a dummy for the group (culture) effects and model 3 includes the group (culture) by trait interaction.

We have a number of options to test whether DIF is present. First, it is possible to compare overall model fit using the likelihood ratio chi-squared test. Uniform DIF is tested by comparing the difference in log likelihood values between model 1 and 2 (df = 1). Non-uniform DIF is tested by comparing models 2 and 3 (df = 1). It is also possible to test whether there is a total DIF effect by directly comparing model 1 vs model 3 (df = 2), testing for the presence of both uniform and non-uniform item bias together. This particular approach uses significance tests based on the difference in chi squares.

As we discussed above, chi square tests are sample size dependent, hence a number of alternative tests have been proposed. These alternatives focus on the size of DIF (hence they are effect size estimates of item bias) rather than whether it is significant. There are two broad types: pseudo $R^2$ (the amount of variance explained by the group effect and group by trait interaction), and raw regression parameters as well as the differences in the regression parameters across models. The interpretation of the pseudo $R^2$ measures have been debated due to scaling issues (see discussions in Choi et al., 2011), but since we are interested in the differences between nested models, their interpretation is relatively straightforward and similar to normal $R^2$ difference estimates. Estimates lower than 0.13 can be seen as indicating negligible DIF, between 0.13 and 0.26

showing moderate DIF and above 0.26 large DIF (Zumbo, 1999). As outlined by Choi and colleagues, some authors have argued that these estimates are too large and lead to under-identification of DIF.

For the regression parameters, it is possible to examine the regression coefficients for the group and the group by trait effects as indicators of the magnitude (Jodoin & Gierl, 2001). It is also possible to examine the difference in the regression coefficient for traits across model 1 and 2 as an indicator of uniform DIF (Crane et al., 2004). If there is a 10% difference in the regression coefficients between model 1 and 2, then this can be seen as a practically meaningful effect (Crane et al., 2004). A convenient feature of the R package that we are describing is that it allows Monte Carlo estimations for detecting DIF thresholds, allowing a computational approach with simulated data for establishing whether items show DIF or not. In other words, the model creates simulated data to estimate how much bias is potentially present in our observed data. The downside is that it is computational demanding and this analysis may take a long time to complete (in our sample using 7 items and 2,000 participants, the analysis took over 60 min to complete).

One of the key differences of IRT based approaches compared to CFA is that it refers to differences in item performances between groups of individuals which are matched on the measured trait. This matching criterion is important because it helps to differentiate between differences in item functioning from meaningful differences in trait levels between groups. One of the crucial problems is how to determine the matching criterion if individual items have DIF. The specific package that we describe below uses an iterative purification process in which the matching criterion is recalculated and rescaled using both the items that are not showing DIF as well as group-specific item parameters for items that are found to show DIF. The program is going through repeated cycles in which items are tested and the overall matching score is recalibrated till an optimal solution is found (as specified by the user). This iterative approach is superior to using just the raw scores, but again these iterative processes are computationally more demanding. For more information on the specific steps and computation process, see Choi, Gibbons, and Crane (2011).

**Logistic Regression to test for DIF in R**

One relevant package that we describe here is lordif (Choi et al., 2011). We chose it because it provides a number of advanced features while being user-friendly. As usual, the package needs to be called as described in the online supplementary materials. We then need to select only the variables used for the analysis (note the use of the paste0 command again):

```
response_data <- example[paste0("help", 1:7)]
```

Importantly, the group variable needs to be specified as a vector and is included in a separate file (which needs to be matching to the main data file). In our case, we are using the package car to recode the data:

```
country <- car::recode(example$country, "'NZ' = 1; 'BRA' = 0")
```

The actual command for running the DIF analysis is straightforward. In our case, we specify an analysis using the Chi square test:

```
countryDIF<-lordif (response_data, country, criterion = "Chisqr", alpha = 0.001, minCell = 5)
```

As before, we create an output object which contains the results. The function is lordif, which first specifies the data set and then the vector which contains the sample or country information. We then have to make a number of choices. The important choice is to define what threshold we want to set for declaring an item as showing DIF. We can select among $\chi^2$ differences between the different models (criterion = "Chisqr", in which case we also need to specify the significance level using the alpha command), $R^2$ (criterion = "R2", we need to select the beta.change threshold, e.g., R2.change = 0.01) and the regression coefficients (criterion = "Beta", we need to select the beta coefficient change, e.g., beta.change = 0.10). These choices can make potentially substantive differences, we urge users to explore their data and decide what criteria is most relevant for their purposes.

A final decision is how to treat minimum cell size (called sparse cell). The analysis proceeds as an ordinal level analysis, if there are few responses to some of the response categories (e.g., very few people ticked 1 or 7 on the Likert scale). We need to specify the minimum value. The default is 5, but we could also specify higher numbers, in which case response categories are collapsed till the minimum cell size is being met by our data. This might mean that instead of having 7 response categories, we may end up with 5 categories only because the extreme response options were combined.

If we use $\chi^2$ differences as a criterion, item 3 for the helping scale is again flagged as showing item bias. McFadden's pseudo-R-square values suggest that moving from model 1 to model 2 increases the explained variance by 0.0100, compared to 0.0046 when moving from model 2 to model 3. Hence, uniform item bias is more likely to be the main culprit. The other pseudo-R2 values also show similar patterns. In contrast, if we use the R2 change criterion (and for example, a change of 0.01 as a criterion), none of the items are flagged as showing DIF.

The relevant code is:

```
countryDIF_r2_change <- lordif(response_data, country,criterion = "R2", R2.change = 0.01, minCell = 5)
```

This highlights the importance that selecting thresholds for detecting DIF have for appropriately identifying items that may show problems.

If we wanted to run the Monte Carlo simulation, we write this function (which specifies the analysis to be checked as well as the alpha level and number of samples to be drawn):

```
countryDIF_MC <-montecarlo(countryDIF, alpha = 0.001, nr=1000)
```

**Evaluation of Logistic Regresssion**

There are multiple advantages of using logistic regression approaches within the larger IRT universe. These techniques allow the most comprehensive, yet flexible and robust analysis of item bias. They assume a non-linear relationship between ability and item parameters, which are independent of the specific sample that is being tested. The data needs to be at least ordinal. Both purely statistical significance driven and effect-size based tests of DIF are possible. One distinct advantage is that the lordif package includes an iterative Monte Carlo approach to provide empirically driven thresholds of item bias. Visualization of item bias is also available through the package (see Choi et al., 2011 for details).

At the same time, there are also a number of downsides. First, as with a number of the other techniques mentioned above (dMACS, alignment), only unidimensional scales can be tested. Second, researchers need to specify thresholds for DIF and the specific choices may lead to quite different outcomes, especially if DIF sizes vary across items. Third, some of the tests are sensitive to sample size and cutoff criteria for DIF differ across the literature. The MonteCarlo simulations are an alternative to construct data-driven cut-offs, but they are computationally intensive. Finally, logistic regression typically requires quite large samples.

### Exploratory factor analysis (EFA) and Principal Component Analysis (PCA)

Exploratory factor analysis as a group of statistical techniques is in many ways similar to CFA, but it does not presuppose a theoretical structure. EFA is often used as a first estimation of the factor structure, which can be confirmed in subsequent studies with CFA. Alternatively, researchers may use EFA to understand why CFA did not show good fit. Therefore, EFA is an integral method in the research process and scale development, either as the starting point for exploring empirical structures at the beginning of a research project or for identifying problems with existing scales.

Similar to CFA, the correlations between all items in a test are used to infer the presence of an underlying variable (in factor analytic terms). The two main approaches are proper Exploratory factor analysis (EFA) and Principal component analysis (PCA). The two methods differ conceptually: PCA is a descriptive reduction technique and EFA is a measurement model (e.g., Borsboom, 2006;

Tabachnick & Fidell, 2007), but practically they often produce similar results. For both methods, Pearson correlations (or covariances) between observed indicators are used as input, and a component or factor loading matrices of items on components or factors (indicating the strength of relationship of the indicators to the factor in EFA) are the output. For simplicity, we will use the term factor to refer to both components in a PCA and factors in an EFA. More detailed treatment of these methods can be found in other publications (Field et al., 2012; Gorsuch, 1983; Tabachnick & Fidell, 2007).

Figure 6.4 shows the main parts of an EFA model, which is conceptually similar to the CFA model. One of the major differences is that all items are allowed to load on all factors. As a result, decisions need to be made about the optimal assignment of loadings to factors (a rotational problem, see below) and what constitutes a meaningful loading (an interpretational problem). Items often show cross-loading, in which an item loads highly on multiple factors simultaneously. Cross-loadings of factors may indicate that an item taps more than one construct or factor (item complexity), problems in the data structure, circumplex structures (there is an underlying organization of the latent variables), or it may indicate factor or component overlap (see Field et al., 2012; Tabachnick & Fidell, 2007). As a crude rule of thumb, factor loadings above .5 on the primary factor and lack of cross-loadings (the next highest loading varies by at least .2) might be good reference points for interpretation.

The principal aim of an EFA is to describe the complex relationship of many indicators with fewer latent factors, but deciding on the number of factors to extract can be tricky. Researchers often use either theoretical considerations and expectations (e.g., the expectation that five factors describe human personality, McCrae & Costa, 1987) or statistical techniques to determine how many factors to extract. Statistical factors take into account how much variance is explained by factors, which is captured by eigenvalues. Eigenvalues represent the variance accounted for by each underlying factor. They are represented by scores that total to the number of items. For example, an instrument with twelve items may capture up to 12 possible underlying factors identified by a single indicator (each item is its own factor). Each factor will have an eigenvalue that indicates the amount of variation that this factor accounts for in the items. The traditional approach to determining the appropriate number of factors was based on Cattell's scree plot and Kaiser's criterion that indicates that factors with eigenvalues greater than 1 (e.g., a factor that explains more variance than any item alone is worth extracting). These methods have been criticized for being too lenient (e.g., Barrett, 1986). Statistically more sophisticated techniques such as Horn's (1965) parallel analysis are now more readily available. Parallel analysis compares the resulting eigenvalues against the eigenvalues obtained from random datasets with the same number of variables and adjusts the obtained eigenvalues (we briefly describe options in the online supplementary materials).

Once a researcher has decided how many factors to extract, a further important question is how to interpret these factors. First, are the factors assumed to be uncorrelated (orthogonal or independent) or correlated (oblique or related). Latent factor intercorrelations can be estimated when oblique rotation is used (Gorsuch, 1983, pp. 203–204). The choice of rotation is primarily a theoretical decision.

Factor rotations are mathematically equivalent. If more than one component or factor has been identified, an infinite number of solutions exist that are all mathematically identical, accounting for the same amount of common variance. These different solutions can be represented graphically as a rotation of a coordinate system with the dimensions representing the factors and the points representing the loadings of the items on the factors. An example of such rotation is given in Figure 6.5. Mathematically, the two solutions are identical. Conceptually, we would draw very different conclusions from both versions of the same rotation. This is the core problem with interpreting factor structures across different cultural groups because this rotational freedom can lead to two groups with identical factor structures showing very different factor loadings (see Table 6.1 for an example – even though the solutions are mathematically identically, they show noticeably different factor loadings). As a consequence, researchers need to rotate their factor structures from the individual groups to similarity before any decisions about factor similarity can be made. The method of choice is orthogonal Procrustes rotation in which the solution from one group is rotated towards the factor structure of the reference group. A good option to decide on the reference group might be to a) use the group in which the instrument was first developed, b) use the larger group (since this reduces the risk of random fluctuations that are more likely to occur in smaller groups) or c) select the group that shows a theoretically clearer or meaningful structure.

After running the Procrustes rotation, the factor structures can be directly compared between the cultural groups. To determine how similar or different the solutions are, we can use a number of different approaches. The most common statistic for comparing factor similarity is Tucker's coefficient of agreement or Tucker's φ (van de Vijver & Leung, 1997). This coefficient is not affected by multiplications of the factor loadings (e.g., factor loadings in one group are multiplied by a constant) but is sensitive to additions (e.g., when a constant is added to loadings in one group). The most stringent index is the correlation coefficient (also called identity coefficient). Other coefficients such as linearity, or additivity can be computed, if necessary (for a general review of these options, see Fischer & Fontaine, 2010; van de Vijver & Leung, 1997). Factor congruence coefficients vary between 0 and 1. Conventionally, values larger than .85 can be judged as showing fair factor similarity and values larger than .95 as showing factor equality (Lorenzo-Seva & ten Berge, 2006), values lower than .85 (ten Berge, 1986) are indicative of incongruence. However, these cut-off criteria might vary for different instruments, and no formal statistical test is associated with these indicators (Paunonen, 1997). It is also informative to compare the different indicators, if they diverge from each other this may suggest that there is a problem with the factor similarity.

**Procrustes Rotation with two Groups using R**

The relevant packages that we need are psych (Revelle, 2018) and GPArotation (Bernaards & Jennrich, 2005). We first need to load these packages and load the relevant data (see the online supplementary materials for further info).

The first step is to run the factor analysis separately for both samples. We could run either a principal component analysis (using the *principal* function) or factor analysis (using the *fa* function).

```
nz_fa <- fa(nz_example[,-1], nfactors = 2, rotate = "varimax")
```

We call the factor analysis function (fa**()**) from the psych package and specify the data we are working on (New Zealand data frame without the first column that contains the column with the country information: nz_example [,-1]), the number of factors we want to extract from the data (nfactors **=** 2), and the rotation we want to use (rotate **=** "varimax"). Because we have a theoretical expectation, we request two factors in each country. We also specify an orthogonal varimax rotation, because we expect the factors to be uncorrelated. Last, we assign the result to an object (nz_fa **<-**) for later use.

Next, we perform the same action for the Brazilian data using the same procedure:

```
br_fa <- fa(br_example[,-1], nfactors = 2, rotate = "varimax")
```

In the next step we can directly rotate the factor loading matrices using New Zealand as target matrix and Brazil as loading matrix. In the ccpsyc package, we included a function called prost, which we adapted from the TargetQ function within the psych package in order to provide the identity coefficient and Tucker's Φ in a straightforward fashion. The *prost* function takes two factor matrices as input and returns the identity coefficient and Tucker's Φ.

```
prost(NZ.fa$loadings, BRA.fa$loadings)
```

We call the output from the factor analyses that we ran above. The $ sign specifies that we only use the factor loadings for the procrustean rotation. In our example, we rotated the Brazilian sample to similarity with the New Zealand sample (first position in our command). We chose NZ as a reference category because NZ is culturally probably more similar to the US where the instrument was developed and the NZ sample was larger, therefore, the solution was expected to be more stable. Tucker's phi was .97 and .98 respectively, indicating the factor structures to be equal. The correlation coefficients on the other hand were lower, .81 for the first factor (helping) and .89 for factor 2 (voice). In addition to the overall factor congruence coefficients, it is also informative to examine the factor structure after rotation to see which items may show higher or lower loadings in each sample.

To do this the prost function has an argument (rotated) which can be set to TRUE (rotated = TRUE). The output now contains the rotated matrix.

```
prost(NZ.fa$loadings, BRA.fa$loadings, rotate=TRUE)
```

Researchers can visually compare the differences in factor loadings between the samples to identify any items that may perform differently. These tests are rather subjective and not clear guidelines are available. The interpretation depends on the overall strength of the factor loadings, the number of items and difference in item performance. Unfortunately, no specific statistical tests are available through R that provide more objective tests at the item level. In our example data, one of the voice items showed strong cross-loadings in one sample. Removing this item, the correlation coefficients increased to .87 for helping and .94 for voice, still not meeting sufficient standards for invariance for at least factor 1 using the more stringent correlation coefficient as a criterion.

**Limitations of the Technique**

A major weakness is that the procedure focuses on the congruence at a factorial level, answering whether similar structures are found in each group compared with the reference group. Therefore, we can only establish configural invariance or structural equivalence. Procrustes rotation does not allow to test for metric invariance as the analysis stays at the factor rather than the item level. Individual items may still show substantial loading differences, and the overall factorial similarity might be misleading. For example, research on the structure of the Eysenck Personality Questionnaire has shown that this issue is not without debate (Bijnen et al., 1986; Bijnen & Poortinga, 1988). The number of items and factors may also influence the congruence levels that a researcher can expect to find (Paunonen, 1997). As mentioned above, it is useful to examine the target and target-rotated loadings as well as the difference between the target loadings and the loadings in a norm group to identify potential anomalies in addition to examining any overall congruence coefficient. This may reveal important and useful information of cross-cultural similarities and differences. Nevertheless, Procrustes rotation can be a useful technique at initial research stages. It is also a useful technique if the data does not allow for a full Multi-Group Confirmatory Factor Analysis to be fitted, for example due to a limited number of indicators per construct. Further, Procrustes rotation can be a useful technique to examine the fit of cultures observed structure to an idealised loading matrix of a construct. This process allows a researcher to investigate whether culture level variables significantly impact structural fit.

**Testing invariance with more than two groups**

The most flexible and versatile technique for testing invariance with more than two groups is multi-group CFA. The approach can easily handle more than two groups and no adjustments to the set-up and testing need to be done. One of the challenges is that lavaan provides $\chi^2$ values for each individual group, but only overall fit indices. Since $\chi^2$ values are sample size dependent, unless sample sizes are equal, it might be difficult to determine which samples and items are problematic when examining an overall poorly fitting multi-group model. One option is to estimate the individual group models because it will provide important clues about possible problems.

The logistic regression approach implemented in lordif can accommodate more than 2 groups. However, the visualization of item bias becomes hard to interpret when more than 2 groups are used.

**Dealing with more than two groups for EFA/PCA**

To conduct an EFA/PCA with Procrustes rotation for multiple groups at the same time different methods can be used. Nevertheless, they all require identical steps to set up the data for analysis (We show how to prepare the data for analysis in the supplementary materials).

**Target rotation with multiple groups**

Conducting an EFA/PCA with a Procrustes rotation to determine configural invariance with more than two samples requires some theoretical considerations:

**Create an ideal matrix as reference group.** In this option a researcher constructs a loading matrix that represents theoretically assumed loadings on a factor with 1, non-loadings with 0, and negative loadings with -1. This approach is most useful for established measures for which strong theoretical assumptions about the structure exist, such as personality traits. This approach yields insight into the fit of the data from each sample compared to the proposed ideal. Below, we show an example using the *prost* function of the ccpsyc package. We provide an example of how to create an ideal matrix in the online supplementary materials.

```
lapply(PCA, function(x){prost(x$loadings, ideal)})
```

**Use the matrix from the instruments' origin.** Most questionnaires were and are developed in a Western context. Therefore, a researcher might want to examine how well a newly translated instrument reproduces the structure in regards to the original structure. While this approach can be useful to validate the structure of newly translated instruments in relation to existing data structures, a substantive drawback of this approach is that it posits the origin culture's structure as de facto correct solution. In our example, we used the results from the first PCA analysis as the target matrix.

```
lapply(PCA, function(x){prost(x$loadings, PCA[[1]]$loadings)})
```

**Creating a pan-cultural matrix.** In this approach an average weighted or unweighted correlation matrix of the items in the structure is created across all cultures of interest. It creates an average matrix, averaging correlations across all items and samples. This does not give priority to any specific cultural group. The resulting correlation matrix can be used as an input to factor analysis and provides a culture-general reference factor loading matrix. This average cultural solution can then be used as the comparison standard for all the individual samples. This approach yields insight into how much each sample corresponds with a common factor solution across all cultures. We show how to create a pan-cultural matrix in the online supplementary materials. Problems emerge if there is misfit in one or more of the samples and the processes needs to become iterative through pruning mis-fitting samples.

```
lapply(PCA, function(x){prost(x$loadings, PAN_PCA$loadings)})
```

**Choosing a target based on sample criteria.** Sample criteria can also be informative when choosing a rotational target. Considerations such as sample size in each culture and factor simplicity can guide the selection (e.g., the largest sample or the sample with the simplest structure may be selected as comparison). This approach can yield good statistical results but might limit the generalizability of the results and the theoretical interpretation.

**Running all pairwise comparisons.** While this approach is free of theoretical considerations, it is only typically feasible or interpretable for a small number of cultures. It is possible to use computational approaches for running cluster analyses of factor similarity, in which we case we attempt to identify groups of samples that show similar factor structures. In the absence of such computational solutions, it might be difficult to make decisions about invariance as one sample might show poor invariance to a second sample, but good invariance to relation to a third sample.

## Overall comparison of methods

Which method should you use? There are a number of theoretical questions that can guide you to decide which approach might be best. A first important question is the data that is available. If only ordinal data are available, then IRT remains the most appropriate option. There are options to run CFA and EFA/PCA with ordinal data in R (after computation of polychoric correlations, but these require some intermediate steps). A second important question is whether the researcher has a theoretical model to test or whether the analysis is exploratory. In the former case, both CFA and logistic regression are good options and can be combined to get the most comprehensive insight into the data (Meade & Lautenschlager, 2004b). In the latter case, EFA and PCA are better. New methods such as ESEM are a hybrid that combined EFA with CFA techniques. Third, only CFA-derived methods and logistic regression allow invariance tests at the individual level and statistical tests of DIF. In contrast, EFA and PCA with Procrustes rotation allow only analyses at the scale or instrument level, therefore, they do not provide metric and scalar invariance tests that then would allow the researcher to compare scores directly across groups. Fourth, all techniques described here require decent (ideally N>200) sample sizes, with logistic regression, CFA and associated techniques such as ESEM and alignment being the most sample-size hungry techniques (Meade & Lautenschlager, 2004a). One major drawback for many practical approaches is that logistic regression and alignment (within the CFA-domain) require analyses of unidimensional scales, whereas CFA in particular is versatile in accommodating more complex theoretical structures. Finally, both CFA and logistic regression techniques provide effect size estimates of DIF, which give researchers options to decide how much of bias is too much. Only logistic regression at this moment provides an easily available (but computationally demanding) way to derive empirically derived item bias parameters.

## Summary

Free software for testing invariance at both basic and advanced levels is now available and is easy to use. Comparisons without establishing or testing invariance and equivalence are open to alternative explanations, therefore, invariance testing is paramount. We have highlighted in this article a number of methods and conceptual approaches to allow researchers to test for invariance of their own data. Easy to implement approaches that are free are available to researcher and hopefully will improve the standards and quality of cross-cultural research.

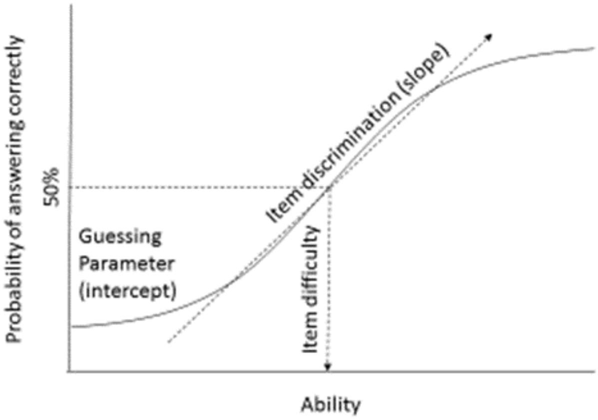



*Figure 6.1 Item difficulty, item discrimination and guessing parameters in a single group*

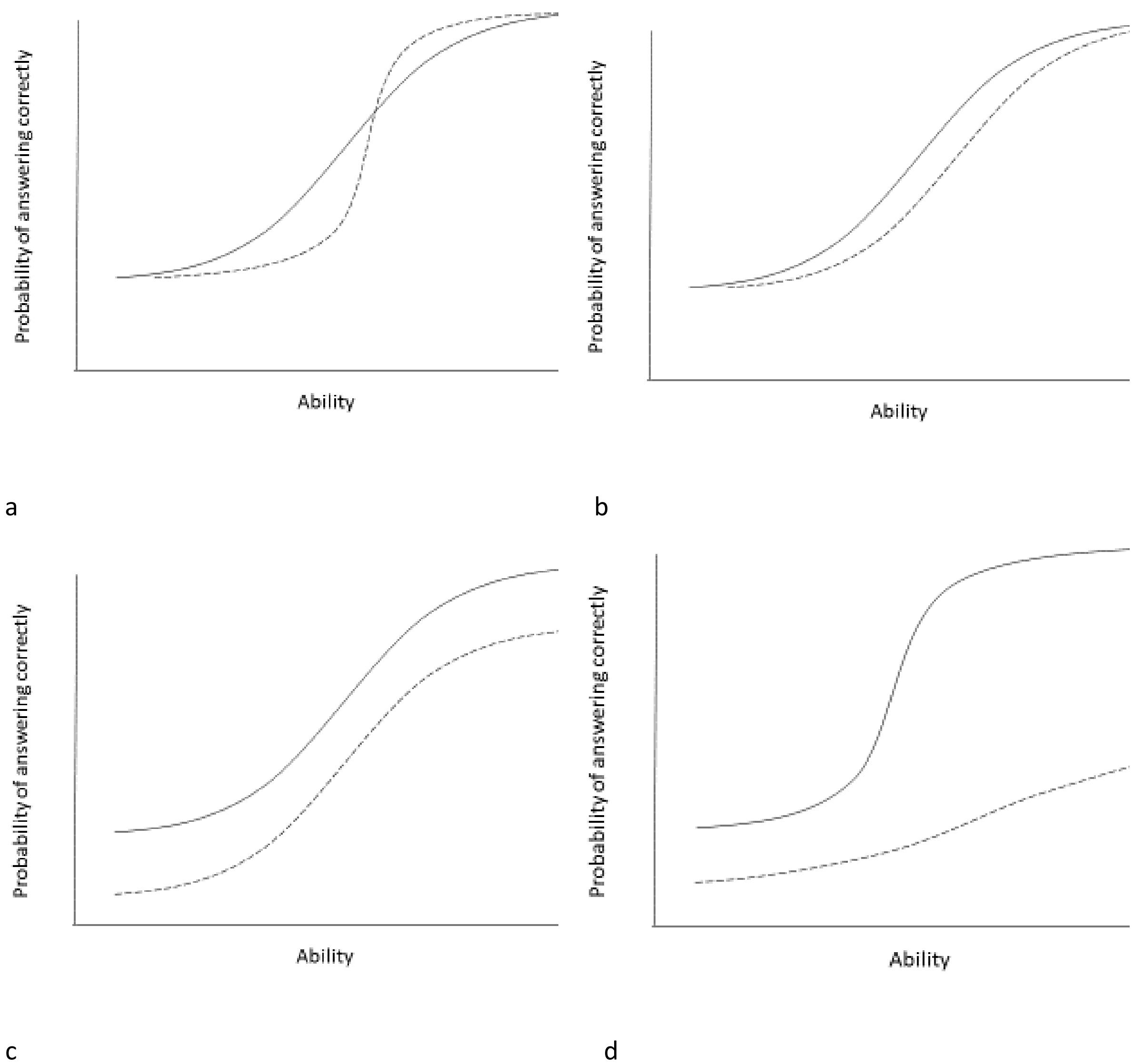



*Figure 6.2 Examples of differential item functioning in two groups.*

The panels show differential item functioning curves for two groups (group 1 indicated by solid line, group 2 indicated by a broken line).

Panel a) shows two groups differing in item discrimination (slope differences). The item differentiates individuals less well in group 1. This is an example of non-uniform item bias

Panel b) shows two groups with different item difficulty. The item is easier (individuals with lower ability are able to correctly answer the item with 50% probability) for the group 1 and more difficult for group 2. Individuals in group 2 need higher ability to answer the items correctly with a 50% probability. This is an example uniform item bias.

Panel c) shows differential guessing or intercept parameters. Group 1 has a higher chance of guessing the item correctly compared to group 2. Scores for group 1 on this item are consistently higher than for group 2, independent of the individual's underlying ability or trait level. This is an example of uniform item bias.

Panel d) shows two groups differing in all three parameters. Group 1 has a higher guessing parameter, the item is easier overall, but also discriminates individuals better at moderate levels of ability compared to group 2. This is an example of both uniform and non-uniform item bias.

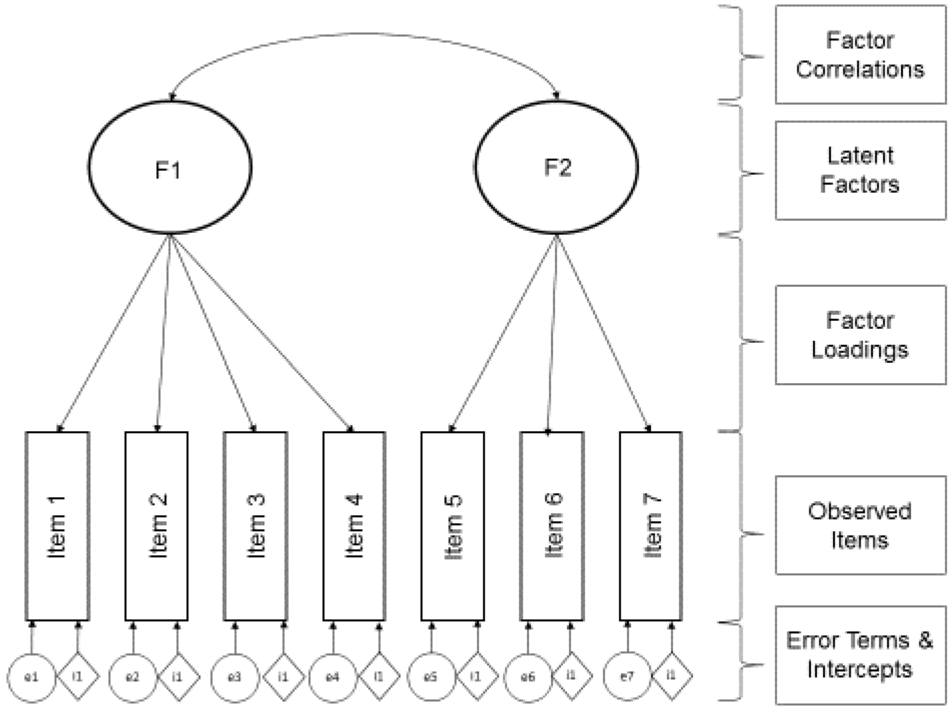



*Figure 6.3 Example of Confirmatory Factor Analysis Model*

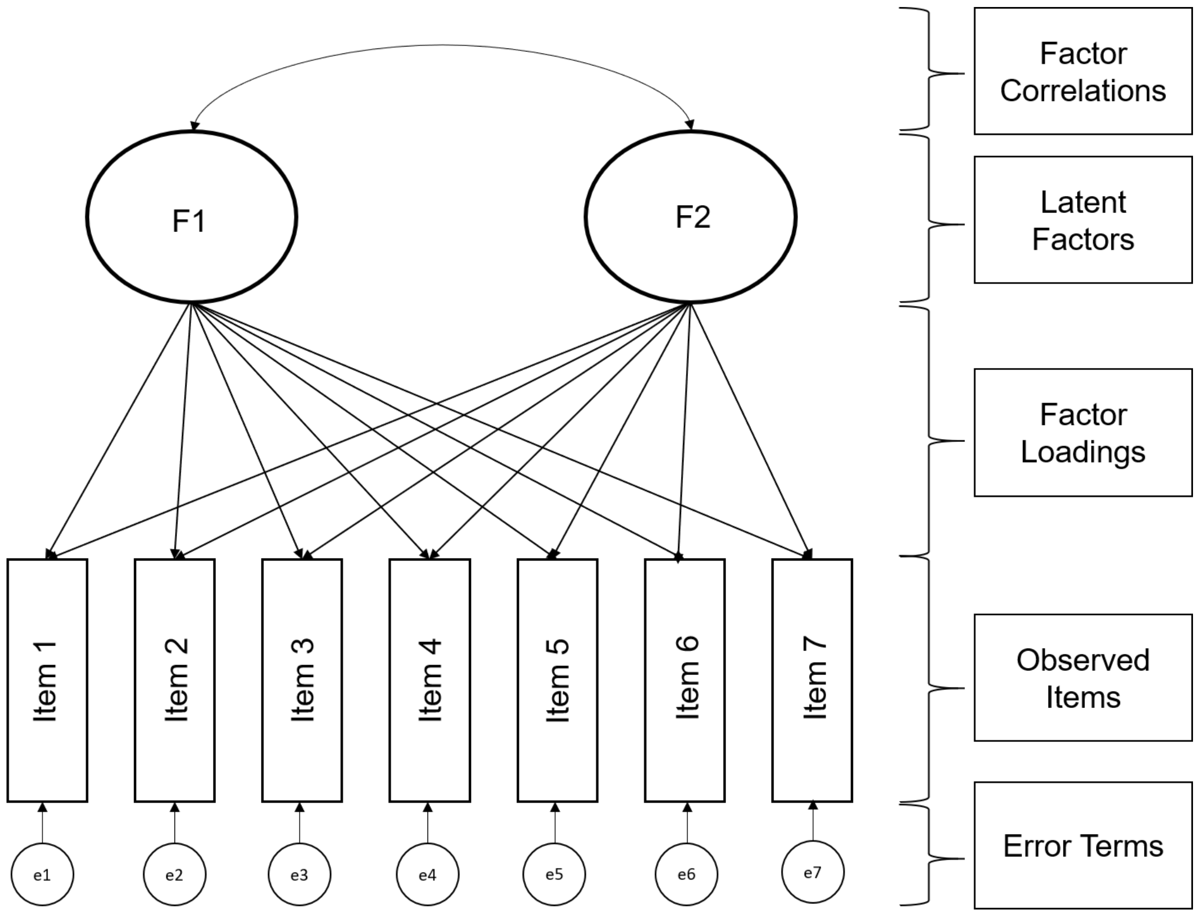



*Figure 6.4 Visual representation of an EFA model*

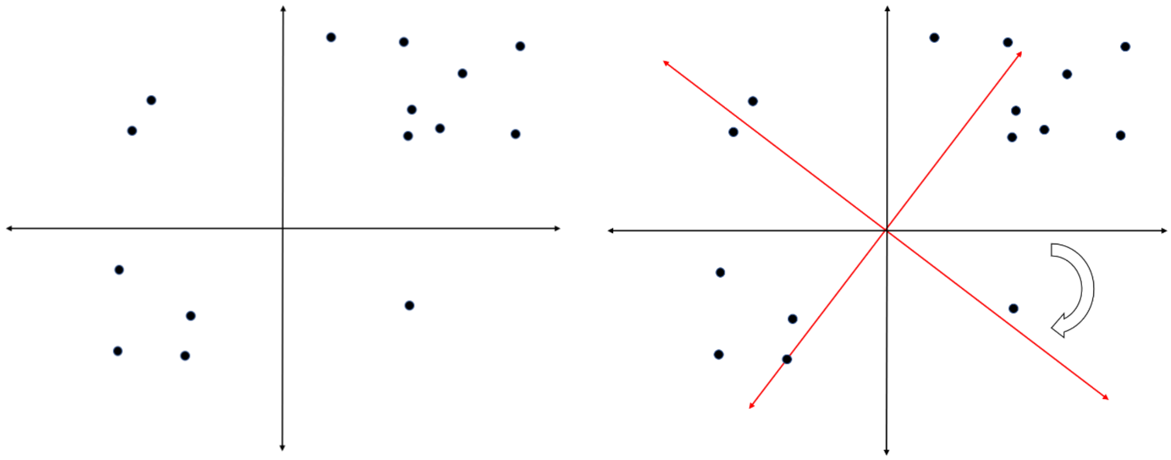


*Figure 6.5 Visualization of Factor Rotations.*

*Table 6.1 An example where two groups identical structures show different factor loadings.*

| | Factor 1 | Factor 2 | | Factor 1 | Factor 2 |
|---|---|---|---|---|---|
| Item 1 | .65 | .30 | | .67 | .19 |
| Item 2 | .66 | .30 | | .69 | .15 |
| Item 3 | .69 | .21 | | .80 | .25 |
| Item 4 | .82 | .24 | | .80 | .25 |
| Item 5 | .79 | .33 | | .67 | .32 |
| Item 6 | .79 | .28 | | .71 | .31 |
| Item 7 | .70 | .34 | | .39 | .59 |
| Item 8 | .44 | .67 | | .22 | .79 |
| Item 9 | .35 | .80 | | .19 | .81 |
| Item 10 | .26 | .81 | | .23 | .76 |
| Item 11 | .30 | .78 | | .43 | .59 |
| Item 12 | .30 | .83 | | .23 | .73 |

## Study 6: The Cross-Cultural Validity of the Five-Facet Mindfulness Questionnaire across 16 Countries[10]

### Preface

In the previous studies of this thesis, I focused on individual level predictors of mindfulness and conceptualizations of mindfulness in a single cultural context. In this final study I aimed to expand the research on trait mindfulness by providing a first global comparative perspective on the measurement of trait mindfulness. Building on the standards and procedures set out in study five, in this study I led a global team of researchers to examine appropriateness of the FFMQ as measurement tool of mindfulness across cultures. The study addressed a number of goals: First, to investigate the universality of the five-factor model of mindfulness and the measurement equivalence of the Five Factor Mindfulness Questionnaire (FFMQ); and second, to examine potential culture level factors that might predict the appropriateness of the FFMQ as measurement tool.

[10] This study has been previously published in *Mindfulness*: Karl, J. A., Méndez Prado, S. M., Gračanin, A., Verhaeghen, P., Ramos, A., Mandal, S. P., Michalak, J., Zhang, C.-Q., Schmidt, C., Tran, U. S., Druica, E., Solem, S., Astani, A., Liu, X., Luciano, J. V., Tkalčić, M., Lilja, J. L., Dundas, I., Wong, S. Y. S. Y., … Fischer, R. (2020). The Cross-cultural Validity of the Five-Facet Mindfulness Questionnaire Across 16 Countries. Mindfulness. https://doi.org/10.1007/s12671-020-01333-6.
Minor revisions and stylistic changes have been made to the manuscript to establish coherence with the rest of the thesis

Is mindfulness a cultural universal? Mindfulness is a principle originating from Buddhist tradition, which was first exported to the West and subsequently exported back to the East in the form of therapeutic interventions and psychological measurements. How valid are such measures to capture mindfulness across cultures? What can cross-cultural research using these measures reveal about mindfulness as a potentially universal psychological trait? In Western psychology, mindfulness is often defined as “paying attention in a particular way; on purpose, in the present moment, and nonjudgmentally” (Kabat-Zinn, 1994, p. 4). Such definitions provided ground for the development of a broad range of mindfulness measures. A combined analysis of multiple available measures resulted in the widely used Five-Facet Mindfulness Questionnaire (FFMQ: Baer et al., 2006), which conceptualizes mindfulness as a higher-order factor subsuming five facets: Acting with Awareness, Non-Judging, Non-Reacting, Describing, and Observing. The measure has been employed in many different cultures and translations exist in major language groups including German, Spanish, Portuguese, and Chinese. The continued research on mindfulness across cultures using the FFMQ indicates an implicit claim to universalism of the five-facet structure of mindfulness. The current approach to the measurement of mindfulness can be considered an imposed-etic approach (Berry, 1989), because mindfulness as a concept originated in a specific Buddhist context, but was transformed into a measurement instrument through a Western lens in measures, such as the FFMQ, and subsequently exported globally to assess mindfulness in different cultures.

A crucial step to support the universality of the construct of mindfulness is to establish measurement equivalence across cultural groups. Equivalence in the current context refers to the comparability of measured scores between cultures and can be broken down into three levels that can be empirically assessed; structural equivalence, metric equivalence, and scalar equivalence (Fontaine, 2005; Vandenberg & Lance, 2000). Structural equivalence implies that the same items can be used to measure the same latent constructs across cultures (Fischer & Fontaine, 2010). In other words, measures show structural equivalence if the same items are used across cultures and these items form the same dimensional structure in all cultures. For example, the item “I’m good at finding words to describe my feelings" would be associated with the Describing facet in all cultures. Metric equivalence implies that items have similar loading strength on the underlying constructs. For example, the item "I'm good at finding words to describe my feelings" would mean equally good indicator of the Describing facet in all cultures (the factor loadings are statistically similar). Finally, scalar equivalence implies that the item intercepts are identical. In other words, respondents with the same level of mindfulness overall would answer identically to each individual question in all cultures and their answers are not affected or shifted by response biases such as acquiescence bias (yes-saying), different referent standards (e.g., reference group effects) or differences in social desirability of a construct across groups (Heine et al., 2002; Van De Vijver & Leung, 2010).

Importantly, these levels of equivalence address the measurement properties of a scale across groups, but do not provide insight into potential domain under-representation across groups. Domain under-representation is present if a concept differs in conceptual scope across cultures, by missing important theoretical elements of the construct within specific cultural settings. In the case of mindfulness, during its transition from a Buddhist context into a Western secular context, metaphysical elements were often omitted to increase the diffusion of the practice (Kucinskas, 2014, 2018).

Whether the FFMQ is equivalent, and at which level, holds important implications for cross-cultural research on mindfulness. Structural equivalence allows exploration of the basic structure of a measure, i.e., if items relate to the proposed theoretical variable. Metric equivalence allows for the cross-cultural comparison of the correlations and score patterns, but no conclusions about cultural differences in mindfulness as a theoretical construct can be made. Only under the condition of scalar equivalence can researchers directly compare mean scores. In other words, researchers can investigate the dimensionality of mindfulness with structural equivalence, can compare the relationship of mindfulness with other measures across cultures with metric equivalence, and can directly compare cultural groups with scalar equivalence. Non-equivalence across a large number of cultural groups indicates that the FFMQ is not a suitable tool for cross-cultural research and that further research is necessary to establish a conceptualization of mindfulness that is valid across cultures.

A crucial part of testing for equivalence is to determine the theoretical structure which can be tested across groups. For the FFMQ, a number of structures have been suggested: a five-facet model in which the individual facets are subsumed under one (Baer et al., 2006) or two higher-order factors (Tran et al., 2013), and five correlated facets without a higher order factor (Van Dam et al., 2012). Further, a number of studies have suggested that the FFMQ should be modelled with positive and negative item-wording factors (Aguado et al., 2015; Van Dam et al., 2012). These item-wording factors model participants' differential responding to positively and negatively worded questions, improving the fit of the structure.

Examining most of the above described possibilities, the FFMQ could be modelled as: 1) five correlated facets with no higher order factor, 2) five correlated facets with uncorrelated methods factors for negatively and positively worded items, 3) five facets subsumed under a single higher-order factor, 4) five facets subsumed under a single higher-order factor with uncorrelated methods factors, 5) five facets subsumed under a single higher-order factor with correlated methods factors, or 6) five facets subsumed under a single higher-order factor with correlated methods factors which in turn are also correlated with the higher-order factor. A visualization of the proposed models can be found in Figure 7.1. Overall, the first step in determining whether the FFMQ is equivalent across cultures is determining the best fitting model within each culture. This analysis will provide first insights into the best conceptual representation of mindfulness across the different contexts in which the instrument has been applied. The most common structure can then be directly tested across all the sites for which data are available.

What is needed is an examination of contextual variables that may influence the replicability of the structure of the FFMQ across groups. This can be achieved by focusing on three major cultural dimensions that might be of relevance for mindfulness. First, monumentalism-flexibility (Minkov et al., 2018) captures important aspects of the stability of self. Minkov et al. (2018) described this important culture-level axis as: “Monumentalism is a metaphor for a cultural tendency to encourage people to be like a monolithic monument: proud, stable, and consistent (made of the same substance outside and inside). Flexibility is the opposite cultural tendency, favoring a modest self-regard, duality, and adaptability.” (p. 12). In other words, monumentalism can be thought of representing self-consistency (being the same person regardless of context). In contrast, flexibility is related to situation-specific behavioral responses, similar to the concept of ‘face’ in Asian societies which requires sensitivity to relationships and being a different person depending on the current context (Hwang, 1987). In line with this reasoning, Minkov et al. (2018) found that countries in Asia (e.g. China, Korea, Singapore) score high on Flexibility, whereas Western cultures cluster around the midpoint of this dimension. Therefore, if the FFMQ is closer to the ideal structure in Asian cultures, this would indicate that the proposed structure of the FFMQ represents an Eastern rather than a Western concept of mindfulness.

Further, individualism-collectivism expresses the level of embeddedness of the individual in the wider society (Hofstede, 2001; Triandis, 1995). Cultures in the West tend to have looser connections between the individual and the in-group, therefore they score higher on individualism. In contrast, individuals in Eastern cultures tend to be more embedded in the wider in-group, therefore those cultures tend to score lower on individualism (Hofstede, 2001; Minkov et al., 2017). A positive relationship between individualism and structural fit can be taken as an indication that the structure of mindfulness proposed in the FFMQ reflects a Western concept of mindfulness.

Tightness-looseness represents societal tendencies to be judgmental and punitive to deviations from cultural norms (Gelfand et al., 2011; Uz, 2015). The greater tendencies of individuals in looser cultures to be non-judgmental of one’s own and others’ deviations from norms might be more in line with current definitions of mindfulness underlying the FFMQ that see mindfulness as non-judgmental awareness (Kabat-Zinn, 1994). The expectation is therefore that societal looseness is associated with greater fit of the five-facet structure, indicating better fit of the FFMQ in cultures less judgmental of norm deviations.

In summary, the current research had three main goals. First, to determine the best fitting structure of the FFMQ in the individual samples and cultures. Second, to test this model across cultures to determine structural, metric, and scalar equivalence, which provides insights about whether the construct of mindfulness can be compared across cultures and what kind of comparisons can be made. The third and final goal was to examine what contextual variables may influence the stability and replicability of mindfulness measurement.

## Method

### Participants

The sample contained 8,541 participants from 16 countries sourced from previously published and unpublished studies: Australia (N = 165; community adults, Beshara et al., 2013), Austria (N = 973; community adults and students, Tran et al., 2013), Chile (N = 398; students, Schmidt & Vinet, 2015), China (N = 214; community adults, Ma et al., 2018), Germany (N = 529; students, Michalak et al., 2016), Spain (N = 1155; adults and students, Aguado et al., 2015), Hong Kong (N = 536; adults and students, Chung et al., 2014; Wong et al., 2017), Croatia (N = 242; adults and students, Gračanin et al., 2017), India (N = 300; community adults, Mandal et al., 2016), Norway (N = 466; adults and students, Dundas et al., 2013; Solem et al., 2015), New Zealand (N = 399, students, Karl & Fischer, 2019), Poland (N = 702; students, Radoń, 2014), Portugal (N = 251; community adults, Ramos et al., 2017), Romania (N = 293; adults and students, Astani, 2016; Druica & Ianole-Calin, 2018), Sweden (N = 495; students, Lilja et al., 2011) , and the United States of America (N = 1422; students, sample 1,2: Verhaeghen, 2018, sample 3,5: 2020; sample 4: Verhaeghen & Aikman, 2020). These studies were identified through a literature search on Google Scholar, PsychInfo, and the Web of science.

**Measures**

The analysis was conducted on the FFMQ-39 (Baer et al., 2006). All questions were measures on a 1-5 scale with verbal anchors in the respective language of the questionnaire (English: *Never or very rarely true*, *Rarely true*, *Sometimes true*, *Often true*, *Very often or always true*). The FFMQ was administered by the original authors in the language relevant to the cultural context. Across all facets and countries, the FFMQ showed acceptable to excellent reliabilities. All facets showed similar average reliability, with Non-Reacting showing the lowest average reliability. Due to space constraints, the full reliability table can be found in Appendix D material (Table 1).

**Data Analyses**

The data analysis has been pre-registered on the OSF (https://osf.io/5q4y2/?view_only=9cfbc3eb46be42c4bc7bffc7ecc286e6). All deviations from the proposed analysis are indicated where necessary. All CFAs and Multi-Group Confirmatory Factor Analyses (MGCFAs) were fitted using lavaan (Rosseel, 2012) in R (R Core Team, 2020), following the procedures set out by Fischer and Karl (2019). Each model was fitted with an MLM estimator to adjust for multivariate non-normality. Further, the variance of all latent variables was fixed to unity to allow estimation of all factor loadings, rather than fixing one item's loading to 1.

**Testing the structure of the FFMQ.** The first step was to test the individual proposed models of the FFMQ outlined in the introduction (Figure 7.1 provides a graphical summary). To determine the best fitting model in each sample separate CFAs were fitted for each of the models in each sample and compared for the relative fit. For each model the following fit indices were reported: $\chi^2$, degrees of freedom, $\chi^2$ / degrees of freedom (for a discussion see: Rasch, 1993), RMSEA (Steiger, 2016) with confidence intervals, SRMR, Comparative Fit Index (CFI; Bentler, 1990), $\hat{\gamma}$ (Fan & Sivo, 2007), and Bayesian Information Criterion (BIC; Schwarz, 1978). Acceptable fit for CFI and $\hat{\gamma}$ was defined as > .90 and good fit was defined as > .95 (Marsh et al., 2004), acceptable fit for the SRMR < .08 (Hu & Bentler, 1999). The RMSEA was evaluated following MacCallum et al. (MacCallum et al., 1996), with less than 0.01, 0.05, and 0.08 to indicate excellent, good, and mediocre fit respectively. To compare nested models, changes in ΔCFI > .01 and $\Delta\hat{\gamma}$ > .001 were used as indicating acceptable fit (Cheung & Rensvold, 2002). Further, the Bayesian Information Criterion (BIC) was used as the deciding criterion. Reductions of 10 between models were taken a strong indication of improvement and a reduction of 5 as a moderate indication of improvement (Berchtold, 2019; Raftery, 1995). Further, Vuong's test of non-nested model comparison with ML estimator (Vuong, 1989) implemented in the nonnest2 package (Merkle & You, 2018) was used to supplement the judgement whether models showed improved fit. The model that showed improved fit from the previous model in the majority of samples was selected.

**Testing the equivalence of the FFMQ.** A commonly employed method to test for measurement equivalence is MGCFA (Fischer & Karl, 2019; Milfont & Fischer, 2010). To test the equivalence of the ideal structure derived in the individual CFAs a MGCFA was used. If multiple samples were present for a culture, the individual samples that successfully converged in the previous step were merged to obtain an overall sample for each culture. To test for structural equivalence, item loadings and intercepts were allowed to vary between cultures. Structural equivalence was present if the model showed acceptable fit across all cultures. For metric equivalence, item loadings were constrained to be equal, but the intercept was allowed to vary between cultures. Metric equivalence was present if the constrained model fits well and there was no substantial drop in model fit from the prior, less restricted model. Substantial drop in fit was defined as ΔCFI ≥ .01, a more stringent cut-off of ΔCFI ≥ .002, and $\Delta\hat{\gamma}$ ≥ .001 (Cheung & Rensvold, 2002). Last, for scalar equivalence the intercept to be was constrained to be equal between cultures. Scalar equivalence was present if the constrained model showed good fit and no substantial drop in fit from the metrically restrained model. The same criteria for deciding on model fit was used as for metric equivalence.

**Exploratory Analyses.** In addition to these confirmatory analyses, several exploratory analyses were specified in the pre-registration. First, in case that the overall FFMQ would not be equivalent, the equivalence of the five facets individually (Acting with Awareness, Non-Reacting, Non-Judging, Observing, and Describing) would be tested. Further, if no equivalence of the FFMQ in most cultures was found alternative solutions using an exploratory factor analysis to determine the common factor solution across cultures would be explored. Finally, the effect of culture level-variables, such as individualism, monumentalism and tightness-looseness, on the appropriateness of the five-factor solution in different cultures was investigated. A Procrustes-analysis, examining the congruence of the loadings of each country's five-factor structure to an ideal solution where items' loadings on the factors were defined as ones and zeros following the original proposed structure of the FFMQ (Baer et al., 2006) was run. Tucker's Φ (Tucker, 1951) was extracted as a measure of similarity between the perfect matrix and the loading matrix of each country. Subsequently, the obtained congruence coefficients were correlated with Minkov's (2017, 2018) individualism-collectivism and monumentalism-flexibility axis as well as two indicators of tightness-looseness (Gelfand et al., 2011; Uz, 2015) to investigate whether the fit of the FFMQ to the idealized structure differs systematically along these cultural dimensions. Data was obtained on both individualism-collectivism and monumentalism-flexibility for all countries from Minkov et al. (2017, 2018), except for Croatia. Tightness scores were obtained from Gelfand et al. (2011) which had data for all countries except four (Chile, Croatia, Romania, and Sweden). Looseness scores were obtained from Uz (Uz, 2015), which had data on all countries except five (Australia, China, Hong Kong, Norway, and New Zealand)

## Results

Following the prior outlined analysis plan, first the fit of the individual models for each individual sample was tested to determine the best fitting structure for the FFMQ. The FFMQ model with correlated facets (Fig. 1, model 1) converged successfully in all samples, but only showed good fit in the Portuguese and New Zealand samples, which together represented 8.70% of all samples (all results are reported in Appendix D, Table 2). This indicates that the correlated five-facet model did not represent the underlying structure of the data in most samples.

The FFMQ model with five-facets subsumed under a higher-order factor (Fig. 1, model 2) showed good fit in two samples (8.70% of the samples), insufficient fit in 20 samples (86.96%), and failed to converge in one sample (4.35% of the samples, all results and comparison of fit with model 1 are reported in Appendix D, model fit: Table 3, comparison Table 8). The change in model fit was examined based on the pre-registered criteria between model 1 and model 2 and which indicated that it did not improve the fit in any of the samples where the models converged (22 out of the 23 samples; 1 sample failed to converge). According to Vuong's test of non-nested models, model 1 showed better fit for 20 out of 23 samples. For two samples, no preferable model could be determined and one sample failed to converge. Overall, this indicates that the FFMQ model with a higher-order factor does not empirically fit the data better compared to the model with correlated facets.

A possibility for the low fit of the FFMQ with a higher order factor could be the presence of positive and negative method factors identified in previous research (e.g., Aguado et al. 2015; Van Dam et al. 2012). The model with correlated facets and uncorrelated method factors (Fig. 1, model 3a) converged successfully in all samples, and showed good fit in 15 samples, which together represented 62.50% of all samples (all results are reported in Appendix D, Table 4). The change in model fit was examined based on the preregistered criteria between model 2 and model 3a and which indicated that it improved the fit in all samples where the models converged (22 out of the 23 samples; 1 sample failed to converge). According to Vuong's test of non-nested models, model 3a showed better fit for all samples that converged. Therefore, the fit of the FFMQ with one higher-order factor and uncorrelated positive/negative method factors was examined next (Fig. 1, model 3b). This model showed acceptable fit in 12 samples (52.17% of the samples), insufficient fit in eight samples (34.78% of the samples) and failed to converge in three samples (13.04% of the samples). The full results are reported in Appendix D (Table 5).

Model 3b showed no difference from model 3a based on CFI, or $\hat{\gamma}$, but was favored by the BIC. This was supported by Vuong's test of non-nested models which indicated that for 15 samples (62.50%) no preferred model could be found. We therefore additionally tested the fit of model 3b against the prior models to determine whether the fit increased. Because model 2 (facets subsumed under higher-order factor) did not show increased fit compared to model 1 (correlated facets), the fit of model 3b (FFMQ with a higher-order factor and positive/negative methods factors) was first compared against model 1 (correlated facets). Model 3b showed substantially higher CFI, $\hat{\gamma}$, and a substantial reduction in BIC for 20 samples (86.96% of all samples), but three samples failed to converge (13.04% of all samples). The comparison of model 3b against model 2 yielded similar results, indicating improved fit of the FFMQ with uncorrelated method factors. This result was further supported by Vuong's test of non-nested models which indicated better fit of model 3b for 20 samples (86.96% of all samples) while three samples failed to converge (13.04% of all samples). This indicates that positive and negative wording method factors were present in most samples and should be modeled.

While the previous finding indicated that the introduction of method factors substantially improves the fit of the FFMQ, it was unclear whether these method factors should be correlated or uncorrelated with each other. Therefore, model 4a allowed the method factors to be correlated with each other. This model showed acceptable fit in 11 samples (47.83% of samples), insufficient fit in six samples (26.09% of samples) and failed to converge in 6 samples (26.09% of the samples). The full results are reported in Appendix D (Table 6). The model with correlated method factors showed improved fit compared to the model with uncorrelated method factors in one sample (4.35% of all samples), no improvement in fit in 16 samples (69.57% of all samples), and six samples did not converge (26.08% of all samples). Vuong's test of non-nested model yielded similar results with model 4a fitting better in two samples (8.70% of all samples), no clear preference between models for 15 samples (65.22% of all samples), and 6 samples (26.09% of all samples) did not converge. Overall, this indicates that the model with correlated method factors does not fit better than the model with uncorrelated method factors in most samples. Furthermore, the uncorrelated method factor model was preferable as it was conceptually simpler and the most parsimonious model.

Last, the model in which the method factors were not only allowed to correlate with each other, but also with the higher-order factor of mindfulness (model 4b), showed acceptable fit in 11 samples (47.83% of samples), insufficient fit in four samples (17.39% of samples), and failed to converge in eight samples (34.78% of samples). The full results are reported in Appendix D (Table 7). In summary, the FFMQ models with uncorrelated positive and negative method factors showed the best fit in the individual samples. Because model 3a (correlated facets with methods factors) and model 3b (facets subsumed under a higher order factor with methods factors) could not be differentiated in most samples both models were selected to be tested for the cross-cultural equivalence of the FFMQ.

**Cross-Cultural Equivalence of the FFMQ**

In the next step, the cross-cultural equivalence of the five-facet model with higher-order factor and uncorrelated method factors was examined. As specified in the pre-registration, the equivalence of the model in all samples that converged in the previous analysis and showed acceptable fit in the individual analysis of fit was examined (the analysis was also run for all countries that successfully converged and showed an identical result. The results are available on the OSF page of this project. If multiple samples per country were available, these were merged to obtain an overall dataset for each country. One sample from the US was excluded, because the model did not converge, and one sample from Norway had to be excluded due to bad fit. No data from Australia, India, Hong Kong, China, Poland, Romania, or Chile were included due to all samples either having bad fit or the model not converging. This left data from only Western countries: Portugal, New Zealand, Germany, USA, Austria, Croatia, Spain, Sweden, and Norway.

Initially, an unconstrained model was fitted to formally test for structural equivalence. This model showed good fit ($\chi2$ (5922) = 10,097.470, $\chi2$ /df = 1.705, CFI = .943, RMSEA = .038[.036, .039], SRMR = .056, BIC = 526,262.200, $\hat{\gamma}$ = .962), indicating that the model was structurally equivalent across cultures. This was the baseline model for the further comparisons. To test metric equivalence, the same model was fitted across cultures, but with all factor loadings on the substantive factors constrained to be equal across cultures (loadings on the method factors were allowed to vary freely, see Van Dam et al. 2012). While the model showed acceptable fit by itself ($\chi2$ (6274) = 11,250.640, $\chi2$ / df = 1.793, CFI = .932, RMSEA = .040[.039, .041], SRMR = .071, BIC = 524,536.000, $\hat{\gamma}$ = .954), it nevertheless showed a substantial drop from the unconstrained model ($\Delta$CFI = − .011, $\Delta\hat{\gamma}$ = − .008) across all countries, indicating that the FFMQ was not metrically equivalent across cultures. We repeated the analysis for model 3a for all samples that showed good individual fit and found identical results. While the model showed acceptable fit by itself ($\chi2$ (7183) = 11927.759, $\chi2$ /df = 1.661, CFI = .943, RMSEA = .038[.036, .039], SRMR = .054, BIC = 602826.531,bγ = .962), it nevertheless showed a substantial drop from the unconstrained model ($\Delta$CFI = − .010, $\Delta\hat{\gamma}$ = − .007). Metric equivalence of a model in which all paths including the method factors were constrained was also tested. This analysis yielded an identical result. The results are available on the OSF page of this project. Based on these results, no further test for scalar equivalence was conducted since the data already failed metric equivalence tests.

**Exploratory Analyses**

**Equivalence of the individual facets of the FFMQ.** The pre-registration specified that in case of poor equivalence of the overall FFMQ, the equivalence of the individual facets would be tested. First, CFAs were run for the separate facets in each sample to determine which samples should be included in the equivalence analysis (due to space constraints the fit for all samples and all facets is reported in Appendix D, Table 8). All samples that showed adequate fit across CFI, RMSEA, SRMS, and $\hat{\gamma}$ were included. If multiple samples in a country showed good fit those were subsequently merged, and the equivalence analysis run across countries.

***Acting with Awareness.*** Acting with Awareness did not show acceptable fit in any of the samples, indicating that a uni-dimensional structure of Acting with Awareness might not be the best fit in most samples.

***Observing*****.** Observing showed a good fit in 82.61% of all samples, indicating that the uni-dimensional structure of Observing fits well in most samples. Because the structure of the observing facet did not fit well in the individual CFA, Australia and Poland were excluded from the equivalence analysis. Across the remaining countries the model showed good structural equivalence ($\chi^2(240)$ = 567.829, $\chi^2/df$ = 2.366, CFI = .963, RMSEA = .056 [.050, .062], SRMR = .035, BIC = 144529.300, $\hat{\gamma}$ = .987). Nevertheless, when tested for metric equivalence the model showed a substantial drop in fit ($\Delta$CFI = -.016, $\Delta\hat{\gamma}$ = -.006), indicating that the Observing facet was not metrically equivalent across the samples studied here. Overall, this indicates that while the Observing items measure a single construct in most countries, the individual items were not equally good indicators in each country.

***Non-Judging.*** The Non-Judging facet showed good fit of the structure in 43.48% of all samples, indicating that the uni-dimensional structure of Non-Judging did not fit well in the majority of samples. For the equivalence analysis Australia, Austria, Chile, Hong Kong, Spain the USA, and Germany were excluded because no sample from these countries showed good fit. Across countries the model showed good structural equivalence ($\chi^2(180)$ = 359.322, $\chi^2/df$ = 1.996, CFI = .976, RMSEA = .061 [.051, .070], SRMR = .033, BIC = 70728.890, $\hat{\gamma}$ = .986). Nevertheless, when tested for metric equivalence a significant drop in model fit ($\Delta$CFI = -.017, $\Delta\hat{\gamma}$ = -.011) was found, implying that the factor loadings were not identical.

***Describing.*** The Describing facet showed good fit of the structure only in 4.35% of all samples. The only sample where the Describing facet showed good fit was Austria. This excluded any test for measurement equivalence.

***Non-Reacting.*** The Non-Reacting facet showed good fit of the structure in 52.17% of all samples, indicating that the uni-dimensional structure of Non-Reacting fits well in the majority of samples. Samples from Germany, Austria, Croatia, Chile, China, Poland, Hong Kong, and Spain were excluded from the further equivalence analysis because the samples did not show acceptable fit in the individual analyses. Across the remaining countries, the model showed good structural equivalence ($\chi^2(112) = 221.483$, $\chi^2/df = 1.978$, CFI = .968, RMSEA = .058 [.047, .069], SRMR = .035, BIC = 58255.100, $\hat{\gamma}$ = .990). Nevertheless, when tested for metric equivalence a significant drop in model fit ($\Delta$CFI = -.013, $\Delta\hat{\gamma}$ = -.005) was found indicating that the Non-Reacting facet was not equivalent across cultures.

To summarize the previous analysis, showed that no single facet of the FFMQ exhibits metric equivalence across all available countries. Further, both Acting with Awareness and Describing did not show good CFA fit when investigated separately from the overall structure of the FFMQ, suggesting that these facets might not be uni-dimensional. Overall, this analysis parallels the finding on the overall structure of the FFMQ. This indicates that neither the FFMQ as a whole nor the individual facets are sufficiently cross-culturally equivalent to allow for cross-cultural comparison of means or even correlations with other constructs.

**Alternative Structure of the FFMQ.** While the FFMQ model with a higher-order factor of mindfulness and uncorrelated methods factors showed good fit in most cultures, a number of cultures, mostly non-Western, still showed below acceptable fit. Overall, both on the level of the total FFMQ and the individual facets, no metric equivalence was found, which indicates that individual items do not load in the same way on the underlying constructs across cultures.

Therefore, an exploratory analysis was conducted to examine the possibility of an alternative structure of the FFMQ across cultures. A sample-size weighted average correlation matrix of the FFMQ items across all cultures was computed and the ideal number of components to be extracted from the correlation matrix calculated using parallel analysis (Dinno, 2018; Horn, 1965). The parallel analysis indicated that six components should be extracted (adjusted Eigenvalues: 7.288, 4.587, 2.609, 2.172, 1.849, 1.042).

As specified in the pre-registration, two separate PCAs were run, once allowing for correlated components using an oblimin rotation and one forcing components to be orthogonal using a varimax rotation. The results for the varimax rotation are reported in Appendix D (Table 9), the results of the oblimin rotation are reported on the OSF page of this project, as they were nearly identical. The overall factor structure in the combined sample suggested that four of the five facets emerged, but that the Acting with Awareness items loaded on two separate factors. One factor was defined by the Acting with Awareness items focusing on behavior, whereas the second factor was defined by Presence items. The results of the cross-cultural PCA indicate that a six-factor structure might fight better across cultures compared to the five-factor structure.

**Sources of Incongruence in the Structure in the FFMQ**

To test the possibility that the previously proposed FFMQ structure was systematically linked to culture-level variables such as individualism-collectivism and monumentalism-flexibility the five-factor solution of each country was rotated towards an idealized loading matrix. The average Φ ranged from .850 to .954 for the individual countries (all results are in Table 10 in Appendix D), where .90 can be considered good fit (Fischer & Fontaine, 2010). Overall 11 countries (Australia, Austria, Chile, China, Germany, Spain, Norway, New Zealand, Portugal, Sweden, USA; 68.75% of all countries) showed good congruence to the ideal structure. To test whether the level of congruence with the ideal matrix can be predicted using country-level cultural information, the average Tucker's Φ in each country was correlated with individualism, monumentalism and tightness vs looseness scores. Individualism ($r$ = .77 [.44, .92], $p$ <.001), but not flexibility ($r$ = -.18 [-.63, .37], $p$ = .52) was significantly related to greater congruence with the proposed structure. Individualism predicted

60.02% of the variance in average congruence. Further, average congruence was significantly related to looseness (*r* = .77 [.32, .94], *p* <.05) measured with the indicator by Uz (2015), indicating better fit in looser cultures and explaining 59.58% of variance in average congruence. When measured with the tightness indicator by Gelfand et al. (2011) the relationship was significant (*r* = -.59 [-.87, -.03], *p* < .05) and in the same direction indicating greater fit in looser cultures explaining 28.88% of the variance.

Finally, the fit of the individual countries to the pooled solution was examined by rotating the six-factor solution in each country towards the pooled loading matrix. The average Φ ranged from .883 to .980 for the individual countries. Overall, only two countries (Hong Kong and India; 12.50% of all countries) showed below acceptable congruence to the pooled structure (all results can be found in Appendix D, Table 11).

To test whether the level of congruence with the ideal matrix can be predicted using country-level cultural information the average Tucker's Φ across all dimensions in each country was correlated with individualism, monumentalism and tightness-looseness. Individualism (*r* = .58 [.39, .91], *p* <.001), but not flexibility (*r* = -.23[-.66, .32], *p* = .42) was significantly related to greater congruence with the proposed structure. Individualism predicted 56.45% of the variance in average congruence towards the pooled structure. In regard to tightness-looseness, the average congruence was not significantly related to looseness (*r* = .35 [-.31, .79], *p* = .29; data by Uz, 2015). For the tightness indicator by Gelfand et al. (2011) the relationship was in the same direction but not significant (*r* = -.38 [-.78, .25], *p* = .22). Overall, this indicates that the five-factor structure of the FFMQ was a better representation of the underlying structure of mindfulness in more individualistic, loose Western rather than more collectivistic, tight non-Western countries. While the six-factor structure of the FFMQ shows no bias based on tightness-looseness, it was still biased in favor of individualistic cultures.

**Discussion**

The current study used a large multi-national data set to provide a systematic analysis of the cross-cultural equivalence of the FFMQ across a wide range of cultures. The main findings were a) the FFMQ structure did not adequately fit across cultures, even when including separate method factors, b) the Acting with Awareness facet broke apart into a behavioral and a presence factor in an exploratory analysis and c) the ideal structure of the FFMQ might be driven by cultural values.

**Implications for the Modelling of the FFMQ**

The FFMQ showed substantially better fit in most countries if it was modelled with positive and negative item-wording factors. These finding support previous research (e.g., Aguado et al., 2015; Van Dam et al., 2012) on the presence of item-wording factors in the FFMQ. These item-wording effects might be more substantial for non-meditators (Van Dam et al., 2009); for an alternative explanation of these findings see: Baer, Samuel, & Lykins, 2011). Overall, this indicates that the inclusion of item-wording factors substantially improved the fit compared to the model proposed by Baer et al. (2006) in the majority of samples. The current results also suggest that these item-wording factors are most likely orthogonal. The presence of item wording factors echoes concerns in the literature about scales, such as the MAAS, that measure mindfulness with only negatively scored items (Grossman, 2011). The use of scales that are not balanced for wording might conflate response tendencies to negatively worded items with substantial variance in mindfulness (for an example of potential variables influencing responses to positive and negative items see: Michaelides et al., 2016).

Across cultures the FFMQ items were best represented as a six-factor structure, with Acting with Awareness divided into awareness of thoughts and awareness of actions. This differentiation of the Acting with Awareness facet suggests that two different processes might underlie this factor and that it should not be treated as a uni-dimensional construct across cultures. The two-factor structure that emerged resembles the distinction made by researchers of consciousness about private cognitive spaces, in other words awareness of the external world and one's behavior in it, and public cognitive spaces, in other words awareness of internal world, e.g. thoughts and images (Gray, 2004). The first sub-factor of Acting with Awareness was characterized by items indicating awareness of one's behavior, aligning with public cognitive spaces. The second sub-factor of Acting with Awareness was characterized by items specific to one's mental processes, aligning with private cognitive spaces. Previous research on the effect of body focused meditation showed that this meditation practice can impair metacognitive efficiency (Schmidt et al., 2019). Using a two-factor structure of Acting with Awareness separating thought from action awareness might provide further insight into the relationship of mindfulness, body-awareness, and meta-cognition.

**Implications for Cross-Cultural Comparisons**

The second aim of the current study was to test the cross-cultural validity of the FFMQ. The cross-cultural equivalence of the FFMQ with uncorrelated methods-factors was examined and the results indicated that this model shows good structural equivalence across the different countries in which the individual CFAs showed good fit. The items were related to the proposed theoretical facets (e.g., showed non-trivial loadings). Nevertheless, no support for metric equivalence was found. This indicates that the items of the FFMQ are not equally good indicators of the individual facets across countries. This non-equivalence precludes both comparisons of correlations between the FFMQ and other variables of interest across countries as well as direct or indirect (profile) mean comparisons between cultures. The FFMQ in its current form is not a suitable tool to assess mindfulness in a cross-cultural context. This does not preclude the use of the FFMQ in mono-cultural studies but highlights the need for a cross-culturally valid measure of mindfulness. Initial steps should start with explicit considerations whether mindfulness is an emic (culture specific) or etic (universal) concept (Berry, 1989; Farh et al., 2006). Current research practice treats mindfulness as de-facto etic construct with scales largely developed in a western context and subsequently translated or adapted into other languages.

One potential reason for this de-facto etic approach to mindfulness in the West is the tendency of Westerners to consider themselves to be essentially culturally neutral, meaning they think of themselves as not introducing a cultural bias into a psychological concept (Bellah, 2007). However, any psychological tests are potentially shaped by the cultural environment in which they were first proposed. The case of mindfulness research shows that this assumption of cultural neutrality is not warranted and instead the concept became more individualistic, focused on personal freedom, and authenticity during the move of mindfulness practice from Asian contexts to North America (Purser & Milillo, 2015). This individualization of mindfulness is not only reflected in theory, but also in the measurement of mindfulness. The current study explored whether the fit of the proposed structure of the FFMQ was systematically linked to previously identified cultural dimensions. The results indicated that cultures higher on individualism and looseness showed better congruence to the proposed structure of the FFMQ. While the effects of monumentalism-flexibility were not significant they still showed an effect in the same direction indicating that more flexible cultures had worse fit to the overall structure. The findings suggest that the FFMQ may capture conceptualizations of mindfulness prevalent in Western and individualistic cultures compared to understandings of mindfulness in more collectivistic cultures, including some of the more collectivistic settings from which the concept originated. This is of concern since it indicates the presence of a systematic Western individualistic bias in the current FFMQ and highlights that to produce a cross-culturally valid measure of mindfulness a translation or adaptation of currently used mindfulness measures might not be sufficient. More conceptual work is needed to adequately understand mindfulness across cultural contexts.

One promising approach for developing a more valid cross-cultural measure can be found in the development of the internationally validated positive and negative affective schedule (Thompson, 2007), which used a mixture of qualitative and quantitative data from a wide range of cultures to determine items and factor structures of the new measure. Overall, to advance cross-cultural research on mindfulness, new measures should be developed utilizing an approach which includes diverse cultural perspectives to minimize the cultural bias of the measure.

## Limitations and Future Research

The strength of the current study lies in the wide range of cultures captured in data set. The samples cover all permanently inhabited continents besides Africa which allows for the examination of the equivalence of the FFMQ from a broad perspective. The major limitation of the current study is the reliance on previously published data on the FFMQ rather than on representative samples from each country. Further, the analysis did not control for meditative experience of the participants, barring comparisons between meditators and non-meditators in different cultures.

Coming back to the initial questions, whether the FFMQ is a valid tool of measurement across cultures, results indicated general problems with the cross-cultural comparability even though the individual samples often showed acceptable fit when considered individually. Importantly, the exploratory analysis suggests that mindfulness as a construct might be biased towards individualistic Western interpretations of the construct. Overall, the FFMQ and the conceptualization of mindfulness in terms of five facets subsumed under a single overall construct might not be suitable for cross-cultural comparisons. To further develop the field of mindfulness research, both a closer exploration of the theoretical structure and cross-culturally valid measurement tools are necessary. Future research could collect data from a wide range of cultures on emic perspectives on mindfulness to aid the creation of a cross-culturally valid measure of mindfulness.

**Data availability**: All raw data, the analytic code, and all materials associated with the study are available on the Open Science Framework (https://osf.io/nftxb/?view_only=2559ee6a8a9c461a9e4b4c838c1230af).

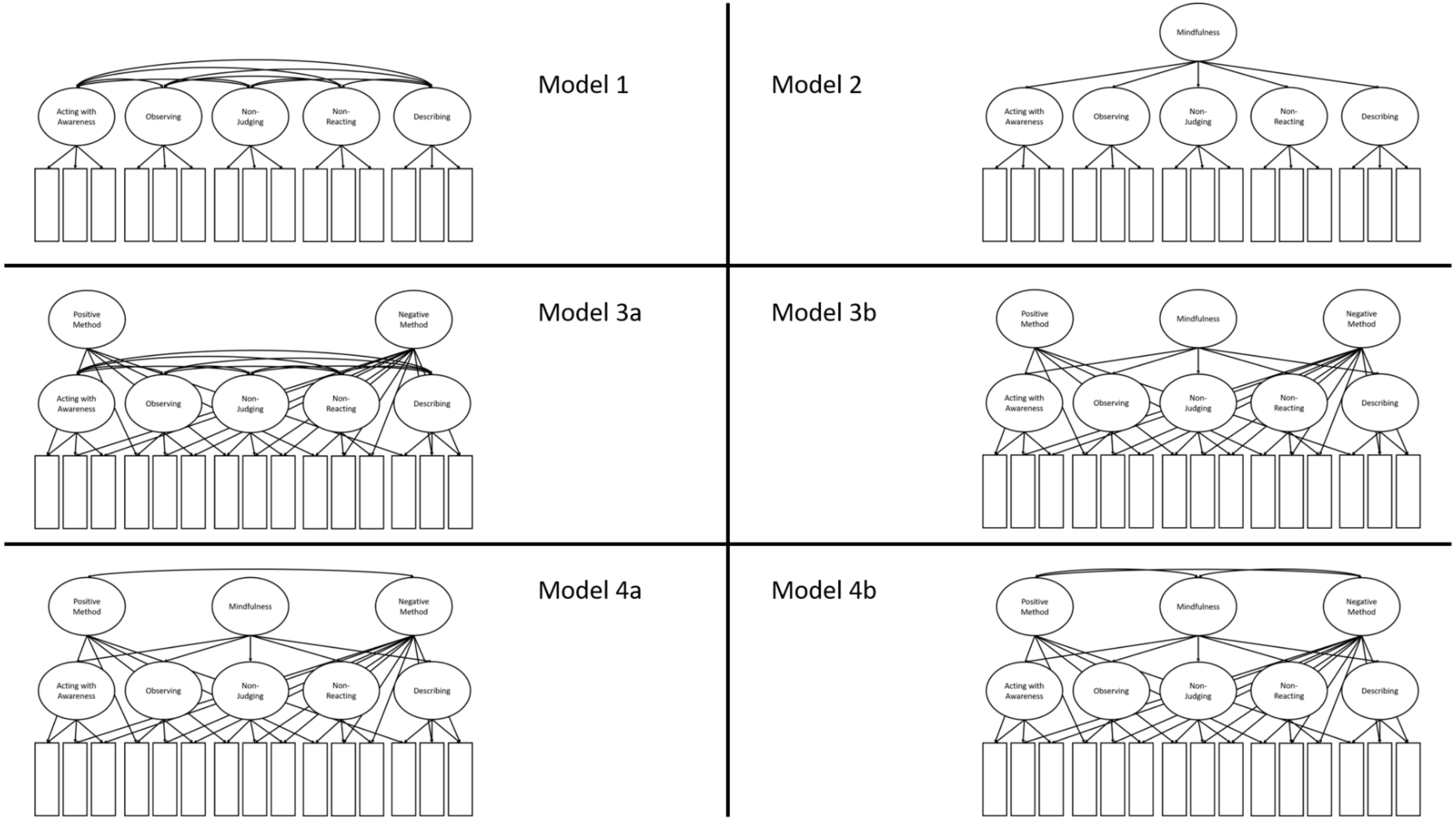



*Figure 7.1 Models tested in each country.*

# General Discussion[11]

This thesis was set against the backdrop of a growing body of trait mindfulness research, growing both in absolute research volume and in relative importance within the larger mindfulness field. Definitions of mindfulness have been shown high diversity (Chiesa, 2013). In this thesis I defined mindfulness as: “paying attention in a particular way; on purpose, in the present moment, and nonjudgmentally” (Kabat-Zinn, 1994, p. 4) due to: (i) the widespread adoption of this definition; and (ii) and it’s elements being reflected in the trait measures used in this thesis (Baer et al., 2006). Importantly, I conceptualized mindfulness traits as descriptors of density distributions of mindfulness states (for similar perspectives see: Mesmer-Magnus et al., 2017). This approach is reflective of the wider field of trait mindfulness research in which “[trait] mindfulness is viewed as the general tendency of a person to show characteristics of nonjudgmental awareness of present-moment experience in their everyday life” (Krägeloh, 2020, p. 64)

In Chapter 1, I investigated the growth and change of the trait mindfulness field. In line with the wider field of psychology (Henrich, 2020) I found a strong bias of the literature towards Western research institutions and topics of importance such as substance abuse. This imbalance raises concerns for two reasons. First, most mindfulness research is using measures derived from originally Buddhist conceptualizations and the current lack of research from East Asian with Buddhist influence beyond China might result in an increased cultural specificity of the construct towards Western conceptualizations. Second, the absence of systematic research on cultural and cross-cultural perspectives on mindfulness might result in what has been termed “imposed-etic” (Berry, 1989). While emic views represent the views of the insider of a culture on a specific cultural system, the etic view represents a cultural outsiders view on the same system (Berry, 1999). In the case of

[11] Parts of this section have been published elsewhere: Karl, J. A., Johnson, F. N., Luisa, B., & Fischer, R. (2021). In search of mindfulness: A review and reconsideration of cultural dynamics from a cognitive perspective. Journal of the Royal Society of New Zealand. https://doi.org/10.1080/03036758.2021.1915804

mindfulness, the emic view would be investigations conducted using a framework guided by Buddhist cultural knowledge, whereas the etic view represents the investigations starting from non-Buddhist cultural frameworks. In an imposed-etic approach conceptual models derived in one-cultural context are treated as de-facto universals and are imposed on different cultures, for example by imposing Western interpretations of mindfulness back onto cultures for which mindfulness has specific cultural and spiritual significance (for an exploration of this see: Feng et al., 2018). It is essential to ensure that the models used are cross-culturally valid and appropriate. Failing to do so limits researcher abilities to produce reliable theories that merge macro and micro level perspectives on mindfulness, and also limit the scope of further research.

In Chapter 1, I also found that a strong preference of the field for two measures of mindfulness, the mindful attention and awareness scale (MAAS: Brown & Ryan, 2003) and the five-facet mindfulness questionnaire (FFMQ: Baer et al., 2006). As can be seen in Figure 8.1 the FFMQ has been steadily growing in use, paralleling the MAAS (Figure 8.1 A). Interestingly, the relative usage of the FFMQ and the MAAS has stabilized around 2011, with yearly citations of the FFMQ now making up around 40% of the combined citations of the measures (Figure 8.1 B). The continued popularity of the MAAS might be a valuable topic for future meta-research on mindfulness as it is likely that specific sub-fields of trait mindfulness prefer it due to the simpler factor structure and the relative shortness of the full MAAS (15 Items, 1 Factor) compared to the full FFMQ (39 Items, 5 Factors). Due to its relative popularity, broader conceptualization allowing more fine-grained analyses of individual components, and the fact that the FFMQ was built bottom-up including questions derived from the MAAS, I decided to utilize FFMQ for the current thesis.

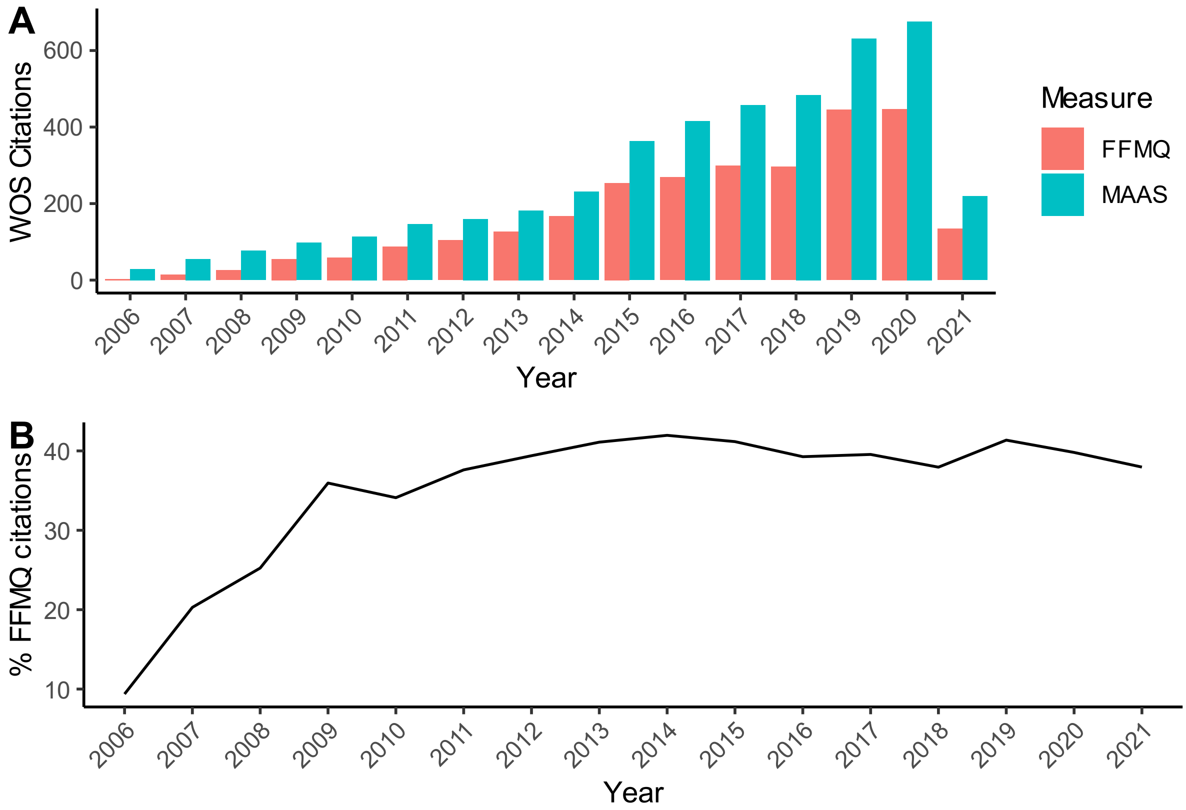



*Figure 8.1 Change in the use of the FFMQ and the MAAS*

To increase confidence in this choice of measure, I examined the conceptual replicability of the FFMQ in Study 2. The main goal of this study was to validate the five-facet structure even if measures were included that were developed after the initial publication of the FFMQ. Further, I aimed to establish the empirical distinguishability of the FFMQ facets from recent measures aimed at capturing non-Buddhist inspired mindfulness, such as the Langer mindfulness scale (Pirson et al., 2018). Overall, I found confirmation of the five-factor structure, but the presence of negative and positive methods factors foreshadowed similar results in a cross-cultural analysis reported in Study 5. The findings also supported the recent move away from treating mindfulness as a unitary concept (e.g., using a total score of mindfulness) towards using the mindfulness facets as individual constructs (Aguado et al., 2015; this recommendation was also made by: Baer et al., 2006; Park et al., 2013). In line with this multicomponent view of mindfulness, I operationalized mindfulness using the individual facets of the FFMQ in Study 3 and 4, to obtain more nuanced insights into mindfulness facets and their links with established individual difference measures. Additionally, I found that

Buddhist inspired measures of mindfulness were empirically separable from the explicitly non-Buddhist inspired Langer mindfulness, supporting the separation of these lines of research on mindfulness.

Taken together, the first set of studies in my thesis highlighted two salient gaps in the trait mindfulness literature. First while individual aspects of the FFMQ can be reliably recovered, less is currently known about how these facets of trait mindfulness facets fit in the larger body of individual differences within psychology in general and personality psychology in particular. Second, while the developmental history of mindfulness can be considered trans-cultural, little is known about cultural dimensions in current mindfulness measurement. In my thesis I attempted to address these fundamental questions.

Moving forward from the first block of studies, I attempted to address the first identified shortcoming of the trait mindfulness literature and investigated potential predictors or individual difference correlates of mindfulness. In Study 3 and 4 I attempted to bring together two perspectives on individual differences that might underly mindfulness, focusing on personality in the big five tradition (McCrae & Costa, 1997; Soto & John, 2017) and reinforcement sensitivity theory (Corr & Cooper, 2016; Gray, 1970). It is important to acknowledge at this point that I do not propose that they are the sole predictors of mindfulness and others have shown promising results using other theoretical measures or by focusing on developmental pathways including attachment (Pepping et al., 2013, 2014; Pepping & Duvenage, 2016). Nevertheless, personality has received the widest interest in relation to mindfulness (Giluk, 2009).

Two major approaches to personality have been investigated in relation to mindfulness: personality as captured by the Big-Five (Giluk, 2009; Haliwa et al., 2020; Hanley et al., 2018; Hanley & Garland, 2017), and reinforcement sensitivity theory (RST; Dolatyar & Walker, 2020; Hamill et al., 2015; Harnett et al., 2016; Reese et al., 2015; Sauer et al., 2011). While these previous studies have investigated the relationship of mindfulness and either the Big-Five or mindfulness and RST, research should include both concepts as they are substantially linked, both conceptually and empirically. Gray (1970, 1981) proposed the Reinforcement sensitivity theory in response to Eysenck's earlier work on Neuroticism and Extraversion to explain some of the neuro-biological underpinnings of personality. Corr and McNaughton (2012) revised the theory by more clearly differentiating the behavioral approach system from the behavioral inhibition system. To recapitulate some of the major components of relevance for this thesis, the behavioral inhibition system (BIS) regulates motivational conflicts. In contrast the behavioral activation system (BAS) regulates the behavioral approach to stimuli. The BAS consists of several subsystems that facilitate goal attainment along the temporo-spatial continuum. These subsystems range from goal identification defined by anticipation and interest (Reward Interest), to goal obtainment via fast action (Impulsivity) or planning and persistence (Goal Drive Persistence) all the way to emotional excitement and behavioral reinforcement when goals are completing or near completion (Reward Reactivity).

The five traits of the Big-Five (Extraversion, Agreeableness, Openness, Conscientiousness and Neuroticism/Emotional Stability ) can be partially map onto the neurobiologically derived RST systems: Neuroticism is strongly correlated with BIS; Extraversion is conceptually related to all BAS subsystems, Conscientiousness is related to BAS Goal-Drive Persistence, but also weakly negatively to BIS (it requires both goal persistence and monitoring goal conflicts), Openness is correlated with BAS Reward Interest (since it is anticipatory and involves reward simulation) and Agreeableness typically shows weak but positive correlations with all BAS components, except BAS Impulsivity (for examples see: Antoniazzi & Klein, 2019; Corr & Cooper, 2016; Pugnaghi et al., 2018). Importantly, RST is thought to capture the expression of an underlying neuro-biological system, whereas the Big-Five are assumed to capture the output of higher-level cognitive processes (Smits & Boeck, 2006).

To elaborate on these links and ground the discussion of the links with mindfulness in the next section, in a recent study I showed together with my supervisor that at a behavioral level items measuring the RST and Big-Five are strongly connected but are not interchangeable (Fischer & Karl, 2020). We used responses of 749 participants and constructed an item-level network which included the Big-Five and RST in addition to personal values. Using recently developed network approaches, we extracted the partial correlations between the individual self-reported items (Figure 8.2 A) and used a regularization approach to remove spurious links (Epskamp et al., 2018; Epskamp & Fried, 2018). We subsequently used Exploratory Graph Analysis (Golino et al., 2020; Golino & Epskamp, 2017), a recently developed powerful community detection approach, to determine whether the conceptually proposed dimensions could be collapsed empirically, in case that they share substantial overlap (Figure 8.2 B). We found substantial relationships between individual aspects of the Big-Five and RST. For example, BIS and Neuroticism were highly related (see the highlighted sections in Figure 8.2 A and B) and in this larger order network could be expressed as one component. Importantly, these two components were similarly strongly related in cross-sectional studies (for example r=.73 in the original study describing the development of the RST-PQ by Corr & Cooper, 2016). In contrast, concepts such behavioral activation and Extraversion that are thought to be similarly related formed clearly distinguishable item communities in our study. Overall, this indicates that Neuroticism and BIS are substantially connected potentially due to cognitive differences in Neuroticism building on biological differences in potential conflict monitoring and threat sensitivity. Therefore, in order to understand the unique relationships of mindfulness with personality-like individual differences, both the big five and RST should be included in order to gain a better understanding of the presumed biological and cognitive correlates of mindfulness. This motivated my approach in Study 3 and 4, in which I examined the relationship between these constructs both cross-sectionally and longitudinally.

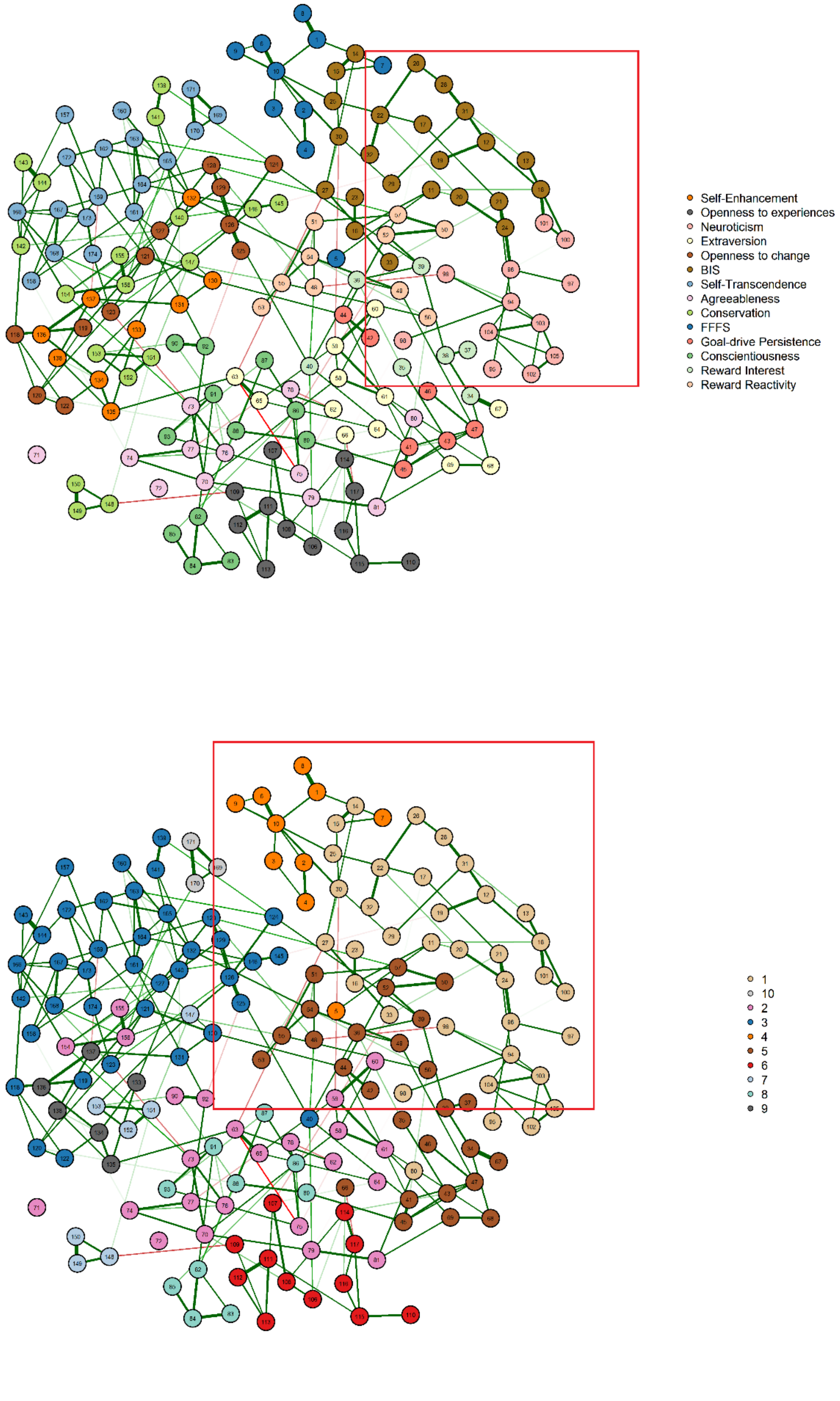



*Figure 8.2 Network of personality, reinforcement sensitivity, and values.*

Figure A, represents the estimated network with colours indicating the original instruments and dimensions, Figure B represents the same network with nodes colored by their grouping after an exploratory graph analysis. (adapted from Fischer & Karl, 2020). As can be seen in the highlighted area, the BIS and Neuroticism items formed a single cluster in the exploratory graph analysis.

The combined perspectives allowed me to provide a more nuanced perspective on the relationship between RST and Big-Five with mindfulness. For example, previous research has indicated that biologically based behavioral inhibition as measured with the RST-R is negatively related to both Non-Reacting and Non-Judging (Dolatyar & Walker, 2020). This would imply that both the ability to stay non-reactive to negative experiences and the ability to be non-judgmental towards the self might be diminished in individuals that have a predisposition to carefully monitor conflicts and are more sensitive to threat overall. In Study 3 using the same measure as Dolatyar and Walker (2020), I found that behavioral inhibition is negatively related to Non-Judging, but not Non-Reacting when also including and controlling for the Big Five personality traits. This implies that neuro-biological differences in threat sensitivity might reduce individuals' ability to be non-judgmental, but this does not seem to influence their ability to regulate their behavioral reactions (Non-reacting). In contrast, I found that Neuroticism is negatively related to Non-Reacting, indicating that higher levels of anxiety, depression, and general emotional volatility might make individuals react more easily to negative experiences. This pattern highlights the importance of investigating the relationship of mindfulness with individual differences simultaneous at different levels because it allows for the formulation and testing of possible mechanistic pathways. Behavioral inhibition has been found to be substantially related to social anxiety (Panayiotou et al., 2014; Ran et al., 2018). High BIS individuals might perceive accidental social transgressions as more threatening to their social standing, and therefore increasingly monitor and judge their own thoughts and behaviors. In contrast, individuals with higher levels of Neuroticism have already elevated levels of anxiety and depression (Roelofs et al., 2008; Yoon et al., 2013) and lower resilience (Lü et al., 2014), which may now more strongly impact their ability to control their emotions (leading to negative associations with Non-reacting but not with Non-judging, controlling for BIS). Overall, this highlights that research on possible dynamics of mindfulness can be advanced by focusing on the facets of mindfulness, especially if the research is embedded more strongly with established individual difference dynamics.

Study 3 relied on cross-sectional data and no causal relationships can be inferred from these patterns. Continuing to push this line of inquiry, in Study 4 I investigated the same relationships longitudinally. To the best of my knowledge no previous study has investigated the longitudinal relationship between RST and mindfulness or the combined effect of RST and the Big-Five on mindfulness over time. To provide a first investigation of potential longitudinal relationships over the span of a few months, I used a random intercept cross-lagged panel model (RICLPM). This model presents an advancement above commonly employed longitudinal models such as the cross-lagged panel model (CLPM) by explicitly modelling stable between-participant differences and separating them from within-participant changes. This statistical separation via the inclusion of a person-specific random intercept allows for an unbiased estimation of within-participant changes over time (Hamaker et al., 2015). To demonstrate that this makes a difference for our understanding of mindfulness, I report a few selected associations in Table 8.1. As can be seen the two models (cross-sectional data analysis and CLPM) that only model between subject relationships and do not clearly separate within and between relationships show similar relationships. Therefore, traditional cross-lagged panel models may converge with cross-sectional study findings. In contrast, the RICLPM indicated a substantially weaker effect of Neuroticism, and did not indicate significant paths from either BIS or Neuroticism to mindfulness. Furthermore, I found the reverse effect for Non-Reacting to BIS, indicating that participants with an increased ability to confront negative experiences without immediately reacting (mindfulness facet) showed reduced expression of threat sensitivity (behavioral inhibition component). It is possible that a process similar to exposure therapy is taking place; non-reactive individuals might stay in a perceived threatening experience longer, allowing their threat detection system to “recalibrate” (Krijn et al., 2004; Powers & Emmelkamp, 2008).

The within person analysis also indicated a potential mechanism for previously observed relationships between mindfulness and health behaviors (Sala et al., 2020). Specifically, I found feedback loops between awareness of ones' actions (Acting with Awareness) and the ability to sustain long term plans (BAS-Goal Drive). Therefore, the experience (and practice) of purposeful behavior within mindfulness and personality differences in pursuing long-term goals may reinforce each other over time. Taken together, in Study 3 and 4 I provided new evidence on the relationship between mindfulness and biological/cognitive individual differences, both cross-sectionally and longitudinally, addressing my first identified gap in the literature as described in study 1.

*Table 8.1 Comparison in the relationship between Non-Judging and Non-Reacting with BIS and Neuroticism across studies.*

| | Neuroticism – Non-Reacting | BIS – Non-Reacting | Neuroticism – Non-Judging | BIS – Non-Judging |
|---|---|---|---|---|
| Cross-Sectional (Study 3) | -.604*** | -.085 | -.121 | -.481*** |
| CLPM (Study 4 OSF) | -.231*** | .112 | -.102 | -.131 |
| RI-CLPM (Study 4) | -.008 | -.218 | -.042 | -.127 |

*Note. Study 4 is reported as cross-lagged panel model (CLPM) in addition to the RI-CLPM recommended by Hamaker et al. (2015), to highlight the importance of clearly separating within and between subject variation.*

Changing the focus to the second gap identified in the bibliometric review (Study 1), I had found that research foci differ substantially between Western countries and the largest non-Western producer (China). I also found that research from non-WEIRD contexts is substantially underrepresented in research on trait mindfulness. In Study 6 I compiled data on the FFMQ across 16 cultures and conducted systematic invariance testing of the measure using multi-group confirmatory factor analysis (using selected methods described in Study 5). I found that the FFMQ cross-culturally is best modelled including negative and positive methods factors (supporting a similar finding in Study 2, which indicated the presence of positive vs negative response factors). Additionally, I found that the FFMQ only showed structural equivalence across those samples in which the measure showed good individual fit. This indicates that the structure of the FFMQ was similar across these countries, but the strength of the loadings of the items differed substantially, preventing comparisons of relationship between the FFMQ and other measures across cultures. This was also true for the individual facets, especially Acting with Awareness. These facets did not satisfy structural equivalence, indicating that cultures are likely to differ in the underlying factorial structure of mindfulness. In a follow up-analysis, I found that Acting with Awareness split into awareness of actions and awareness of thoughts across cultures, indicating that monist conceptualizations of awareness, equating awareness of the mind and awareness of the body, might not be suitable across cultures. Finally, I examined the fit of the individual country-specific structures towards the idealized structure of the FFMQ, using Procrustes rotation. I found that divergence from the ideal structure was significantly related to individualism and cultural looseness, features often associated with Western cultures (Gelfand et al., 2011; Minkov et al., 2017; Uz, 2015). This provides empirical support for the notion I raised in Study 1: The Western skew in the research of mindfulness manifests itself in one of the most common mindfulness measures. Standard measures such as the FFMQ cannot be considered universal or "culture free", but rather represent a Western individualist understanding of mindfulness. This difference in conceptualizing mindfulness is especially pressing as recent studies indicate that individualistic tendencies might also alter the response to mindfulness interventions, with some studies indicating that mindfulness for individualistic participants results in greater narcissism (Poulin et al., 2021). Clearly, the cross-cultural validity and functionality of mindfulness requires greater attention.

**Limitations**

In this thesis I provided novel perspectives on trait mindfulness, both from the perspective of individual differences and from cultural perspectives. Nevertheless, similar to previous research that I criticized, one of the main limitations of this thesis is the choice of population. Study 2, 3, and 4 relied on student samples and this replicates the wider bias within psychology (for an exploration of this, see: Henrich, 2020). The original study describing the FFMQ (Baer et al., 2006) relied on student samples and my current work is constrained by similar generalizability issues. I tried to address this by providing a diverse cultural examination in Study 6. By focusing on a wider cultural perspective, I aimed to bring a diversity to mindfulness research that is often lacking. Building on this work I have recently speculated on emic cultural perspectives on mindfulness specific to New Zealand (Karl, Johnson, et al., 2021).

A second shortcoming of the current thesis is the time scale of the longitudinal study; the time period of four months does not allow me to explore potential age and cohort effects (for some initial evidence on the role of age and mindfulness see: Mahlo & Windsor, 2020). With the current sample, I was limited by the duration that participants are available for participating in this study. On a positive side, I studied a period that previously has been shown to be marked by personality changes. Nevertheless, it might be possible that developmental processes such as the ones examined in my longitudinal study might unfold differently if examined across a longer timespan.

Last, while I showed in Study 2 that overall mindfulness measures can be subsumed in the five-facet structure of the FFMQ, this does not imply that they have the same cross-cultural measurement issues that the FFMQ has. For example, the MAAS due to its simpler structure might show better cross-cultural measurement properties. If this was the case, this would imply that the facets may be experienced differently but not the overall state of mindfulness. I have strongly argued for a facet focus of mindfulness and based on the current evidence, I would not expect stronger equivalence with shorter measures such as the MAAS. Nevertheless, future research might explore the suitability of different mindfulness constructs. Importantly, this should not be a one-way street from WEIRD to non-WEIRD countries but should also include the test of measures more firmly based in Buddhist writing (Ng & Wang, 2021) actively tested and explored in non-Buddhist contexts.

**Central Take-Aways**

Finally, I want to summarize the three main insights of my thesis and contributions to the literature. First, I have shown that the five-facets of mindfulness can be largely replicated in a New Zealand sample, even while accounting for novel measures. I further showed that Buddhist-inspired measures of mindfulness can be empirically distinguished from non-Buddhist measures of mindfulness. This provides additional validation that these measures form the basis of distinct research lines and indicates the robustness of the facets of mindfulness beyond the original measures.

Second, I have shown that the links between mindfulness, RST, and the Big-Five can be refined if researchers account for the overlap between the different personality constructs. I provide the first longitudinal evidence which simultaneously involves cognitive and biological perspectives. This allows future research to start generating more specific hypothesis how mindfulness is related to cognitive and biological personality differences. I also showed that mindfulness, RST, and the Big-Five are longitudinally linked, but the changes within-individuals deviated substantially from cross-sectional results. These findings highlight the need of further research on longitudinal links between mindfulness and personality.

Last, I have provided the first large scale investigation into the cross-cultural validity of the FFMQ. This study substantially expands on previous studies by providing a global coverage of samples. I found little support of the FFMQ as a suitable instrument to reliably measure cross-cultural differences. Further, I found, in line with empirical studies of the trait mindfulness literature, that the FFMQ showed substantially better measurement properties in Western countries. While this does not imply that the FFMQ cannot be used in mono-cultural studies, it alerts researchers to take greater care when trying to compare mindfulness across cultures. This study has motivated more recent explorations into culture specific measurements of mindfulness and related constructs (S. Chen et al., 2021; Van Doren et al., 2021). Further development in the domain of cross-cultural comparability might allow researchers to design measurement instruments which allow for more refined insights into cultural factor influencing participants mindfulness.

**Personal Reflections**

To conclude, I want to take a step back and consider how my personal view on the topic of trait mindfulness has changed during this thesis. At the beginning, I was motivated by the question where differences in trait mindfulness may arise. I was confident that this would be a well-explored area.  Given the sheer body of work that is being produced on mindfulness (as can be seen in Study 1), I was surprised to find relative disconnect and scarcity of material on this question. On both aspects that interested me (individual differences and cultural differences) less information and integration was available than I had thought. In my thesis I attempted to provide empirical insights into the topics I perceived to have the most pressing need for each aspect. While for individual level predictors such as personality quality evidence was available, it is often not well integrated in the wider network of individual differences even though the theories represent conceptually related approaches (as can be seen with the Big-Five and Reinforcement Sensitivity). For cultural differences in trait mindfulness, very little evidence on cultural similarity or differences in mindfulness was present. While a small amount of good qualitative evidence on different understandings exists (for example: Feng et al., 2018) only few studies assessed cultural comparability, and if authors reported these analyses, they were conducted with a limited scope. Looking into the future from my perspective, trait mindfulness needs further engagement with the wider individual difference literature, but also with its developmental history both on conceptual and empirical levels. In closing, I hope that my thesis not only helped to show where the road could lead, but also helped to pave the beginning.

**Bibliography**

Adams, C. E., McVay, M. A., Kinsaul, J., Benitez, L., Vinci, C., Stewart, D. W., & Copeland, A. L. (2012). Unique relationships between facets of mindfulness and eating pathology among female smokers. *Eating Behaviors*, *13*(4), 390–393. https://doi.org/10.1016/j.eatbeh.2012.05.009

Adams, C. E., McVay, M. A., Stewart, D. W., Vinci, C., Kinsaul, J., Benitez, L., & Copeland, A. L. (2014). Mindfulness ameliorates the relationship between weight concerns and smoking behavior in female smokers: A cross-sectional investigation. *Mindfulness*, *5*(2), 179–185. https://doi.org/10.1007/s12671-012-0163-9

Aguado, J., Luciano, J. V., Cebolla, A., Serrano-Blanco, A., Soler, J., & García-Campayo, J. (2015). Bifactor analysis and construct validity of the five facet mindfulness questionnaire (FFMQ) in non-clinical Spanish samples. *Frontiers in Psychology*, *6*, 404. https://doi.org/10.3389/fpsyg.2015.00404

Anālayo. (2004). *Satipaṭṭhāna: The direct path to realization*. Windhorse Publications.

Anālayo, B. (2018). Mindfulness constructs in early Buddhism and Theravāda: Another contribution to the memory debate. *Mindfulness*, *9*(4), 1047–1051. https://doi.org/10.1007/s12671-018-0967-3

Anālayo, B. (2019). Adding historical depth to definitions of mindfulness. *Current Opinion in Psychology*, *28*, 11–14. https://doi.org/10.1016/j.copsyc.2018.09.013

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411–423. https://doi.org/10.1037/0033-2909.103.3.411

Andrei, F., Vesely, A., & Siegling, A. B. (2016). An examination of concurrent and incremental validity of four mindfulness scales. *Journal of Psychopathology and Behavioral Assessment*, *38*(4), 559–571. https://doi.org/10.1007/s10862-016-9546-x

Antoniazzi, D., & Klein, R. (2019). Risky riders: A comparison of personality theories on motorcyclist riding behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour*, *62*, 33–44. https://doi.org/10.1016/j.trf.2018.12.008

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397–438. https://doi.org/10.1080/10705510903008204

Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(4), 495–508. https://doi.org/10.1080/10705511.2014.919210

Astani, A.-Iuliana. (2016). Mindfulness and unconditional self-acceptance as protective factors against thin ideal internalization. *Annals of A.I. I. Cuza University. Psychology Series*, *25*(1), 37–46.

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

Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, *13*(1), 27–45. https://doi.org/10.1177/1073191105283504

Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., Walsh, E., Duggan, D., & Williams, J. M. G. (2008). Construct validity of the Five Facet Mindfulness Questionnaire in meditating and nonmeditating samples. *Assessment*, *15*(3), 329–342. https://doi.org/10.1177/1073191107313003

Barnes, S., Brown, K. W., Krusemark, E., Campbell, W. K., & Rogge, R. D. (2007). The role of mindfulness in romantic relationship satisfaction and responses to relationship stress.

*Journal of Marital and Family Therapy*, *33*(4), 482–500. https://doi.org/10.1111/j.1752-0606.2007.00033.x

Barnhofer, T., Duggan, D. S., & Griffith, J. W. (2011). Dispositional mindfulness moderates the relation between neuroticism and depressive symptoms. *Personality and Individual Differences*, *51*(8), 958–962. https://doi.org/10.1016/j.paid.2011.07.032

Barrett, P. (1986). Factor comparison: An examination of three methods. *Personality and Individual Differences*, *7*(3), 327–340. https://doi.org/10.1016/0191-8869(86)90008-5

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, *42*(5), 815–824. https://doi.org/10.1016/j.paid.2006.09.018

Bellah, R. N. (2007). *Habits of the heart: Individualism and commitment in American life* (3rd edition). University of California Press.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Berchtold, A. (2019). Sequence analysis and transition models. In J. C. Choe (Ed.), *Encyclopedia of Animal Behavior (Second Edition)* (pp. 506–517). Academic Press. https://doi.org/10.1016/B978-0-12-809633-8.01241-3

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

Bergomi, C., Tschacher, W., & Kupper, Z. (2013b). The assessment of mindfulness with self-report measures: Existing scales and open issues. *Mindfulness*, *4*(3), 191–202. https://doi.org/10.1007/s12671-012-0110-9

Bernaards, C. A., & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, *65*(5), 676–696. https://doi.org/10.1177/0013164404272507

Bernstein, J. H. (2015). Transdisciplinarity: A review of Its origins, development, and current issues. *Journal of Research Practice*, *11*(1), 1–20.

Berry, J. W. (1989). Imposed etics-amics-derived etics: The operationalization of a compelling idea. *International Journal of Psychology*, *24*(2–6), 721–735. https://doi.org/10.1080/00207598908247841

Berry, J. W. (1999). Emics and etics: A symbiotic conception. *Culture & Psychology*, *5*(2), 165–171. https://doi.org/10.1177/1354067X9952004

Beshara, M., Hutchinson, A. D., & Wilson, C. (2013). Does mindfulness matter? Everyday mindfulness, mindful eating and self-reported serving size of energy dense foods among a sample of South Australian adults. *Appetite*, *67*, 25–29. https://doi.org/10.1016/j.appet.2013.03.012

Beukeboom, C. J., Tanis, M., & Vermeulen, I. E. (2013). The language of extraversion: Extraverted people talk more abstractly, introverts are more concrete. *Journal of Language and Social Psychology*, *32*(2), 191–201. https://doi.org/10.1177/0261927X12460844

Bijnen, E. J., & Poortinga, Y. H. (1988). The questionable value of cross-cultural comparisons with the Eysenck personality questionnaire. *Journal of Cross-Cultural Psychology*, *19*(2), 193–202. https://doi.org/10.1177/0022022188192005

Bijnen, E. J., Van Der Net, T. Z. J., & Poortinga, Y. H. (1986). On cross-cultural comparative studies with the Eysenck personality questionnaire. *Journal of Cross-Cultural Psychology*, *17*(1), 3–16. https://doi.org/10.1177/0022002186017001001

Blanke, E. S., & Brose, A. (2017). Mindfulness in daily life: A Multidimensional approach. *Mindfulness*, *8*(3), 737–750. https://doi.org/10.1007/s12671-016-0651-4

Blasberg, S. A., Rogers, K. H., & Paulhus, D. L. (2014). The bidimensional impression management index (BIMI): Measuring agentic and communal forms of impression management. *Journal of Personality Assessment*, *96*(5), 523–531. https://doi.org/10.1080/00223891.2013.862252

Bleidorn, W., Hopwood, C. J., & Lucas, R. E. (2018). Life events and personality trait change. *Journal of Personality*, *86*(1), 83–96. https://doi.org/10.1111/jopy.12286

Bleidorn, W., Klimstra, T. A., Denissen, J. J. A., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2013). Personality maturation around the world: A cross-cultural examination of social-investment theory. *Psychological Science*, *24*(12), 2530–2540. https://doi.org/10.1177/0956797613498396

Boer, D., Hanke, K., & He, J. (2018). On detecting systematic measurement error in cross-cultural research: A review and critical reflection on equivalence and invariance tests. *Journal of Cross-Cultural Psychology*, *49*(5), 713–734. https://doi.org/10.1177/0022022117749042

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

Bollen, K. A. (1989). *Structural equations with latent variables* (pp. xiv, 514). John Wiley & Sons. https://doi.org/10.1002/9781118619179

Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science*, *5*(1), 184–186. https://doi.org/10.1177/2167702616657069

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, *71*(3), 425–440. https://doi.org/10.1007/s11336-006-1447-6

Bowen, S., Chawla, N., Collins, S. E., Witkiewitz, K., Hsu, S., Grow, J., Clifasefi, S., Garner, M., Douglass, A., Larimer, M. E., & Marlatt, A. (2009). Mindfulness-based relapse prevention for substance use disorders: A pilot efficacy trial. *Substance Abuse*, *30*(4), 295–305. https://doi.org/10.1080/08897070903250084

Bowen, S., Witkiewitz, K., Clifasefi, S. L., Grow, J., Chawla, N., Hsu, S. H., Carroll, H. A., Harrop, E., Collins, S. E., Lustyk, M. K., & Larimer, M. E. (2014). Relative efficacy of mindfulness-based relapse prevention, standard relapse prevention, and treatment as usual for substance use disorders: A randomized clinical trial. *JAMA Psychiatry*, *71*(5), 547–556. https://doi.org/10.1001/jamapsychiatry.2013.4546

Bravo, A. J., Boothe, L. G., & Pearson, M. R. (2016). Getting personal with mindfulness: A latent profile analysis of mindfulness and psychological outcomes. *Mindfulness*, *7*(2), 420–432. https://doi.org/10.1007/s12671-015-0459-7

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, *84*(4), 822–848. https://doi.org/10.1037/0022-3514.84.4.822

Brown, K. W., Weinstein, N., & Creswell, J. D. (2012). Trait mindfulness modulates neuroendocrine and affective responses to social evaluative threat. *Psychoneuroendocrinology*, *37*(12), 2037–2041. https://doi.org/10.1016/j.psyneuen.2012.04.003

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456–466. https://doi.org/10.1037/0033-2909.105.3.456

Cardaciotto, L., Herbert, J. D., Forman, E. M., Moitra, E., & Farrow, V. (2008). The assessment of present-moment awareness and acceptance. *Assessment*, *15*(2), 204–223. https://doi.org/10.1177/1073191107311467

Carlson, L. E., & Brown, K. W. (2005). Validation of the Mindful Attention Awareness Scale in a cancer population. *Journal of Psychosomatic Research*, *58*(1), 29–33. https://doi.org/10.1016/j.jpsychores.2004.04.366

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

Carpenter, J. K., Conroy, K., Gomez, A. F., Curren, L. C., & Hofmann, S. G. (2019). The relationship between trait mindfulness and affective symptoms: A meta-analysis of the Five Facet Mindfulness Questionnaire (FFMQ). *Clinical Psychology Review*, *74*, 101785. https://doi.org/10.1016/j.cpr.2019.101785

Chadwick, P., Hember, M., Symes, J., Peters, E., Kuipers, E., & Dagnan, D. (2008). Responding mindfully to unpleasant thoughts and images: Reliability and validity of the Southampton mindfulness questionnaire (SMQ). *British Journal of Clinical Psychology*, *47*(4), 451–455. https://doi.org/10.1348/014466508X314891

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

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, *95*(5), 1005–1018. https://doi.org/10.1037/a0013193

Chen, S., Cui, H., Zhou, R., Jia, P. E. A. of L. K. B. Y., Beijing, U. N., Beijing, 100875, & China. (2012). Revision of Mindful Attention Awareness Scale(MAAS). *Chinese Journal of Clinical Psychology*, *2*, 148–151.

Chen, S., You, B., & Jackson, T. (2021). Facets of mindfulness as predictors of emotional distress among chinese adults with chronic musculoskeletal pain. *Mindfulness*, *12*(3), 775–783. https://doi.org/10.1007/s12671-020-01548-7

Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology*, *31*(2), 187–212. https://doi.org/10.1177/0022022100031002003

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Chiesa, A. (2013). The Difficulty of Defining Mindfulness: Current Thought and Critical Issues. *Mindfulness*, *4*(3), 255–268. https://doi.org/10.1007/s12671-012-0123-4

Chiesa, A., & Malinowski, P. (2011). Mindfulness-based approaches: Are they all the same? *Journal of Clinical Psychology*, *67*(4), 404–424. https://doi.org/10.1002/jclp.20776

Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and monte carlo simulations. *Journal of Statistical Software*, *39*(8), 1–30. https://doi.org/10.18637/jss.v039.i08

Christopher, M. S., Charoensuk, S., Gilbert, B. D., Neary, T. J., & Pearce, K. L. (2009). Mindfulness in Thailand and the United States: A case of apples versus oranges? *Journal of Clinical Psychology*, *65*(6), 590–612. https://doi.org/10.1002/jclp.20580

Christopher, M. S., Christopher, V., & Charoensuk, S. (2009). Assessing “Western” Mindfulness Among Thai Theravāda Buddhist Monks. *Mental Health, Religion & Culture*, *12*(3), 303–314. https://doi.org/10.1080/13674670802651487

Chung, P.-K., Zhang, C.-Q., & Zhang, C.-Q. (2014). Psychometric validation of the Toronto mindfulness scale – trait version in Chinese college students. *Europe’s Journal of Psychology*, *10*(4), 726–739. https://doi.org/10.5964/ejop.v10i4.776

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. *Journal of Informetrics*, *5*(1), 146–166. https://doi.org/10.1016/j.joi.2010.10.002

Coffey, K. A., & Hartman, M. (2008). Mechanisms of action in the inverse relationship between mindfulness and psychological distress. *Complementary Health Practice Review*, *13*(2), 79–91. https://doi.org/10.1177/1533210108316307

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Contributors,  semTools. (2016). *{semTools}: Useful tools for structural equation modeling*.

Corr, P. J., & Cooper, A. J. (2016). The Reinforcement Sensitivity Theory of Personality Questionnaire (RST-PQ): Development and validation. *Psychological Assessment*, *28*(11), 1427–1440. https://doi.org/10.1037/pas0000273

Corr, P. J., DeYoung, C. G., & McNaughton, N. (2013). Motivation and personality: A neuropsychological perspective. *Social and Personality Psychology Compass*, *7*(3), 158–175. https://doi.org/10.1111/spc3.12016

Corr, P. J., & Matthews, Gerald. (Eds.). (2020). *The Cambridge handbook of personality psychology.* (2nd ed.). Cambridge University Press.

Corr, P. J., & McNaughton, N. (2012). Neuroscience and approach/avoidance personality traits: A two stage (valuation–motivation) approach. *Neuroscience & Biobehavioral Reviews*, *36*(10), 2339–2354. https://doi.org/10.1016/J.NEUBIOREV.2012.09.013

Cosme, D., & Wiens, S. (2015). Self-reported trait mindfulness and affective reactivity: A motivational approach using multiple psychophysiological measures. *PLOS ONE*, *10*(3), e0119466. https://doi.org/10.1371/journal.pone.0119466

Costa, P. T., & McCrae, R. R. (2006). Age changes in personality and their origins: Comment on Roberts, Walton, and Viechtbauer (2006). *Psychological Bulletin*, *132*(1), 26–28. https://doi.org/10.1037/0033-2909.132.1.26

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *The Behavioral and Brain Sciences*, *33*(2–3), 137–150; discussion 150-193. https://doi.org/10.1017/S0140525X09991567

Crane, P. K., Belle, G. van, & Larson, E. B. (2004). Test bias in a cognitive test: Differential item functioning in the CASI. *Statistics in Medicine*, *23*(2), 241–256. https://doi.org/10.1002/sim.1713

Creswell, J. D. (2015). Biological pathways linking mindfulness with health. In *Handbook of mindfulness: Theory, research, and practice* (pp. 426–440). The Guilford Press.

Creswell, J. D. (2017). Mindfulness Interventions. *Annual Review of Psychology*, *68*(1), 491–516. https://doi.org/10.1146/annurev-psych-042716-051139

Creswell, J. D., & Lindsay, E. K. (2014). How does mindfulness training affect health? A mindfulness stress buffering account. *Current Directions in Psychological Science*, *23*(6), 401–407. https://doi.org/10.1177/0963721414547415

Creswell, J. D., Way, B. M., Eisenberger, N. I., & Lieberman, M. D. (2007). Neural correlates of dispositional mindfulness during affect labeling. *Psychosomatic Medicine*, *69*(6), 560–565. https://doi.org/10.1097/PSY.0b013e3180f6171f

Curtis, M. (2019). *An exploratory thematic analysis of mindfulness definitions, test instruments, and methods used in current research* [Ph.D.]. Ashford University.

De Raad, B., Barelds, D. P. H., Timmerman, M. E., De Roover, K., Mlačić, B., & Church, A. T. (2014). Towards a pan-cultural personality structure: Input from 11 psycholexical studies. *European Journal of Personality*, *28*(5), 497–510. https://doi.org/10.1002/per.1953

Demarzo, M. M. P. M., Montero-Marin, J. P. D., Stein, P. K. P. D., Cebolla, A. P. D., Provinciale, J. G. P. D., & García-Campayo, J. M. (2014). Mindfulness may both moderate and mediate the effect of physical fitness on cardiovascular responses to stress: A speculative hypothesis. *Frontiers in Physiology*, *5*. https://doi.org/10.3389/fphys.2014.00105

Deng, Y.-Q., Li, S., Tang, Y.-Y., Zhu, L.-H., Ryan, R., & Brown, K. (2012). Psychometric properties of the Chinese translation of the Mindful Attention Awareness Scale (MAAS). *Mindfulness*, *3*(1), 10–14. https://doi.org/10.1007/s12671-011-0074-1

DeYoung, C. G. (2015). Cybernetic Big Five theory. *Journal of Research in Personality*, *56*, 33–58. https://doi.org/10.1016/J.JRP.2014.07.004

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., won Choi, D., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, *97*(2), 143–156. https://doi.org/10.1007/s11205-009-9493-y

Dolatyar, K., & Walker, B. R. (2020). Reinforcement sensitivity theory and mindfulness. *Personality and Individual Differences*, *163*(1), 110089. https://doi.org/10.1016/j.paid.2020.110089

Druica, E., & Ianole-Calin, R. (2018). Simply clustering. Making new sense in the Five Facets Mindfulness Questionnaire. *Romanian Statistical Review*, *66*(1), 61–81.

Dundas, I., Vøllestad, J., Binder, P.-E., & Sivertsen, B. (2013). The Five Factor Mindfulness Questionnaire in Norway. *Scandinavian Journal of Psychology*, *54*(3), 250–260. https://doi.org/10.1111/sjop.12044

Epskamp, S. (2020). Psychometric network models from time-series and panel data. *Psychometrika*, *85*(1), 206–231. https://doi.org/10.1007/s11336-020-09697-3

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, *42*(3), 509–529. https://doi.org/10.1080/00273170701382864

Farb, N. A. S., Segal, Z. V., & Anderson, A. K. (2013). Mindfulness meditation training alters cortical representations of interoceptive attention. *Social Cognitive and Affective Neuroscience*, *8*(1), 15–26. https://doi.org/10.1093/scan/nss066

Farh, J.-L., Cannella, A. A., & Lee, C. (2006). Approaches to scale development in Chinese management research. *Management and Organization Review*, *2*(3), 301–318. https://doi.org/10.1111/j.1740-8784.2006.00055.x

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

Feng, X. J., Krägeloh, C. U., Billington, D. R., & Siegert, R. J. (2018). To what extent is mindfulness as presented in commonly used mindfulness questionnaires different from how it is conceptualized by senior ordained buddhists? *Mindfulness*, *9*(2), 441–460. https://doi.org/10.1007/s12671-017-0788-9

Fernandez, A. C., Wood, M. D., Stein, L. A. R., & Rossi, J. S. (2010). Measuring mindfulness and examining its relationship with alcohol use and negative consequences. *Psychology of Addictive Behaviors*, *24*(4), 608–616. https://doi.org/10.1037/a0021742

Field, A. P., Miles, J., & Field, Zoë. (2012). *Discovering statistics using R*. SAGE.

Fischer, R. (2017). *Personality, values, culture: An evolutionary approach*. Cambridge University Press. https://doi.org/10.1017/9781316091944

Fischer, R., Bortolini, T., Karl, J. A., Zilberberg, M., Robinson, K., Rabelo, A., Gemal, L., Wegerhoff, D., Nguyễn, T. B. T., Irving, B., Chrystal, M., & Mattos, P. (2020). Rapid review and meta-meta-analysis of self-guided interventions to address anxiety, depression, and stress during

COVID-19 social distancing. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.563876

Fischer, R., Ferreira, M. C., Jiang, D.-Y. Y., Cheng, B.-S. S., Achoui, M. M., Wong, C. C., Baris, G., Mendoza, S., van Meurs, N., Achmadi, D., Hassan, A., Zeytinoglu, G., Dalyan, F., Harb, C., Darwish, D. D., & Assmar, E. M. (2011). Are perceptions of organizational justice universal? An exploration of measurement invariance across thirteen cultures. *Social Justice Research*, *24*(4), 297–313. https://doi.org/10.1007/s11211-011-0142-7

Fischer, R., Ferreira, M. C., Van Meurs, N., Gok, K., Jiang, D.-Y., Fontaine, J. R. J., Harb, C., Cieciuch, J., Achoui, M., Mendoza, M. S. D., Hassan, A., Achmadi, D., Mogaji, A. A., & Abubakar, A. (2019). Does organizational formalization facilitate voice and helping organizational citizenship behaviors? It depends on (national) uncertainty norms. *Journal of International Business Studies*, *50*(1), 125–134. https://doi.org/10.1057/s41267-017-0132-6

Fischer, R., & Fontaine, J. R. J. (2010). Methods for investigating structural equivalence. In D. Matsumoto & F. J. R. E. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 179–215). Cambridge University Press. https://doi.org/10.1017/CBO9780511779381.010

Fischer, R., & Karl, J. A. (2019). A primer to (cross-cultural) multi-group invariance testing possibilities in R. *Frontiers in Psychology*, *10*, 1507. https://doi.org/10.3389/fpsyg.2019.01507

Fischer, R., & Karl, J. A. (2020). The network architecture of individual differences: Personality, reward-sensitivity, and values. *Personality and Individual Differences*, *160*(1), 109922. https://doi.org/10.1016/j.paid.2020.109922

Fischer, R., & Poortinga, Y. H. (2018). Addressing methodological challenges in culture-comparative research. *Journal of Cross-Cultural Psychology*, *49*(5), 691–712. https://doi.org/10.1177/0022022117738086

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, *80*(6), 1011–1027. https://doi.org/10.1037/0022-3514.80.6.1011

Fleeson, W., & Jayawickreme, E. (2015). Whole Trait Theory. *Journal of Research in Personality*, *56*, 82–92. https://doi.org/10.1016/j.jrp.2014.10.009

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*(4), 466–491. https://doi.org/10.1037/1082-989X.9.4.466

Foa, E. B., Johnson, K. M., Feeny, N. C., & Treadwell, K. R. (2001). The child PTSD Symptom Scale: A preliminary examination of its psychometric properties. *Journal of Clinical Child Psychology*, *30*(3), 376–384. https://doi.org/10.1207/S15374424JCCP3003_9

Fontaine, J. (2005). Equivalence. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 1, pp. 803–813).

Frewen, P. A., Evans, E. M., Maraj, N., Dozois, D. J. A., & Partridge, K. (2008). Letting go: Mindfulness and negative automatic thinking. *Cognitive Therapy and Research*, *32*(6), 758–774. https://doi.org/10.1007/s10608-007-9142-1

Garland, E. L., Beck, A. C., Lipschitz, D. L., & Nakamura, Y. (2015). Dispositional mindfulness predicts attenuated waking salivary cortisol levels in cancer survivors: A latent growth curve analysis. *Journal of Cancer Survivorship*, *9*(2), 215–222. https://doi.org/10.1007/s11764-014-0402-2

Garland, E. L., Farb, N. A., Goldin, P. R., & Fredrickson, B. L. (2015). The Mindfulness-to-Meaning Theory: Extensions, applications, and challenges at the attention–appraisal–emotion interface. *Psychological Inquiry*, *26*(4), 377–387. https://doi.org/10.1080/1047840X.2015.1092493

Gehlert, S., Murray, A., Sohmer, D., McClintock, M., Conzen, S., & Olopade, O. (2010). The importance of transdisciplinary collaborations for understanding and resolving health disparities. *Social Work in Public Health*, *25*(3), 408–422. https://doi.org/10.1080/19371910903241124

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., Aycan, Z., Boehnke, K., Boski, P., Cabecinhas, R., Chan, D., Chhokar, J., D'Amato, A., Ferrer, M., Fischlmayr, I. C., … Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, *332*(6033), 1100–1104. https://doi.org/10.1126/science.1197754

Gethin, R. (2011). On some definitions of mindfulness. *Contemporary Buddhism*, *12*(1), 263–279. https://doi.org/10.1080/14639947.2011.564843

Gilpin, R. (2008). The use of Theravāda Buddhist practices and perspectives in mindfulness-based cognitive therapy. *Contemporary Buddhism*, *9*(2), 227–251. https://doi.org/10.1080/14639940802556560

Giluk, T. L. (2009). Mindfulness, Big Five personality, and affect: A meta-analysis. *Personality and Individual Differences*, *47*(8), 805–811. https://doi.org/10.1016/J.PAID.2009.06.026

Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, *55*(3), 377–393. https://doi.org/10.1177/0013164495055003002

Goldberg, L. R. (2006). Doing it all Bass-Ackwards: The development of hierarchical factor structures from the top down. *Journal of Research in Personality*, *40*(4), 347–358. https://doi.org/10.1016/j.jrp.2006.01.001

Goldin, P. R., & Gross, J. J. (2010). Effects of mindfulness-based stress reduction (MBSR) on emotion regulation in social anxiety disorder. *Emotion*, *10*(1), 83–91. https://doi.org/10.1037/a0018441

Golino, H., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLOS ONE*, *12*(6), e0174035. https://doi.org/10.1371/journal.pone.0174035

Golino, H., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., Sadana, R., Thiyagarajan, J. A., & Martinez-Molina, A. (2020). Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychological Methods*, *25*(3), 292–320. https://doi.org/10.1037/met0000255

Göllner, R., Damian, R. I., Rose, N., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). Is doing your homework associated with becoming more conscientious? *Journal of Research in Personality*, *71*, 1–12. https://doi.org/10.1016/j.jrp.2017.08.007

Gorsuch, R. L. (1983). *Factor Analysis, 2nd Edition* (2nd edition). Lawrence Erlbaum Associates.

Gračanin, A., Gunjača, V., Tkalčić, M., Kardum, I., Bajšanski, I., & Perak, B. (2017). Struktura usredotočene svjesnosti i njezina povezanost s crtama ličnosti i emocionalnim reagiranjem. *Psihologijske teme*, *26*(3), 675–700. https://doi.org/10.31820/pt.26.3.9

Gray, J. A. (1970). The psychophysiological basis of introversion-extraversion. *Behaviour Research and Therapy*, *8*(3), 249–266. https://doi.org/10.1016/0005-7967(70)90069-0

Gray, J. A. (2004). *Consciousness: Creeping up on the hard problem*. Oxford University Press.

Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness, empathy, and helping: A person × situation perspective. *Journal of Personality and Social Psychology*, *93*(4), 583–599. https://doi.org/10.1037/0022-3514.93.4.583

Grossman, P. (2011). Defining mindfulness by how poorly I think I pay attention during everyday awareness and other intractable problems for psychology’s (re)invention of mindfulness: Comment on Brown et al. (2011). *Psychological Assessment*, *23*(4), 1034–1040. https://doi.org/10.1037/a0022713

Grossman, P., & Dam, N. T. V. (2011). Mindfulness, by any other name…: Trials and tribulations of sati in western psychology and science. *Contemporary Buddhism*, *12*(1), 219–239. https://doi.org/10.1080/14639947.2011.564841

Grossman, P., Niemann, L., Schmidt, S., & Walach, H. (2004). Mindfulness-based stress reduction and health benefits: A meta-analysis. *Journal of Psychosomatic Research*, *57*(1), 35–43. https://doi.org/10.1016/S0022-3999(03)00573-7

Gu, J., Strauss, C., Crane, C., Barnhofer, T., Karl, A., Cavanagh, K., & Kuyken, W. (2016). Examining the factor structure of the 39-item and 15-item versions of the Five Facet Mindfulness Questionnaire before and after mindfulness-based cognitive therapy for people with recurrent depression. *Psychological Assessment*, *28*(7), 791–802. https://doi.org/10.1037/pas0000263

Guàrdia-Olmos, J., Peró-Cebollero, M., Benítez-Borrego, S., & Fox, J. (2013). Using SEM Library in R software to Analyze Exploratory Structural Equation Models. *Proceedings of the 59th ISI World Statistics Congress*, 4600–4605. https://www.statistics.gov.hk/wsc/CPS105-P6-S.pdf

Haas, B. W., Omura, K., Constable, R. T., & Canli, T. (2007). Emotional conflict and neuroticism: Personality-dependent activation in the amygdala and subgenual anterior cingulate. *Behavioral Neuroscience*, *121*(2), 249–256. https://doi.org/10.1037/0735-7044.121.2.249

Haliwa, I., Wilson, J. M., Spears, S. K., Strough, J., & Shook, N. J. (2020). Exploring facets of the mindful personality: Dispositional mindfulness and the Big Five. *Personality and Individual Differences*, 110469. https://doi.org/10.1016/j.paid.2020.110469

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Hambleton, R. K., & Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*, *12*(3), 38–47. https://doi.org/10.1111/j.1745-3992.1993.tb00543.x

Hamill, T. S., Pickett, S. M., Amsbaugh, H. M., & Aho, K. M. (2015). Mindfulness and acceptance in relation to Behavioral Inhibition System sensitivity and psychological distress. *Personality and Individual Differences*, *72*, 24–29. https://doi.org/10.1016/J.PAID.2014.08.007

Hanley, A. W. (2016). The mindful personality: Associations between dispositional mindfulness and the Five Factor Model of personality. *Personality and Individual Differences*, *91*, 154–158. https://doi.org/10.1016/J.PAID.2015.11.054

Hanley, A. W., Baker, A. K., & Garland, E. L. (2018). The mindful personality II: Exploring the metatraits from a cybernetic perspective. *Mindfulness*, *9*(3), 972–979. https://doi.org/10.1007/s12671-017-0836-5

Hanley, A. W., & Garland, E. L. (2017). The mindful personality: A meta-analysis from a cybernetic perspective. *Mindfulness*, *8*(6), 1456–1470. https://doi.org/10.1007/s12671-017-0736-8

Hargus, E., Crane, C., Barnhofer, T., & Williams, J. M. G. (2010). Effects of mindfulness on meta-awareness and specificity of describing prodromal symptoms in suicidal depression. *Emotion (Washington, D.C.)*, *10*(1), 34–42. https://doi.org/10.1037/a0016825

Harnett, P. H., Reid, N., Loxton, N. J., & Lee, N. (2016). The relationship between trait mindfulness, personality and psychological distress: A revised reinforcement sensitivity theory perspective. *Personality and Individual Differences*, *99*, 100–105. https://doi.org/10.1016/j.paid.2016.04.085

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (pp. xvii, 507). Guilford Press.

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What’s wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology*, *82*(6), 903–918.

Hendriks, T., Warren, M. A., Schotanus-Dijkstra, M., Hassankhan, A., Graafsma, T., Bohlmeijer, E., & Jong, J. de. (2019). How WEIRD are positive psychology interventions? A bibliometric analysis of randomized controlled trials on the science of well-being. *The Journal of Positive Psychology*, *14*(4), 489–501. https://doi.org/10.1080/17439760.2018.1484941

Henrich, J. (2020). *The WEIRDest people in the world: Jow the West became psychologically peculiar and particularly prosperous*. Farrar, Straus and Giroux.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29–29. https://doi.org/10.1038/466029a

Higgins, J., & Eden, R. (2018). Emerging understandings of mindfulness through experiential awareness. *Learning: Research and Practice*, *4*(1), 102–111. https://doi.org/10.1080/23735082.2018.1428144

Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. SAGE.

Holgado–Tello, F. P., Chacón–Moscoso, S., Barbero–García, I., & Vila–Abad, E. (2008). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity*, *44*(1), 153. https://doi.org/10.1007/s11135-008-9190-y

Hölzel, B. K., Lazar, S. W., Gard, T., Schuman-Olivier, Z., Vago, D. R., & Ott, U. (2011). How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. *Perspectives on Psychological Science*, *6*(6), 537–559. https://doi.org/10.1177/1745691611419671

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

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

Hwang, K. (1987). Face and favor: The Chinese power game. *American Journal of Sociology*, *92*(4), 944–974. https://doi.org/10.1086/228588

Iani, L., Lauriola, M., Cafaro, V., & Didonna, F. (2017). Dimensions of mindfulness and their relations with psychological well-being and neuroticism. *Mindfulness*, *8*(3), 664–676. https://doi.org/10.1007/s12671-016-0645-2

Jodoin, M. G., & Gierl, M. J. (2001). Evaluating type I rrror and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, *14*(4), 329–349. https://doi.org/10.1207/S15324818AME1404_2

Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results. *General Hospital Psychiatry*, *4*(1), 33–47. https://doi.org/10.1016/0163-8343(82)90026-3

Kabat-Zinn, J. (1990). *Full catastrophe living*. Delacorte Press.

Kabat-Zinn, J. (1994). *Wherever you go, there you are: Mindfulness meditation in everyday life.* Hyperion.

Kabat-Zinn, J. (2011). Some reflections on the origins of MBSR, skillful means, and the trouble with maps. *Contemporary Buddhism*, *12*(1), 281–306. https://doi.org/10.1080/14639947.2011.564844

Kaiser, T. (2018). *ESEM in R: a tutorial*. https://tkaiser.science/esemR.html

Kandler, C. (2012). Nature and nurture in personality development: The case of neuroticism and extraversion. *Current Directions in Psychological Science*, *21*(5), 290–296. https://doi.org/10.1177/0963721412452557

Kang, C., & Whittingham, K. (2010). Mindfulness: A dialogue between Buddhism and clinical psychology. *Mindfulness*, *1*(3), 161–173. https://doi.org/10.1007/s12671-010-0018-1

Kang, Y., McNeish, D. M., & Hancock, G. R. (2016). The role of measurement quality on practical guidelines for assessing measurement and structural invariance. *Educational and Psychological Measurement*, *76*(4), 533–561. https://doi.org/10.1177/0013164415603764

Karl, J. A., & Fischer, R. (2019). Individual differences and mindfulness. *PsyArXiv*. https://doi.org/10.31234/OSF.IO/Z2CX6

Karl, J. A., & Fischer, R. (2020). Revisiting the five-facet structure of mindfulness. *Measurement Instruments for the Social Sciences*, *2*(1), 7. https://doi.org/10.1186/s42409-020-00014-3

Karl, J. A., & Fischer, R. (2021). *Affect and state mindfulness*. https://doi.org/10.31234/osf.io/jqhu7

Karl, J. A., Fischer, R., & Jose, P. E. (2021). The development of mindfulness in young adults: The relationship of personality, reinforcement sensitivity, and mindfulness. *Mindfulness*. https://doi.org/10.1007/s12671-020-01576-3

Karl, J. A., Johnson, F., Bucci, L., & Fischer, R. (2021). In search of mindfulness: A review and reconsideration of cultural dynamics from a cognitive perspective. *Journal of the Royal Society of New Zealand*.

Karl, J. A., Méndez Prado, S. M., Gračanin, A., Verhaeghen, P., Ramos, A., Mandal, S. P., Michalak, J., Zhang, C.-Q., Schmidt, C., Tran, U. S., Druica, E., Solem, S., Astani, A., Liu, X., Luciano, J. V., Tkalčić, M., Lilja, J. L., Dundas, I., Wong, S. Y. S. Y., … Fischer, R. (2020). The cross-cultural validity of the Five-Facet Mindfulness Questionnaire across 16 countries. *Mindfulness*. https://doi.org/10.1007/s12671-020-01333-6

Karyadi, K. A., VanderVeen, J. D., & Cyders, M. A. (2014). A meta-analysis of the relationship between trait mindfulness and substance use behaviors. *Drug and Alcohol Dependence*, *143*, 1–10. https://doi.org/10.1016/j.drugalcdep.2014.07.014

Kee, Y. H., Li, C., Kong, L. C., Tang, C. J., & Chuang, K.-L. (2019). Scoping review of mindfulness research: A topic modelling approach. *Mindfulness*, *10*(8), 1474–1488. https://doi.org/10.1007/s12671-019-01136-4

Keng, S.-L., Smoski, M. J., & Robins, C. J. (2011). Effects of mindfulness on psychological health: A review of empirical studies. *Clinical Psychology Review*, *31*(6), 1041–1056. https://doi.org/10.1016/j.cpr.2011.04.006

Keune, P. M., Bostanov, V., Kotchoubey, B., & Hautzinger, M. (2012). Mindfulness versus rumination and behavioral inhibition: A perspective from research on frontal brain asymmetry. *Personality and Individual Differences*, *53*(3), 323–328. https://doi.org/10.1016/j.paid.2012.03.034

Kiken, L. G., Garland, E. L., Bluth, K., Palsson, O. S., & Gaylord, S. A. (2015). From a state to a trait: Trajectories of state mindfulness in meditation during intervention predict changes in trait mindfulness. *Personality and Individual Differences*, *81*, 41–46. https://doi.org/10.1016/j.paid.2014.12.044

Klackl, J., Jonas, E., & Fritsche, I. (2018). Neural evidence that the behavioral inhibition system is involved in existential threat processing. *Social Neuroscience*, *13*(3), 355–371. https://doi.org/10.1080/17470919.2017.1308880

Koestner, R., Bernieri, F., & Zuckerman, M. (1992). Self-regulation and consistency between attitudes, traits, and behaviors. *Personality and Social Psychology Bulletin*, *18*(1), 52–59. https://doi.org/10.1177/0146167292181008

Krägeloh, C. (2020). *Mindfulness research and terminology science*.

Krijn, M., Emmelkamp, P. M. G., Olafsson, R. P., & Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: A review. *Clinical Psychology Review*, *24*(3), 259–281. https://doi.org/10.1016/j.cpr.2004.04.001

Kucinskas, J. (2014). The unobtrusive tactics of religious movements. *Sociology of Religion*, *75*(4), 537–550. https://doi.org/10.1093/socrel/sru055

Kucinskas, J. (2018). *The mindful elite: Mobilizing from the inside out*. Oxford University Press.

Lau, M. A., Bishop, S. R., Segal, Z. V., Buis, T., Anderson, N. D., Carlson, L., Shapiro, S., Carmody, J., Abbey, S., & Devins, G. (2006). The Toronto Mindfulness Scale: Development and validation. *Journal of Clinical Psychology*, *62*(12), 1445–1467. https://doi.org/10.1002/jclp.20326

Lavender, J. M., Gratz, K. L., & Tull, M. T. (2011). Exploring the relationship between facets of mindfulness and eating pathology in women. *Cognitive Behaviour Therapy*, *40*(3), 174–182. https://doi.org/10.1080/16506073.2011.555485

Li, C.-H. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Lilja, J. L., Frodi-Lundgren, A., Hanse, J. J., Josefsson, T., Lundh, L.-G., Sköld, C., Hansen, E., & Broberg, A. G. (2011). Five Facets Mindfulness Questionnaire—reliability and factor structure: A Swedish version. *Cognitive Behaviour Therapy*, *40*(4), 291–303. https://doi.org/10.1080/16506073.2011.580367

Lilja, J. L., Lundh, L.-G., Josefsson, T., & Falkenström, F. (2013). Observing as an essential facet of mindfulness: A comparison of FFMQ patterns in meditating and non-meditating individuals. *Mindfulness*, *4*(3), 203–212. https://doi.org/10.1007/s12671-012-0111-8

Lindsay, E. K., & Creswell, J. D. (2017). Mechanisms of mindfulness training: Monitor and Acceptance Theory (MAT). *Clinical Psychology Review*, *51*, 48–59. https://doi.org/10.1016/J.CPR.2016.10.011

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research*, *32*(1), 53–76. https://doi.org/10.1207/s15327906mbr3201_3

Little, T. D., Card, N. A., Slegers, D. W., & Ledford, E. C. (2007). Representing contextual effects in multiple-group MACS models. In *Modeling contextual effects in longitudinal studies.* (pp. 121–147). Lawrence Erlbaum Associates Publishers.

Lodi-Smith, J., & Roberts, B. W. (2016). Beyond the cross-lagged panel model: Next-generation statistical tools for analyzing interdependencies across the life course. *Personality and Social Psychology Review*, *11*(1), 68–86. https://doi.org/10.1177/1088868306294590

Long, J. (1983). *Confirmatory Factor Analysis*. SAGE Publications, Inc. https://doi.org/10.4135/9781412983778

Long, J. S. (1983). *Covariance Structure Models: An Introduction to LISREL*. SAGE Publications, Inc.

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *2*(2), 57–64. https://doi.org/10.1027/1614-2241.2.2.57

Lü, W., Wang, Z., Liu, Y., & Zhang, H. (2014). Resilience as a mediator between extraversion, neuroticism and happiness, PA and NA. *Personality and Individual Differences*, *63*, 128–133. https://doi.org/10.1016/j.paid.2014.01.015

Lüdtke, O., Roberts, B. W., Trautwein, U., & Nagy, G. (2011). A random walk down university avenue: Life paths, life events, and personality trait change at the transition to university life. *Journal of Personality and Social Psychology*, *101*(3), 620–637. https://doi.org/10.1037/a0023743

Lyu, H.-M., Wang, G.-F., Cheng, W.-C., & Shen, S.-L. (2017). Tornado hazards on June 23 in Jiangsu Province, China: Preliminary investigation and analysis. *Natural Hazards*, *85*(1), 597–604. https://doi.org/10.1007/s11069-016-2588-2

Ma, Y., She, Z., Siu, A. F.-Y., Zeng, X., & Liu, X. (2018). Effectiveness of online mindfulness-based interventions on psychological distress and the mediating role of emotion regulation. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02090

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, *51*(1), 201–226. https://doi.org/10.1146/annurev.psych.51.1.201

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Magnusson, D. (1992). Back to the phenomena: Theory, methods, and statistics in psychological research. *European Journal of Personality*, *6*(1), 1–14. https://doi.org/10.1002/per.2410060102

Mahlo, L., & Windsor, T. D. (2020). Older and more mindful? Age differences in mindfulness components and well-being. *Aging & Mental Health*, *0*(0), 1–12. https://doi.org/10.1080/13607863.2020.1734915

Mahlo, L., & Windsor, T. D. (2021). State mindfulness and affective well-being in the daily lives of middle-aged and older adults. *Psychology and Aging*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/pag0000596

Mai, Y., Zhang, Z., & Wen, Z. (2018). Comparing exploratory structural equation modeling and existing approaches for multiple regression with latent variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(5), 737–749. https://doi.org/10.1080/10705511.2018.1444993

Mandal, S. P., Arya, Y. K., & Pandey, R. (2016). Validation of the factor structure of the Five Facet Mindfulness Questionnaire. *Indian Journal of Health & Wellbeing*, *7*(1), 61–66.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 439–476. https://doi.org/10.1080/10705510903008220

Masselink, M., Roekel, E. V., Hankin, B. L., Keijsers, L., Lodder, G. M. A., Vanhalst, J., Verhagen, M., Young, J. F., & Oldehinkel, A. J. (2018). The longitudinal association between self-esteem and depressive symptoms in adolescents: Separating between-person effects from within-person effects. *European Journal of Personality*, *32*(6), 653–671. https://doi.org/10.1002/per.2179

Maydeu-Olivares, A. (2017). Assessing the size of model misfit in structural equation models. *Psychometrika*, *82*(3), 533–558. https://doi.org/10.1007/s11336-016-9552-7

McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, *61*(3), 204–217. https://doi.org/10.1037/0003-066X.61.3.204

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*(1), 81–90.

McCrae, R. R., & Costa, P. T. Jr. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*(5), 509–516. https://doi.org/10.1037/0003-066X.52.5.509

McCrae, R. R., & Löckenhoff, C. E. (2010). Self-regulation and the five-factor model of personality traits. In *Handbook of Personality and Self-Regulation* (pp. 145–168). Wiley-Blackwell. https://doi.org/10.1002/9781444318111.ch7

McDonald, R. P. (1989). An index of goodness-of-fit based on noncentrality. *Journal of Classification*, *6*(1), 97–103. https://doi.org/10.1007/BF01908590

McMahan, D. L. (2008). *The making of Buddhist modernism* (1 edition). Oxford University Press.

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412–433. https://doi.org/10.1037/met0000144

McNeish, D., An, J., & Hancock, G. R. (2018). The thorny relation between measurement quality and fit index cutoffs in latent variable models. *Journal of Personality Assessment*, *100*(1), 43–52. https://doi.org/10.1080/00223891.2017.1281286

Meade, A. W., & Lautenschlager, G. J. (2004a). A monte-carlo study of confirmatory factor analytic tests of measurement equivalence/invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(1), 60–72. https://doi.org/10.1207/S15328007SEM1101_5

Meade, A. W., & Lautenschlager, G. J. (2004b). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, *7*(4), 361–388. https://doi.org/10.1177/1094428104268027

Medvedev, O. N., Siegert, R. J., Feng, X. J., Billington, D. R., Jang, J. Y., & Krägeloh, C. U. (2016). Measuring Trait Mindfulness: How to Improve the Precision of the Mindful Attention Awareness Scale Using a Rasch Model. *Mindfulness*, *7*(2), 384–395. https://doi.org/10.1007/s12671-015-0454-z

Mein Smith, P. (2011). *A concise history of New Zealand* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139196574

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. https://doi.org/10.1007/BF02294825

Merkle, E., & You, D. (2018). *nonnest2: Tests of Non-Nested Models*. https://cran.r-project.org/package=nonnest2

Mesmer-Magnus, J., Manapragada, A., Viswesvaran, C., & Allen, J. W. (2017). Trait mindfulness at work: A meta-analysis of the personal and professional correlates of trait mindfulness. *Human Performance*, *30*(2–3), 79–98. https://doi.org/10.1080/08959285.2017.1307842

Michaelides, M. P., Koutsogiorgi, C., & Panayiotou, G. (2016). Method effects on an adaptation of the Rosenberg Self-Esteem Scale in Greek and the role of personality traits. *Journal of Personality Assessment*, *98*(2), 178–188. https://doi.org/10.1080/00223891.2015.1089248

Michalak, J., Zarbock, G., Drews, M., Otto, D., Mertens, D., Ströhle, G., Schwinger, M., Dahme, B., & Heidenreich, T. (2016). Erfassung von Achtsamkeit mit der Deutschen version des Five Facet Mindfulness Questionnaires (FFMQ-D). *Zeitschrift Für Gesundheitspsychologie*, *24*(1), 1–12. https://doi.org/10.1026/0943-8149/a000149

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-. *International Journal of Psychological Research*, *3*(1), 111–121. https://doi.org/10.1007/s11135-007-9143-x

Minkov, M., Bond, M. H., Dutt, P., Schachner, M., Morales, O., Sanchez, C., Jandosova, J., Khassenbekov, Y., & Mudd, B. (2018). A reconsideration of Hofstede's fifth dimension: New Flexibility versus Monumentalism data from 54 countries. *Cross-Cultural Research*, *52*(3), 309–333. https://doi.org/10.1177/1069397117727488

Minkov, M., Dutt, P., Schachner, M., Morales, O., Sanchez, C., Jandosova, J., Khassenbekov, Y., & Mudd, B. (2017). A revision of Hofstede's individualism-collectivism dimension: A new national index from a 56-country study. *Cross Cultural & Strategic Management*, *24*(3), 386–404. https://doi.org/10.1108/CCSM-11-2016-0197

Mund, M., & Nestler, S. (2019). Beyond the Cross-Lagged Panel Model: Next-generation statistical tools for analyzing interdependencies across the life course. *Advances in Life Course Research*, *41*, 100249. https://doi.org/10.1016/j.alcr.2018.10.002

Muthén, L. K., & Muthén, B. O. (2018). *Mplus User's Guide. Eighth Edition.* Muthén & Muthén.

Ng, S., & Wang, Q. (2021). Measuring mindfulness grounded in the original Buddha's discourses on meditation practice. In A. L. Ai, P. Wink, R. F. Paloutzian, & K. A. Harris (Eds.), *Assessing Spirituality in a Diverse World* (pp. 355–381). Springer International Publishing. https://doi.org/10.1007/978-3-030-52140-0_15

Nilsson, H., & Kazemi, A. (2016). Reconciling and thematizing definitions of mindfulness: The big five of mindfulness. *Review of General Psychology*, *20*(2), 183–193. https://doi.org/10.1037/gpr0000074

Nitzan-Assayag, Y., Aderka, I. M., & Bernstein, A. (2015). Dispositional mindfulness in trauma recovery: Prospective relations and mediating mechanisms. *Journal of Anxiety Disorders*, *36*, 25–32. https://doi.org/10.1016/j.janxdis.2015.07.008

Nye, C. D., & Drasgow, F. (2011). Effect size indices for analyses of measurement equivalence: Understanding the practical importance of differences between groups. *Journal of Applied Psychology*, *96*(5), 966–980. https://doi.org/10.1037/a0022955

Panayiotou, G., Karekla, M., & Panayiotou, M. (2014). Direct and indirect predictors of social anxiety: The role of anxiety sensitivity, behavioral inhibition, experiential avoidance and self-consciousness. *Comprehensive Psychiatry*, *55*(8), 1875–1882. https://doi.org/10.1016/j.comppsych.2014.08.045

Pang, D., & Ruch, W. (2018). Scrutinizing the components of mindfulness: Insights from current, past, and non-meditators. *Mindfulness*, 1–14. https://doi.org/10.1007/s12671-018-0990-4

Park, T., Reilly-Spong, M., & Gross, C. R. (2013). Mindfulness: A systematic review of instruments to measure an emergent patient-reported outcome (PRO). *Quality of Life Research*, *22*(10), 2639–2659. https://doi.org/10.1007/s11136-013-0395-8

Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, *60*(2), 307–317. https://doi.org/10.1037/0022-3514.60.2.307

Paunonen, S. V. (1997). On chance and factor congruence following orthogonal Procrustes rotation. *Educational and Psychological Measurement*, *57*(1), 33–59. https://doi.org/10.1177/0013164497057001003

Pearson, M. R., Lawless, A. K., Brown, D. B., & Bravo, A. J. (2015). Mindfulness and emotional outcomes: Identifying subgroups of college students using latent profile analysis. *Personality and Individual Differences*, *76*, 33–38. https://doi.org/10.1016/j.paid.2014.11.009

Pennebaker, J. W., & Beall, S. K. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. *Journal of Abnormal Psychology*, *95*(3), 274–281. https://doi.org/10.1037/0021-843X.95.3.274

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, *77*(6), 1296–1312. https://doi.org/10.1037/0022-3514.77.6.1296

Pepping, C. A., Davis, P. J., & O'Donovan, A. (2013). Individual differences in attachment and dispositional mindfulness: The mediating role of emotion regulation. *Personality and Individual Differences*, *54*(3), 453–456. https://doi.org/10.1016/J.PAID.2012.10.006

Pepping, C. A., & Duvenage, M. (2016). The origins of individual differences in dispositional mindfulness. *Personality and Individual Differences*, *93*, 130–136. https://doi.org/10.1016/J.PAID.2015.05.027

Pepping, C. A., O'Donovan, A., & Davis, P. J. (2014). The differential relationship between mindfulness and attachment in experienced and inexperienced meditators. *Mindfulness*, *5*(4), 392–399. https://doi.org/10.1007/s12671-012-0193-3

Peters, G.-J. Y. (2018). *{userfriendlyscience}: Quantitative analysis made accessible*. https://doi.org/10.17605/osf.io/txequ

Pirson, M. A., Langer, E. J., Bodner, T., & Zilcha, S. (2012). The development and validation of the Langer Mindfulness Scale—Enabling a socio-cognitive perspective of mindfulness in organizational contexts. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2158921

Pirson, M. A., Langer, E., Zilcha, ·Sigal, & Zilcha, S. (2018). Enabling a socio-cognitive perspective of mindfulness: The development and validation of the Langer Mindfulness Scale. *Journal of Adult Development*, *25*(3), 168–185. https://doi.org/10.1007/s10804-018-9282-4

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Poulin, M., Ministero, L., Gabriel, S., Morrison, C., & Naidu, E. (2021). *Minding your own business? Mindfulness decreases prosocial behavior for those with independent self-construals*. PsyArXiv. https://doi.org/10.31234/osf.io/xhyua

Powers, M. B., & Emmelkamp, P. M. G. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of Anxiety Disorders*, *22*(3), 561–569. https://doi.org/10.1016/j.janxdis.2007.04.006

Pugnaghi, G., Cooper, A., Ettinger, U., & Corr, P. J. (2018). The psychometric properties of the German language Reinforcement Sensitivity Theory-Personality Questionnaire (RST-PQ). *Journal of Individual Differences*, *39*(3), 182–190. https://doi.org/10.1027/1614-0001/a000262

Purser, R. (2019). *McMindfulness: How mindfulness became the new capitalist spirituality*. Repeater.

Purser, R., & Milillo, J. (2015). Mindfulness revisited. *Journal of Management Inquiry*, *24*(1), 3–24. https://doi.org/10.1177/1056492614532315

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Quaglia, J. T., Braun, S. E., Freeman, S. P., McDaniel, M. A., & Brown, K. W. (2016). Meta-analytic evidence for effects of mindfulness training on dimensions of self-reported dispositional mindfulness. *Psychological Assessment*, *28*(7), 803–818. https://doi.org/10.1037/pas0000268

R Core Team. (2020). *R: A language and environment for statistical computing* [Manual]. https://www.R-project.org/

Raad, B. D., & Oudenhoven, J. P. V. (2008). Factors of values in the Dutch Language and their relationship to factors of personality. *European Journal of Personality*, *22*(2), 81–108. https://doi.org/10.1002/per.667

Raad, B. D., & Oudenhoven, J. P. V. (2011). A psycholexical study of virtues in the Dutch language, and relations between virtues and personality. *European Journal of Personality*, *25*(1), 43–52. https://doi.org/10.1002/per.777

Radoń, S. (2014). Validation of the Polish adaptation of the Five Facet Mindfulness Questionnaire. *Roczniki Psychologiczne/Annals of Psychology*, *17*(4), 737–760.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163. https://doi.org/10.2307/271063

Ramos, A., Rosado, A. F. B., Serpa, S. O., Cangas, A. J., Gallego, J., & Ramos, L. (2017). Validity evidence of the Portuguese version of the Five Facet Mindfulness Questionnaire. *Revista de Psicologia Del Deporte*, *27*(2), 87–98.

Ran, G., Zhang, Q., & Huang, H. (2018). Behavioral inhibition system and self-esteem as mediators between shyness and social anxiety. *Psychiatry Research*, *270*, 568–573. https://doi.org/10.1016/j.psychres.2018.10.017

Rasch, G. (1993). *Probabilistic Models for Some Intelligence and Attainment Tests*. MESA Press, 5835 S.

Rasmussen, M. K., & Pidgeon, A. M. (2011). The direct and indirect benefits of dispositional mindfulness on self-esteem and social anxiety. *Anxiety, Stress, and Coping*, *24*(2), 227–233. https://doi.org/10.1080/10615806.2010.515681

Reese, E. D., Zielinski, M. J., & Veilleux, J. C. (2015). Facets of mindfulness mediate behavioral inhibition systems and emotion dysregulation. *Personality and Individual Differences*, *72*, 41–46. https://doi.org/10.1016/j.paid.2014.08.008

Revelle, W. (2018). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University. https://cran.r-project.org/package=psych

Roberts, B. W., & Davis, J. P. (2016). Young adulthood is the crucible of personality development. *Emerging Adulthood*, *4*(5), 318–326. https://doi.org/10.1177/2167696816653052

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Robitzsch, A. (2019). *sirt: Supplementary item response theory models*. https://cran.r-project.org/package=sirt

Roelofs, J., Huibers, M., Peeters, F., & Arntz, A. (2008). Effects of neuroticism on depression and anxiety: Rumination as a possible mediator. *Personality and Individual Differences*, *44*(3), 576–586. https://doi.org/10.1016/j.paid.2007.09.019

Rosenstreich, E., & Ruderman, L. (2017). A dual-process perspective on mindfulness, memory, and consciousness. *Mindfulness*, *8*(2), 505–516. https://doi.org/10.1007/s12671-016-0627-4

Rosseel, Y. (2012). {lavaan}: An {R} package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, *74*(1), 31–57. https://doi.org/10.1177/0013164413498257

Sala, M., Rochefort, C., Lui, P. P., & Baldwin, A. S. (2020). Trait mindfulness and health behaviours: A meta-analysis. *Health Psychology Review*, *14*(3), 345–393. https://doi.org/10.1080/17437199.2019.1650290

Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *ASA 1988 Proceedings of the Business and Economic Statistics Section*, 308–313.

Sauer, S., Walach, H., & Kohls, N. (2011). Gray's Behavioural Inhibition System as a mediator of mindfulness towards well-being. *Personality and Individual Differences*, *50*(4), 506–511. https://doi.org/10.1016/j.paid.2010.11.019

Sauer, S., Walach, H., Schmidt, S., Hinterberger, T., Lynch, S., Büssing, A., & Kohls, N. (2013). Assessment of mindfulness: Review on state of the art. *Mindfulness*, *4*(1), 3–17. https://doi.org/10.1007/s12671-012-0122-5

Schmidt, C., Reyes, G., Barrientos, M., Langer, Á. I., & Sackur, J. (2019). Meditation focused on self-observation of the body impairs metacognitive efficiency. *Consciousness and Cognition*, *70*, 116–125. https://doi.org/10.1016/j.concog.2019.03.001

Schmidt, C., & Vinet, E. V. (2015). Mindfulness: Validation of the Five Facet Mindfulness Questionnaire (FFMQ) in Chilean university students. *Terapia Psicológica*, *33*(2), 93–102. https://doi.org/10.4067/S0718-48082015000200004

Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lönnqvist, J.-E., Demirutku, K., Dirilen-Gumus, O., & Konty, M. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology*, *103*(4), 663–688. https://doi.org/10.1037/a0029393

Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse* (pp. xiv, 351). Guilford Press.

Seli, P., Ralph, B. C. W., Risko, E. F., W Schooler, J., Schacter, D. L., & Smilek, D. (2017). Intentionality and meta-awareness of mind wandering: Are they one and the same, or distinct dimensions? *Psychonomic Bulletin & Review*, *24*(6), 1808–1818. https://doi.org/10.3758/s13423-017-1249-0

Sharf, R. H. (1995). Buddhist modernism and the rhetoric of meditative experience. *Numen*, *42*(3), 228–283.

Siegling, A. B., & Petrides, K. V. (2014). Measures of trait mindfulness: Convergent validity, shared dimensionality, and linkages to the five-factor model. *Frontiers in Psychology*, *5*, 1164. https://doi.org/10.3389/fpsyg.2014.01164

Siegling, A. B., & Petrides, K. V. (2016). Zeroing in on mindfulness facets: Similarities, validity, and dimensionality across three independent measures. *PLOS ONE*, *11*(4), e0153073. https://doi.org/10.1371/journal.pone.0153073

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., Bahník, Š., Bai, F., Bannard, C., Bonnier, E., Carlsson, R., Cheung, F., Christensen, G., Clay, R., Craig, M. A., Dalla

Rosa, A., Dam, L., Evans, M. H., Flores Cervantes, I., … Nosek, B. A. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science*, *1*(3), 337–356. https://doi.org/10.1177/2515245917747646

Slobodskaya, H. R., & Kuznetsova, V. B. (2013). The role of reinforcement sensitivity in the development of childhood personality. *International Journal of Behavioral Development*, *37*(3), 248–256. https://doi.org/10.1177/0165025413475895

Smits, D. J. M., & Boeck, P. D. (2006). From BIS/BAS to the Big Five. *European Journal of Personality*, *20*(4), 255–270. https://doi.org/10.1002/per.583

Solem, S., Thunes, S. S., Hjemdal, O., Hagen, R., & Wells, A. (2015). A metacognitive perspective on mindfulness: An empirical investigation. *BMC Psychology*, *3*(1), 24. https://doi.org/10.1186/s40359-015-0081-4

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Spinhoven, P., Huijbers, M. J., Zheng, Y., Ormel, J., & Speckens, A. E. M. M. (2017). Mindfulness facets and Big Five personality facets in persons with recurrent depression in remission. *Personality and Individual Differences*, *110*, 109–114. https://doi.org/10.1016/J.PAID.2017.01.045

Steiger, J. H. (2016). Notes on the Steiger–Lind (1980) handout. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(6), 777–781. https://doi.org/10.1080/10705511.2016.1217487

Stevenson, J. C., Emerson, L.-M. M., & Millings, A. (2017). The relationship between adult attachment orientation and mindfulness: A systematic review and meta-analysis. *Mindfulness*, *8*(6), 1438–1455. https://doi.org/10.1007/s12671-017-0733-y

Stieger, M., Wepfer, S., Rüegger, D., Kowatsch, T., Roberts, B. W., & Allemand, M. (2020). Becoming more conscientious or more open to experience? Effects of a two-week smartphone-based intervention for personality change. *European Journal of Personality*, *34*(3), 345–366. https://doi.org/10.1002/per.2267

Sun, J. (2014). Mindfulness in context: A historical discourse analysis. *Contemporary Buddhism*, *15*(2), 394–415. https://doi.org/10.1080/14639947.2014.978088

Tabachnick, B. G., & Fidell, L. S. (2007). Using multivariate statistics, 5th ed. In *Using multivariate statistics, 5th ed.* Allyn & Bacon/Pearson Education.

Tellis, G. J. (2017). Interesting and impactful research: On phenomena, theory, and writing. *Journal of the Academy of Marketing Science*, *45*(1), 1–6. https://doi.org/10.1007/s11747-016-0499-0

ten Berge, J. M. F. (1986). Rotation to perfect congruence and the cross validation of component weights across populations. *Multivariate Behavioral Research*, *21*(1), 41–64. https://doi.org/10.1207/s15327906mbr2101_3

Thera, V. N. A. (1998). *Heart of Buddhist meditation: A handbook of mental training based on the Buddha’s way of mindfulness* (2nd ed.). Buddhist Publication Society,Sri Lanka.

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the positive negative affect schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*, 227–242. https://doi.org/10.1177/0022022106297301

Tomlinson, E. R., Yousaf, O., Vittersø, A. D., & Jones, L. (2018). Dispositional mindfulness and psychological health: A systematic review. *Mindfulness*, *9*(1), 23–43. https://doi.org/10.1007/s12671-017-0762-6

Tran, U. S., Glück, T. M., & Nader, I. W. (2013). Investigating the Five Facet Mindfulness Questionnaire (FFMQ): Construction of a short form and evidence of a two-factor higher order structure of mindfulness. *Journal of Clinical Psychology*, *69*(9), 951–965. https://doi.org/10.1002/jclp.21996

Triandis, H. C. (1995). *Individualism and collectivism.* Westview Press.

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology*, *7*, 769. https://doi.org/10.3389/fpsyg.2016.00769

Truong, Q. C., Krägeloh, C. U., Siegert, R. J., Landon, J., & Medvedev, O. N. (2020). Applying Generalizability Theory to Differentiate Between Trait and State in the Five Facet Mindfulness Questionnaire (FFMQ). *Mindfulness*, *11*(4), 953–963. https://doi.org/10.1007/s12671-020-01324-7

Tucker, LJ. (1951). A method for synthesis of factor analysis studies. *Personnel Research Report Dep. Army.*, *984*.

Uz, I. (2015). The index of cultural tightness and looseness among 68 countries. *Journal of Cross-Cultural Psychology*, *46*(3), 319–335. https://doi.org/10.1177/0022022114563611

Valerio, A. (2016). Owning mindfulness: A bibliometric analysis of mindfulness literature trends within and outside of Buddhist contexts. *Contemporary Buddhism*, *17*(1), 157–183. https://doi.org/10.1080/14639947.2016.1162425

Van Dam, N. T., Hobkirk, A. L., Danoff-Burg, S., & Earleywine, M. (2012). Mind your words: Positive and negative items create method effects on the five facet mindfulness questionnaire. *Assessment*, *19*(2), 198–204. https://doi.org/10.1177/1073191112438743

Van Dam, N. T., van Vugt, M. K., Vago, D. R., Schmalzl, L., Saron, C. D., Olendzki, A., Meissner, T., Lazar, S. W., Kerr, C. E., Gorchov, J., Fox, K. C. R., Field, B. A., Britton, W. B., Brefczynski-Lewis, J. A., & Meyer, D. E. (2018). Mind the hype: A critical evaluation and prescriptive agenda for research on mindfulness and meditation. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science*, *13*(1), 36–61. https://doi.org/10.1177/1745691617709589

van de Vijver, F. J. R., & Leung, K. (1997). Methods and data analysis for cross-cultural research. In *Methods and data analysis for cross-cultural research.* Sage Publications, Inc.

Van De Vijver, F. J. R., & Leung, K. (2010). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F. J. R. van de Vijver (Eds.), *Cross-Cultural Research Methods in Psychology* (pp. 17–45). Cambridge University Press. https://doi.org/10.1017/CBO9780511779381.003

Van Doren, N., Zainal, N. H., & Newman, M. G. (2021). Cross-Cultural and gender invariance of emotion regulation in the United States and India. *Journal of Affective Disorders*. https://doi.org/10.1016/j.jad.2021.04.089

Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

Vazsonyi, A. T., Ksinan, A., Mikuška, J., & Jiskrova, G. (2015). The Big Five and adolescent adjustment: An empirical test across six cultures. *Personality and Individual Differences*, *83*, 234–244. https://doi.org/10.1016/j.paid.2015.03.049

Verhaeghen, P. (2018). The mindfulness manifold: Exploring how self-preoccupation, self-compassion, and self-transcendence translate mindfulness into positive psychological outcomes. *Mindfulness*, *10*(1), 1–15. https://doi.org/10.1007/s12671-018-0959-3

Verhaeghen, P. (2020). The examined life is wise living: The relationship between mindfulness, wisdom, and the moral foundations. *Journal of Adult Development*, *27*(4), 305–322. https://doi.org/10.1007/s10804-019-09343-y

Verhaeghen, P., & Aikman, S. N. (2020). How the mindfulness manifold relates to the five moral foundations, prejudice, and awareness of privilege. *Mindfulness*, *11*, 241–254. https://doi.org/10.1007/s12671-019-01243-2

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, *57*(2), 307–333. https://doi.org/10.2307/1912557

Walach, H., Buchheld, N., Buttenmüller, V., Kleinknecht, N., & Schmidt, S. (2006). Measuring mindfulness—The Freiburg Mindfulness Inventory (FMI). *Personality and Individual Differences*, *40*(8), 1543–1555. https://doi.org/10.1016/J.PAID.2005.11.025

Wang, X., Xu, M., Song, Y., Li, X., Zhen, Z., Yang, Z., & Liu, J. (2014). The network property of the thalamus in the default mode network is correlated with trait mindfulness. *Neuroscience*, *278*, 291–301. https://doi.org/10.1016/j.neuroscience.2014.08.006

Way, B. M., Creswell, J. D., Eisenberger, N. I., & Lieberman, M. D. (2010). Dispositional mindfulness and depressive symptomatology: Correlations with limbic and self-referential neural activity during rest. *Emotion (Washington, D.C.)*, *10*(1), 12–24. https://doi.org/10.1037/a0018312

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*(3), 374–385. https://doi.org/10.1016/j.jrp.2008.12.008

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, *8*, 84. https://doi.org/10.2307/270754

White, H. D., & McCain, K. W. (1989). Bibliometrics. *Annual Review of Information Science and Technology*, *24*, 119–186.

Williams, J. M. G., & Kabat-Zinn, J. (2013). *Mindfulness: Diverse perspectives on its meaning, origins and applications*. Routledge.

Williams, M. J., Dalgleish, T., Karl, A., & Kuyken, W. (2014). Examining the factor structures of the Five Facet Mindfulness Questionnaire and the Self-Compassion Scale. *Psychological Assessment*, *26*(2), 407–418. https://doi.org/10.1037/a0035566

Wong, S. Y., Zhang, D., Li, C. C., Yip, B. H., Chan, D. C., Ling, Y., Lo, C. S., Woo, D. M., Sun, Y., Ma, H., Mak, W. W., Gao, T., Lee, T. M., & Wing, Y. (2017). Comparing the effects of mindfulness-based cognitive therapy and sleep psycho-education with exercise on chronic insomnia: A randomised controlled trial. *Psychotherapy and Psychosomatics*, *86*(4), 241–253. https://doi.org/10.1159/000470847

Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review*, *21*(3), 253–277. https://doi.org/10.1177/1088868316652279

Yang, T., Huang, L., & Wu, Z. (2003). [The application of Chinese health questionnaire for mental disorder screening in community settings in mainland China]. *Zhonghua Liu Xing Bing Xue Za Zhi = Zhonghua Liuxingbingxue Zazhi*, *24*(9), 769–773.

Yarkoni, T. (2010). Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality*, *44*(3), 363–373. https://doi.org/10.1016/j.jrp.2010.04.001

Yoon, K. L., Maltby, J., & Joormann, J. (2013). A pathway from neuroticism to depression: Examining the role of emotion regulation. *Anxiety, Stress, & Coping*, *26*(5), 558–572. https://doi.org/10.1080/10615806.2012.734810

Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and likert-type (ordinal) item scores*. Directorate of Human Resources Research and Evaluation, Department of National Defense.

**Appendix A**

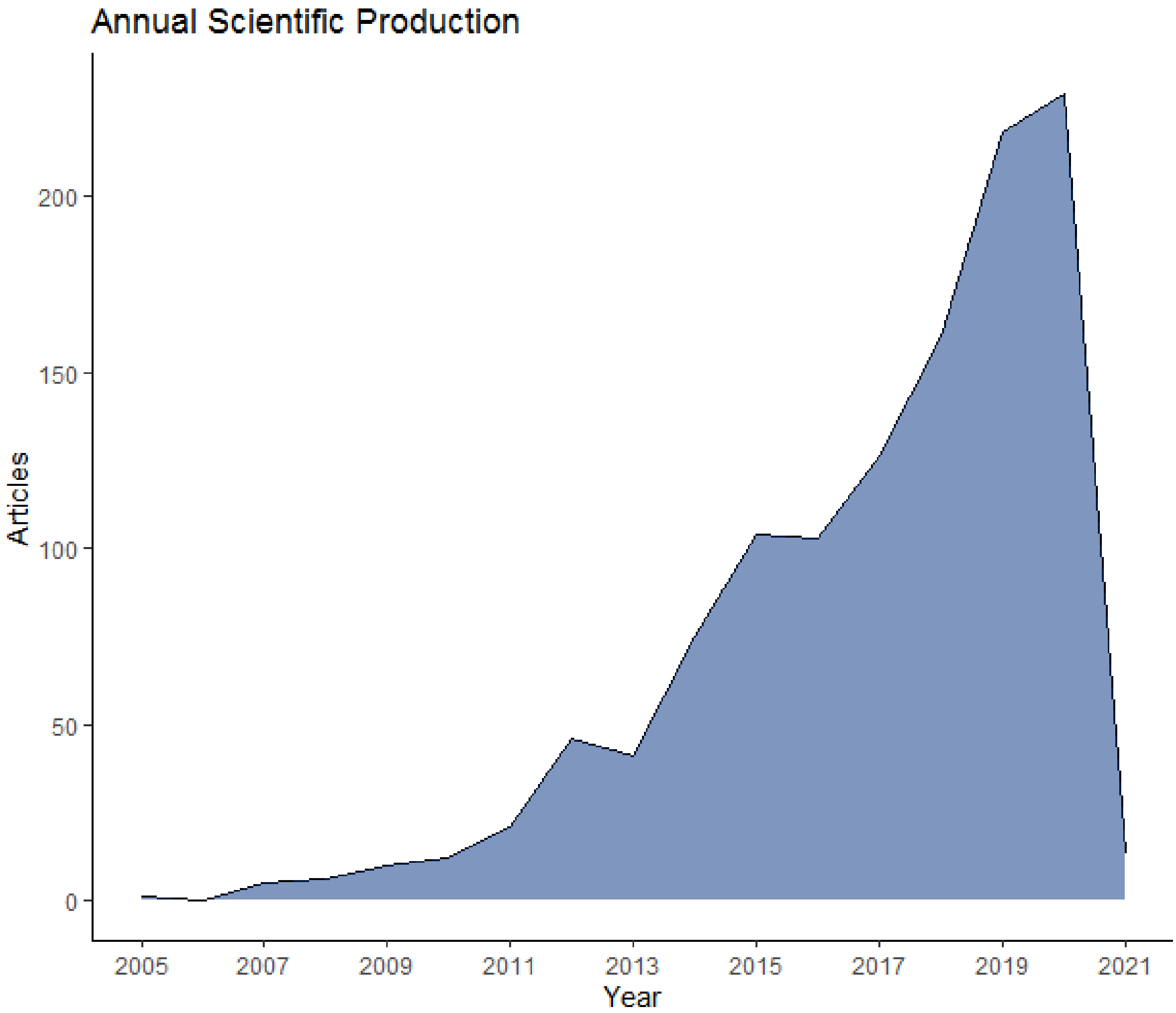
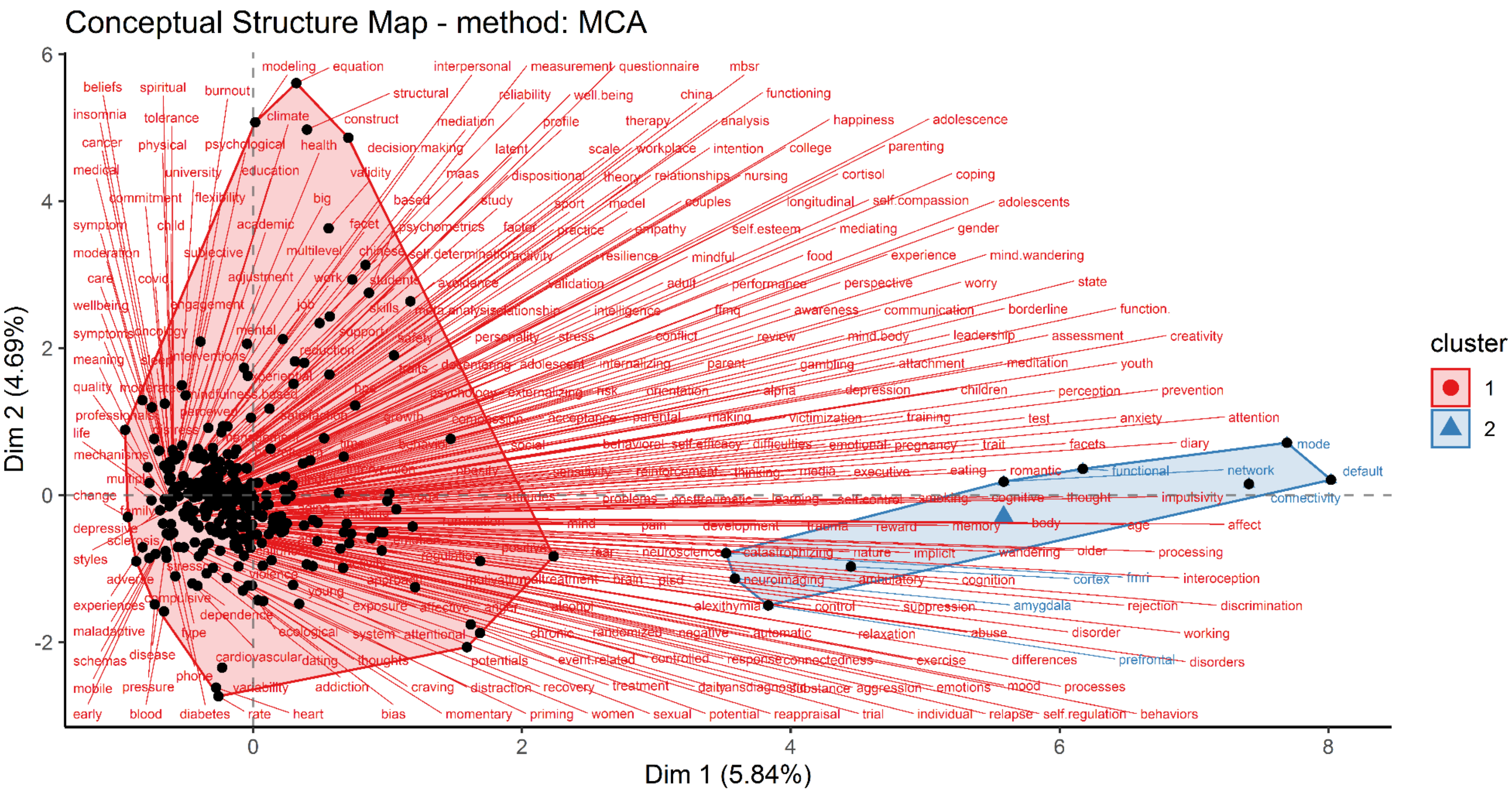



**Figure 1** *Documents per year covered by the search terms*


Conceptual Structure Map - method: MCA
Dim 2 (4.69%)
Dim 1 (5.84%)
6
4
2
0
-2
0
2
4
6
8
cluster
1
2
modeling
equation
interpersonal
measurement
questionnaire
mbsr
beliefs
spiritual
burnout
structural
reliability
well.being
china
functioning
insomnia
tolerance
climate
construct
mediation
profile
therapy
analysis
happiness
adolescence
cancer
physical
psychological
health
decision.making
latent
scale
workplace
intention
college
parenting
medical
university
education
validity
maas
dispositional
theory
relationships
nursing
cortisol
coping
commitment
flexibility
big
based
study
sport
model
couples
longitudinal
self.compassion
adolescents
symptom
child
academic
facet
psychometrics
factor
practice
empathy
self.esteem
mediating
gender
moderation
subjective
multilevel
chinese
self.determination
resilience
mindful
food
experience
mind.wandering
care
covid
adjustment
work
students
avoidance
validation
adult
performance
perspective
worry
state
wellbeing
engagement
job
skills
intelligence
ffmq
awareness
communication
borderline
function
symptoms
oncology
mental
support
personality
stress
conflict
review
mind.body
leadership
assessment
creativity
meaning
interventions
safety
reduction
decentering
adolescent
internalizing
parent
gambling
attachment
meditation
youth
quality
traits
psychology
externalizing
risk
orientation
alpha
depression
children
perception
prevention
professionals
life
growth
compassion
acceptance
parental
making
victimization
training
test
anxiety
attention
mechanisms
social
behavioral
self.efficacy
difficulties
emotional
pregnancy
trait
facets
diary
mode
sensitivity
reinforcement
thinking
media
executive
eating
romantic
functional
network
default
change
attitudes
problems
posttraumatic
learning
self.control
smoking
cognitive
thought
impulsivity
connectivity
depressive
mind
pain
development
trauma
reward
memory
body
age
affect
sclerosis
positive
fear
neuroscience
catastrophizing
nature
implicit
wandering
older
processing
styles
stress
regulation
motivation
maltreatment
brain
ptsd
neuroimaging
ambulatory
cognition
cortex
fmri
interoception
adverse
violence
young
compulsive
exposure
affective
anger
alcohol
alexithymia
control
suppression
amygdala
rejection
discrimination
experiences
dependence
maladaptive
type
ecological
system
attentional
chronic
randomized
negative
automatic
relaxation
abuse
disorder
working
schemas
disease
cardiovascular
dating
thoughts
potentials
event.related
controlled
response
connectedness
exercise
differences
prefrontal
disorders
mobile
pressure
phone
variability
addiction
craving
distraction
recovery
treatment
daily
transdiagnostic
substance
aggression
emotions
mood
processes
early
blood
diabetes
rate
heart
bias
momentary
priming
women
sexual
potential
reappraisal
trial
individual
relapse
self.regulation
behaviors

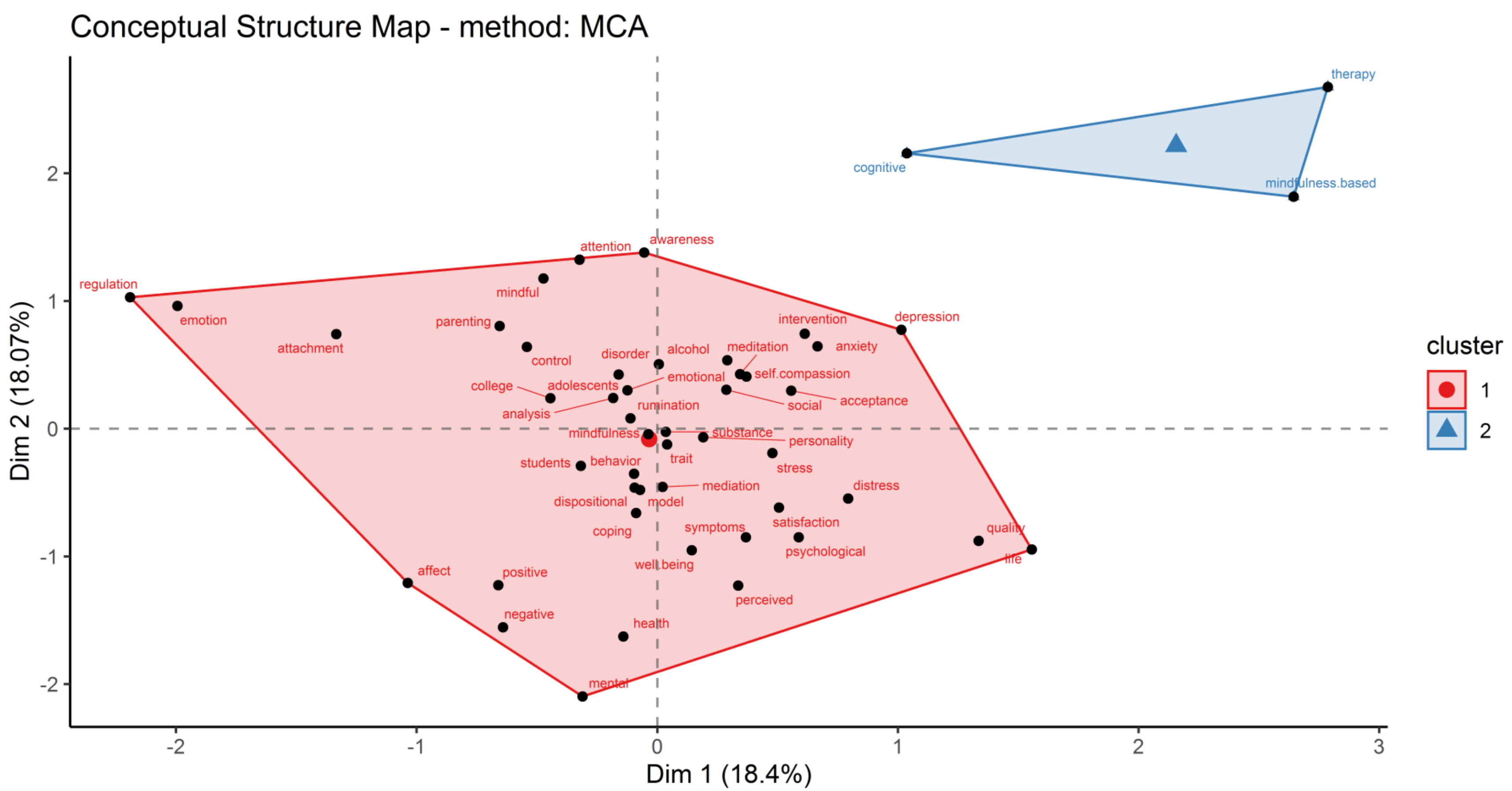



**Figure 2** *MCA clustering of author keywords at 5 and 25 minimal degrees.*

**Table 1**

*Top ten countries by research output*

| Country | Documents | Percent of Total | Single-Country Documents | Multi-Country Documents | MCP to SCP Ratio % |
|---|---|---|---|---|---|
| USA | 511 | 43.16 | 465 | 46 | 9 |

| | | | | | |
|---|---|---|---|---|---|
| China | 108 | 9.12 | 76 | 32 | 29.6 |
| Canada | 93 | 7.85 | 81 | 12 | 12.9 |
| Australia | 75 | 6.33 | 64 | 11 | 14.7 |
| United Kingdom | 64 | 5.41 | 48 | 16 | 25 |
| Netherlands | 39 | 3.29 | 18 | 21 | 53.8 |
| Germany | 36 | 3.04 | 26 | 10 | 27.8 |
| Spain | 36 | 3.04 | 25 | 11 | 30.6 |
| Italy | 34 | 2.87 | 27 | 7 | 20.6 |
| New Zealand | 22 | 1.86 | 13 | 9 | 40.9 |

## Appendix B

### Response factor

We ran a CFA in which each item loaded on one of six factors, based on its highest loading in the principal components analysis. The model yielded poor fit ($\chi^2$(8240) = 10467.298, CFI = .781, RMSEA = .026[.025, .028]), due to the number of items included and the substantial number of cross loadings present. Nevertheless, including a general response factor on which all items were allowed to load freely substantially improved the fit of the model (CFI nevertheless still stayed under recommended cut-offs): $\chi^2$(8110) = 9455.708, CFI = .868, RMSEA = .020[.019, .022]. This indicates that response behaviour to all items accounts for some of the variance in our data.

**STable 1**
Fit of the individual mindfulness scales in previous studies.

| Instrument | Study | Sample size | Number of latent factors | Chi square | df | $X^2$ /df | CFI | SRMR | RMSEA |
|---|---|---|---|---|---|---|---|---|---|
| FMI | (Karatepe & Yavuz, 2019) | 206 | 1 | NA | NA | 2.750 | .890 | NA | .064 |
| FMI | (Bruggeman-Everts, Van der Lee, Van ‘t Hooft, & Nyklíček, 2017) | 158 | 1 | NA | NA | 1.999 | .865 | NA | .080 |
| MAAS | (Cebolla, Luciano, DeMarzo, Navarro-Gil, & Campayo, 2013) | 251 | 1 | 185.43 | NA | NA | .94 | .05 | .07 |
| MAAS | (Osman, Lamis, Bagge, Freedenthal, & Barnes, 2016) | 1,200 | 1 | 338.78 | 90 | NA | .94 | NA | .06 |
| PHLMS | (Cardaciotto, Herbert, Forman, Moitra, & Farrow, 2008) | 280 | 2 | NA | NA | 1.6 | .91 | NA | .05 |
| KIMS | (Baer, Smith, & Allen, 2004) | 215 | 5 | NA | NA | NA | .95 | NA | .07 |
| KIMS | (Baum et al., 2010) | 234 | 5 | 159.03 | 53 | 3.00 | .91 | NA | .09 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| CAMS-R | (Chan, Lo, Lin, & Thompson, 2016) | 215 | 4 | 101.7 | 48 | NA | .89 | NA | .069 |
| LMS | (Pirson, Langer, Zilcha, & Zilcha, 2018) | 2258 | 3 | 536.1 | 72 | NA | .95 | NA | .052 |

We do not claim that this list is exhaustive. One important consideration is that we used item-level analyses, whereas previous studies (e.g., Baer et al., 2004; Baum et al., 2010) also used parcelling, which greatly affects fit statistics.

## Appendix C

Table 1
Mean, SD, and correlation for the full sample at time 1

| | M | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Non-Reacting | 2.85 | 0.72 | | | | | | | | | | | | | | | |
| Non-Judging | 2.87 | 0.78 | .15** | | | | | | | | | | | | | | |
| Observing | 3.55 | 0.74 | .13** | -.18** | | | | | | | | | | | | | |
| Describing | 3.14 | 0.8 | .28** | .25** | .10** | | | | | | | | | | | | |
| Acting with Awareness | 2.99 | 0.69 | 0.03 | .38** | -.07+ | .20** | | | | | | | | | | | |
| Fight Flight Freeze Sensitivity | 2.49 | 0.57 | -.18** | -.20** | 0 | -.17** | -.26** | | | | | | | | | | |
| BIS | 2.74 | 0.55 | -.37** | -.58** | .15** | -.36** | -.41** | .42** | | | | | | | | | |
| BAS-Impulsiveness | 2.6 | 0.57 | 0.03 | -.13** | .19** | 0.03 | -.39** | .17** | .20** | | | | | | | | |
| BAS-Reward Reactivity | 2.88 | 0.49 | .10** | -0.04 | .26** | .17** | -.11** | .23** | .12** | .41** | | | | | | | |
| BAS-Goal Drive Persistence | 2.95 | 0.57 | .13** | 0 | .16** | .28** | .10** | .11** | 0.03 | .10** | .51** | | | | | | |
| BAS-Reward Interest | 2.62 | 0.6 | .21** | .06+ | .25** | .19** | 0 | -0.05 | -.11** | .34** | .47** | .56** | | | | | |
| Agreeableness | 3.7 | 0.66 | -0.01 | .11** | .12** | .13** | .19** | 0.05 | -0.06 | -.10** | .18** | .21** | .18** | | | | |
| Conscientiousness | 3.2 | 0.72 | 0.06 | .12** | .08* | .15** | .37** | 0.01 | -.17** | -.25** | .13** | .52** | .26** | .31** | | | |
| Neuroticism | 3.14 | 0.88 | -.53** | -.45** | -0.01 | -.33** | -.28** | .28** | .72** | 0.05 | -.12** | -.22** | -.33** | -.12** | -.25** | | |
| Openness | 3.63 | 0.67 | .07+ | 0.04 | .36** | .21** | .08* | -.17** | -0.04 | 0.04 | .11** | .17** | .29** | .12** | .07* | -.11** | |
| Extraversion | 3.13 | 0.79 | .11** | .24** | .06+ | .31** | .16** | -.18** | -.37** | .25** | .33** | .35** | .53** | .14** | .17** | -.35** | .18** |

*Notes.* *** $p < .001$, ** $p < .01$, * $p < .05$

Table 2
Mean, SD, and correlation for the full sample at time 2

| | M | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Non-Reacting | 2.83 | 0.72 | | | | | | | | | | | | | | | |
| Non-Judging | 2.74 | 0.78 | .23** | | | | | | | | | | | | | | |
| Observing | 3.49 | 0.79 | 0.06 | -.08+ | | | | | | | | | | | | | |
| Describing | 3.14 | 0.79 | .25** | .21** | .07+ | | | | | | | | | | | | |
| Acting with Awareness | 2.96 | 0.73 | 0.04 | .38** | 0.01 | .26** | | | | | | | | | | | |
| Fight Flight Freeze Sensitivity | 2.34 | 0.61 | -.22** | -.15** | 0 | -.10* | -.13** | | | | | | | | | | |
| BIS | 2.69 | 0.56 | -.43** | -.56** | .15** | -.32** | -.38** | .35** | | | | | | | | | |
| BAS-Impulsiveness | 2.57 | 0.56 | -0.03 | -.16** | .13** | -0.06 | -.34** | .16** | .22** | | | | | | | | |
| BAS-Reward Reactivity | 2.81 | 0.47 | 0.03 | -0.05 | .29** | .14** | 0.02 | .21** | .11** | .39** | | | | | | | |
| BAS-Goal Drive Persistence | 2.83 | 0.54 | .10* | .11** | .13** | .24** | .24** | .15** | -0.04 | 0 | .42** | | | | | | |
| BAS-Reward Interest | 2.56 | 0.56 | .21** | .10* | .24** | .20** | .09* | -0.03 | -.13** | .31** | .46** | .51** | | | | | |
| Agreeableness | 3.71 | 0.68 | 0 | .11** | 0.07 | .09* | .18** | .08+ | -0.05 | -0.04 | .19** | .23** | .19** | | | | |
| Conscientiousness | 3.09 | 0.71 | 0.06 | .20** | -0.02 | .15** | .42** | .09* | -.24** | -.27** | .09* | .45** | .23** | .27** | | | |
| Neuroticism | 3.16 | 0.91 | -.60** | -.44** | 0.03 | -.29** | -.29** | .24** | .75** | .10* | -.07+ | -.19** | -.29** | -.10* | -.29** | | |
| Openness | 3.62 | 0.71 | 0.02 | 0.07 | .40** | .23** | .10* | -.16** | 0.03 | 0.05 | .19** | .18** | .31** | .13** | 0.01 | -0.04 | |
| Extraversion | 3.2 | 0.76 | .16** | .17** | 0.06 | .32** | .18** | -.12** | -.36** | .25** | .36** | .33** | .51** | .13** | .22** | -.40** | .13** |

*Notes.* *** $p < .001$, ** $p < .01$, * $p < .05$

Table 3
Mean, SD, and correlation for the full sample at time 3

| | M | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Non-Reacting | 2.90 | .70 | | | | | | | | | | | | | | | |
| Non-Judging | 2.86 | .81 | .22** | | | | | | | | | | | | | | |
| Observing | 3.56 | .78 | .07 | -.09* | | | | | | | | | | | | | |
| Describing | 3.15 | .77 | .25** | .27** | .10* | | | | | | | | | | | | |
| Acting with Awareness | 3.01 | .69 | .04 | .37** | -.08* | .26** | | | | | | | | | | | |
| Fight Flight Freeze Sensitivity | 2.40 | .59 | -.11** | -.18** | .06 | -.15** | -.10* | | | | | | | | | | |
| BIS | 2.67 | .54 | -.44** | -.57** | .13** | -.31** | -.34** | .38** | | | | | | | | | |
| BAS-Impulsiveness | 2.52 | .55 | -.05 | -.09* | .11* | .03 | -.32** | .16** | .25** | | | | | | | | |
| BAS-Reward Reactivity | 2.84 | .5 | .05 | -.03 | .30** | .20** | -.02 | .26** | .21** | .45** | | | | | | | |
| BAS-Goal Drive Persistence | 2.86 | .58 | .04 | .10* | .25** | .28** | .17** | .12** | .11** | .11** | .47** | | | | | | |
| BAS-Reward Interest | 2.55 | .59 | .11** | .16** | .25** | .20** | .07 | .05 | -.03 | .35** | .50** | .57** | | | | | |
| Agreeableness | 3.69 | .69 | 0 | .10* | .13** | .14** | .20** | .05 | -.03 | 0 | .25** | .25** | .23** | | | | |
| Conscientiousness | 3.20- | .72 | .07 | .16** | .11** | .15** | .34** | .02 | -.15** | -.24** | .14** | .50** | .27** | .39** | | | |
| Neuroticism | 3.09 | .90 | -.56** | -.47** | -.02 | -.33** | -.32** | .22** | .71** | .14** | -.04 | -.16** | -.25** | -.19** | -.30** | | |
| Openness | 3.62 | .73 | 0 | .07 | .38** | .23** | .03 | -.09* | .04 | .15** | .18** | .26** | .34** | .20** | .05 | -.06 | |
| Extraversion | 3.07 | .76 | .06 | .23** | .13** | .33** | .18** | -.15** | -.27** | .26** | .33** | .37** | .56** | .17** | .21** | -.31** | .20** |

*Notes.* *** $p < .001$, ** $p < .01$, * $p < .05$

**Temporal Invariance of the Measures**

Due to our interest in the relationship of these variables over time, we first tested the temporal invariance of each construct across the three waves to examine the suitability of the measures for cross-temporal comparisons. Because past research found that the FFMQ is best modelled with positive and negative methods factors (Aguado et al., 2015; Karl et al., 2020) we tested the fit of this structure over time using a multi-group confirmatory factor analysis with an MLR estimator and time-point as grouping variable. We first examined a configural model, in which factor loadings and intercepts were freely estimated across time. A test of metric invariance constrained the factor loadings to be equal across time. A test of scalar invariance added further equality constraints on the intercepts. These invariance tests were conducted at the facet level for the FFMQ and RST-R-RQ and at the trait level for the BFI-2. Based on the commonly accepted criterion of $\Delta CFI < .01$ (Fischer & Karl, 2019), we found scalar invariance for all but three personality dimensions (FFFS, Reward-Reactivity, and Neuroticism) across the three waves. This implies that the intercepts of items measuring these specific personality facets were not identical across time-points. Overall, our finding implies that we can compare the relationships between all constructs as well as the mean differences for all measures, but these three personality facets.  The average fit of the scalar model across all time points was as follows: CFI = .932 (Min = .901, Max = .970), RMSEA = .083 (Min = .044, Max = .127), SRMR = .054 (Min = .039, Max = .097). We show the fit for all facets and models in Table 4.

Table 4
Equivalence of measures over time.

| | CFI | RMSEA | LC | UC | SRMR | Δ CFI |
|---|---|---|---|---|---|---|
| **Agreeableness** | | | | | | |
| Structural | .926 | .104 | .090 | .119 | .041 | |
| Metric | .927 | .087 | .075 | .100 | .045 | -.001 |
| Scalar | .927 | .077 | .066 | .088 | .047 | .000 |
| **Conscientiousness** | | | | | | |
| Structural | .933 | .089 | .075 | .103 | .039 | |
| Metric | .936 | .074 | .062 | .086 | .041 | -.003 |
| Scalar | .927 | .070 | .059 | .081 | .045 | .009 |
| **Extraversion** | | | | | | |
| Structural | .941 | .105 | .090 | .120 | .043 | |
| Metric | .941 | .088 | .076 | .101 | .046 | .000 |
| Scalar | .941 | .077 | .066 | .089 | .047 | .000 |
| **Neuroticism** | | | | | | |
| Structural | .931 | .127 | .113 | .141 | .040 | |
| Metric | .930 | .110 | .098 | .122 | .047 | .002 |
| Scalar | .920 | .104 | .094 | .115 | .053 | **.010** |
| **Openness** | | | | | | |
| Structural | .956 | .068 | .052 | .083 | .033 | |
| Metric | .954 | .059 | .046 | .072 | .042 | .002 |
| Scalar | .954 | .053 | .041 | .065 | .044 | .001 |
| **Behavioral Inhibition System (BIS)** | | | | | | |
| Structural | .909 | .064 | .061 | .067 | .050 | |
| Metric | .909 | .062 | .059 | .065 | .054 | .000 |
| Scalar | .904 | .062 | .059 | .065 | .056 | .004 |
| **Fight-Flight-Freeze System (FFFS)** | | | | | | |
| Structural | .929 | .064 | .057 | .072 | .040 | |
| Metric | .930 | .059 | .052 | .066 | .044 | -.001 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Scalar | .914 | .061 | .054 | .067 | .049 | **.015** |
| Goal Drive Persistence | | | | | | |
| Structural | .970 | .063 | .053 | .074 | .031 | |
| Metric | .968 | .059 | .050 | .068 | .041 | .002 |
| Scalar | .962 | .059 | .050 | .067 | .044 | .006 |
| Impulsivity | | | | | | |
| Structural | .901 | .087 | .078 | .097 | .049 | |
| Metric | .905 | .077 | .068 | .086 | .050 | -.004 |
| Scalar | .899 | .072 | .064 | .081 | .052 | .007 |
| Reward Interest | | | | | | |
| Structural | .918 | .107 | .096 | .118 | .045 | |
| Metric | .917 | .095 | .085 | .105 | .050 | .002 |
| Scalar | .912 | .089 | .080 | .098 | .053 | .004 |
| Reward Reactivity | | | | | | |
| Structural | .923 | .073 | .066 | .081 | .046 | |
| Metric | .923 | .067 | .060 | .075 | .051 | .001 |
| Scalar | .912 | .067 | .060 | .073 | .054 | **.011** |
| Five-Facet Mindfulness Questionnaire (FFMQ) | | | | | | |
| Structural | .941 | .044 | .040 | .047 | .053 | |
| Metric | .942 | .041 | .037 | .044 | .056 | -.001 |
| Scalar | .938 | .041 | .038 | .044 | .057 | .004 |

*Notes.* All models were fitted with an MLR estimator to adjust for multi-variate non-normality and robust fit indices are presented. Δ CFI above recommended cut-offs (Fischer & Karl, 2019), indicating non-equivalence, are bolded. The FFMQ was modelled as facets subsumed under a higher order factor of mindfulness with positive and negative methods factors. BFI and RST variables were modelled individually.

Table 5. Longitudinal effects on mindfulness facets.

| Acting with Awareness | | | | | |
|---|---|---|---|---|---|
| Predictor | *B* | SE | *p* | CI_low | CI_high |
| Acting with Awareness | -.132 | .143 | .357 | -.413 | .149 |
| Agreeableness | -.111 | .135 | .413 | -.377 | .155 |

| | | | | | |
|---|---|---|---|---|---|
| BIS | -.038 | .214 | .858 | -.458 | .382 |
| Conscientiousness | .155 | .169 | .361 | -.177 | .486 |
| Describing | -.006 | .095 | .953 | -.191 | .180 |
| Extraversion | .036 | .139 | .793 | -.236 | .309 |
| FFFS | .041 | .160 | .795 | -.271 | .354 |
| Goal-Drive Persistence | .369 | .166 | .026 | .044 | .693 |
| Impulsiveness | -.104 | .199 | .600 | -.493 | .285 |
| Neuroticism | -.086 | .156 | .583 | -.392 | .220 |
| Non-Judging | -.030 | .075 | .692 | -.178 | .118 |
| Non-Reacting | -.034 | .105 | .743 | -.239 | .171 |
| Observing | .007 | .091 | .939 | -.171 | .185 |
| Openness | -.050 | .135 | .711 | -.314 | .214 |
| Reward Interest | -.031 | .182 | .865 | -.388 | .326 |
| Reward Reactivity | -.046 | .169 | .786 | -.376 | .284 |
| | | | Describing | | |
| Acting with Awareness | .148 | .113 | .189 | -.073 | .370 |
| Agreeableness | -.247 | .135 | .067 | -.511 | .017 |
| BIS | -.145 | .263 | .580 | -.661 | .370 |
| Conscientiousness | -.200 | .224 | .373 | -.640 | .240 |
| Describing | .161 | .154 | .294 | -.140 | .463 |
| Extraversion | -.135 | .187 | .469 | -.501 | .231 |
| FFFS | -.084 | .211 | .691 | -.497 | .330 |
| Goal-Drive Persistence | .297 | .203 | .143 | -.100 | .695 |
| Impulsiveness | .090 | .224 | .689 | -.350 | .530 |
| Neuroticism | -.021 | .159 | .897 | -.332 | .291 |
| Non-Judging | -.045 | .076 | .554 | -.193 | .104 |
| Non-Reacting | -.023 | .092 | .804 | -.203 | .157 |
| Observing | .163 | .113 | .150 | -.059 | .386 |
| Openness | -.005 | .153 | .976 | -.305 | .296 |
| Reward Interest | -.037 | .190 | .846 | -.410 | .336 |
| Reward Reactivity | -.143 | .196 | .465 | -.527 | .241 |

| | | | | | |
|---|---|---|---|---|---|
| Non-Judging | | | | | |
| Acting with Awareness | -.269 | .112 | .016 | -.488 | -.050 |
| Agreeableness | -.166 | .143 | .244 | -.445 | .113 |
| BIS | -.127 | .246 | .607 | -.609 | .356 |
| Conscientiousness | .250 | .193 | .196 | -.129 | .628 |
| Describing | -.144 | .110 | .189 | -.360 | .071 |
| Extraversion | -.209 | .168 | .212 | -.538 | .119 |
| FFFS | -.138 | .168 | .412 | -.468 | .192 |
| Goal-Drive Persistence | -.005 | .193 | .979 | -.384 | .374 |
| Impulsiveness | .020 | .235 | .933 | -.441 | .480 |
| Neuroticism | -.043 | .162 | .792 | -.360 | .274 |
| Non-Judging | .041 | .119 | .730 | -.192 | .274 |
| Non-Reacting | .050 | .101 | .620 | -.148 | .248 |
| Observing | .009 | .106 | .933 | -.198 | .216 |
| Openness | .100 | .187 | .593 | -.266 | .465 |
| Reward Interest | .140 | .164 | .392 | -.181 | .462 |
| Reward Reactivity | -.053 | .198 | .790 | -.441 | .335 |
| Non-Reacting | | | | | |
| Acting with Awareness | -.079 | .106 | .455 | -.286 | .128 |
| Agreeableness | .080 | .127 | .529 | -.169 | .329 |
| BIS | -.218 | .210 | .299 | -.629 | .194 |
| Conscientiousness | -.177 | .176 | .315 | -.523 | .169 |
| Describing | -.067 | .103 | .514 | -.268 | .134 |
| Extraversion | .044 | .149 | .769 | -.249 | .337 |
| FFFS | .194 | .158 | .221 | -.117 | .505 |
| Goal-Drive Persistence | -.067 | .169 | .691 | -.398 | .264 |
| Impulsiveness | -.087 | .193 | .652 | -.466 | .292 |
| Neuroticism | -.008 | .149 | .960 | -.299 | .284 |
| Non-Judging | .110 | .078 | .158 | -.043 | .262 |
| Non-Reacting | .016 | .116 | .890 | -.210 | .242 |
| Observing | .184 | .101 | .068 | -.014 | .383 |

| | | | | | |
|---|---|---|---|---|---|
| Openness | -.088 | .134 | .511 | -.350 | .174 |
| Reward Interest | .011 | .175 | .950 | -.332 | .353 |
| Reward Reactivity | -.150 | .199 | .449 | -.540 | .239 |
| | | | Observing | | |
| Acting with Awareness | .138 | .115 | .233 | -.088 | .364 |
| Agreeableness | -.062 | .141 | .660 | -.338 | .214 |
| BIS | -.337 | .267 | .208 | -.861 | .187 |
| Conscientiousness | .188 | .201 | .350 | -.206 | .581 |
| Describing | .075 | .121 | .538 | -.163 | .312 |
| Extraversion | .037 | .164 | .821 | -.285 | .359 |
| FFFS | .213 | .169 | .207 | -.118 | .545 |
| Goal-Drive Persistence | .261 | .184 | .157 | -.100 | .623 |
| Impulsiveness | .317 | .231 | .169 | -.135 | .769 |
| Neuroticism | .250 | .159 | .117 | -.062 | .562 |
| Non-Judging | .023 | .078 | .772 | -.130 | .175 |
| Non-Reacting | -.066 | .094 | .485 | -.250 | .119 |
| Observing | .185 | .150 | .215 | -.108 | .479 |
| Openness | -.156 | .177 | .378 | -.504 | .191 |
| Reward Interest | -.176 | .182 | .333 | -.533 | .180 |
| Reward Reactivity | -.185 | .185 | .318 | -.547 | .178 |

**Appendix D**

**Table 1.**
Reliability of the FFMQ facets in the individual samples.

| | Acting with Awareness | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | α | α Low | α High | ω | ω Low | ω High | GLB | H |
| Norway 1 | .860 | .831 | .888 | .861 | .834 | .889 | .908 | .869 |
| India | .810 | .777 | .843 | .811 | .779 | .844 | .873 | .828 |
| Hong Kong 1 | .835 | .809 | .861 | .841 | .816 | .865 | .892 | .870 |
| Portugal | .905 | .887 | .923 | .905 | .887 | .923 | .930 | .918 |
| New Zealand | .877 | .859 | .895 | .878 | .860 | .896 | .906 | .893 |
| Germany | .799 | .772 | .825 | .781 | .752 | .810 | .850 | .883 |
| USA 1 | .902 | .885 | .920 | .902 | .885 | .920 | .945 | .904 |
| USA 2 | .854 | .830 | .878 | .854 | .831 | .878 | .912 | .869 |
| USA 3 | .873 | .849 | .898 | .872 | .848 | .897 | .910 | .891 |
| USA 4 | .850 | .827 | .873 | .848 | .825 | .871 | .883 | .873 |
| USA 5 | .913 | .892 | .934 | .914 | .894 | .934 | .939 | .916 |
| Austria | .794 | .774 | .814 | .791 | .771 | .812 | .851 | .835 |
| Croatia | .825 | .791 | .859 | .817 | .782 | .853 | .870 | .865 |
| Chile | .841 | .816 | .865 | .844 | .821 | .867 | .868 | .854 |
| Romania 1 | .873 | .837 | .910 | .875 | .840 | .911 | .915 | .884 |
| Romania 2 | .870 | .841 | .898 | .872 | .844 | .900 | .902 | .888 |
| Australia | .848 | .813 | .884 | .848 | .812 | .883 | .907 | .880 |
| China | .875 | .850 | .901 | .876 | .851 | .901 | .907 | .896 |
| Poland | .757 | .729 | .784 | .756 | .729 | .784 | .839 | .788 |
| Spain | .891 | .881 | .900 | .892 | .882 | .901 | .917 | .900 |
| Hong Kong 2 | .911 | .891 | .932 | .913 | .894 | .933 | .934 | .919 |
| Sweden | .851 | .831 | .871 | .852 | .832 | .872 | .891 | .854 |
| Norway 2 | .824 | .789 | .858 | .825 | .792 | .859 | .876 | .828 |
| Average | .854 | .829 | .879 | .853 | .829 | .878 | .897 | .874 |
| | Describing | | | | | | | |

| | α | α Low | α High | ω | ω Low | ω High | GLB | H |
|---|---|---|---|---|---|---|---|---|
| Norway 1 | .879 | .855 | .903 | .882 | .859 | .905 | .920 | .906 |
| India | .802 | .768 | .837 | .804 | .770 | .837 | .841 | .824 |
| Hong Kong 1 | .831 | .804 | .857 | .833 | .807 | .859 | .856 | .852 |
| Portugal | .892 | .872 | .913 | .895 | .875 | .914 | .911 | .908 |
| New Zealand | .914 | .901 | .927 | .915 | .902 | .928 | .913 | .925 |
| Germany | .884 | .869 | .899 | .885 | .870 | .900 | .902 | .896 |
| USA 1 | .921 | .906 | .935 | .922 | .908 | .936 | .947 | .930 |
| USA 2 | .885 | .866 | .903 | .886 | .867 | .904 | .916 | .898 |
| USA 3 | .872 | .847 | .896 | .875 | .851 | .898 | .906 | .885 |
| USA 4 | .858 | .836 | .879 | .861 | .841 | .882 | .917 | .871 |
| USA 5 | .926 | .908 | .943 | .925 | .907 | .943 | .960 | .938 |
| Austria | .890 | .879 | .900 | .890 | .880 | .901 | .886 | .893 |
| Croatia | .876 | .852 | .899 | .878 | .855 | .901 | .910 | .888 |
| Chile | .779 | .746 | .811 | .780 | .747 | .813 | .841 | .789 |
| Romania 1 | .884 | .851 | .917 | .888 | .857 | .920 | .904 | .902 |
| Romania 2 | .856 | .823 | .888 | .858 | .827 | .889 | .912 | .889 |
| Australia | .865 | .834 | .896 | .865 | .834 | .896 | .909 | .872 |
| China | .909 | .890 | .927 | .910 | .891 | .928 | .933 | .913 |
| Poland | .688 | .653 | .723 | .679 | .642 | .715 | .777 | .708 |
| Spain | .899 | .891 | .908 | .901 | .893 | .910 | .927 | .911 |
| Hong Kong 2 | .715 | .649 | .781 | .706 | .640 | .773 | .857 | .845 |
| Sweden | .891 | .877 | .906 | .892 | .878 | .906 | .919 | .895 |
| Norway 2 | .886 | .865 | .908 | .888 | .867 | .909 | .916 | .894 |
| Average | .861 | .837 | .885 | .862 | .838 | .886 | .899 | .880 |

Non-Judging

| | α | α Low | α High | ω | ω Low | ω High | GLB | H |
|---|---|---|---|---|---|---|---|---|
| Norway 1 | .917 | .901 | .934 | .918 | .902 | .935 | .925 | .926 |
| India | .773 | .734 | .812 | .775 | .737 | .814 | .798 | .796 |
| Hong Kong 1 | .767 | .731 | .802 | .770 | .735 | .806 | .824 | .790 |
| Portugal | .866 | .841 | .891 | .869 | .845 | .894 | .888 | .882 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| New Zealand | .931 | .921 | .942 | .933 | .923 | .943 | .930 | .937 |
| Germany | .897 | .884 | .910 | .898 | .884 | .911 | .896 | .915 |
| USA 1 | .937 | .926 | .948 | .937 | .926 | .949 | .945 | .948 |
| USA 2 | .869 | .848 | .890 | .870 | .849 | .891 | .900 | .895 |
| USA 3 | .898 | .879 | .918 | .899 | .880 | .919 | .929 | .912 |
| USA 4 | .886 | .869 | .903 | .887 | .871 | .904 | .900 | .906 |
| USA 5 | .932 | .915 | .948 | .933 | .917 | .949 | .913 | .937 |
| Austria | .856 | .842 | .869 | .856 | .842 | .870 | .898 | .868 |
| Croatia | .867 | .841 | .892 | .869 | .844 | .894 | .870 | .887 |
| Chile | .855 | .833 | .876 | .856 | .835 | .877 | .892 | .875 |
| Romania 1 | .844 | .800 | .887 | .845 | .801 | .889 | .903 | .869 |
| Romania 2 | .858 | .826 | .889 | .859 | .828 | .890 | .887 | .866 |
| Australia | .849 | .814 | .884 | .849 | .814 | .884 | .881 | .851 |
| China | .829 | .794 | .864 | .836 | .803 | .869 | .870 | .850 |
| Poland | .787 | .763 | .810 | .788 | .764 | .812 | .785 | .805 |
| Spain | .918 | .911 | .925 | .918 | .911 | .925 | .907 | .925 |
| Hong Kong 2 | .807 | .763 | .851 | .814 | .771 | .856 | .869 | .843 |
| Sweden | .892 | .878 | .907 | .893 | .879 | .907 | .921 | .903 |
| Norway 2 | .878 | .856 | .901 | .880 | .857 | .903 | .900 | .909 |
| Average | .866 | .842 | .889 | .867 | .844 | .891 | .888 | .882 |
| | | | | Non-Reacting | | | | |
| | α | α Low | α High | ω | ω Low | ω High | GLB | H |
| Norway 1 | .795 | .754 | .836 | .799 | .758 | .839 | .851 | .817 |
| India | .607 | .540 | .675 | .614 | .548 | .681 | .646 | .639 |
| Hong Kong 1 | .558 | .489 | .627 | .569 | .502 | .637 | .668 | .612 |
| Portugal | .717 | .663 | .770 | .721 | .669 | .774 | .758 | .736 |
| New Zealand | .844 | .820 | .867 | .845 | .822 | .868 | .871 | .852 |
| Germany | .782 | .753 | .811 | .787 | .760 | .815 | .842 | .810 |
| USA 1 | .881 | .860 | .903 | .883 | .862 | .904 | .917 | .886 |
| USA 2 | .773 | .736 | .810 | .776 | .740 | .812 | .757 | .798 |
| USA 3 | .798 | .759 | .837 | .802 | .765 | .840 | .799 | .826 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| USA 4 | .760 | .725 | .796 | .763 | .727 | .798 | .805 | .780 |
| USA 5 | .882 | .855 | .910 | .883 | .855 | .911 | .920 | .893 |
| Austria | .669 | .637 | .701 | .670 | .638 | .702 | .719 | .678 |
| Croatia | .710 | .653 | .766 | .712 | .657 | .768 | .789 | .748 |
| Chile | .618 | .561 | .676 | .620 | .563 | .677 | .720 | .645 |
| Romania 1 | .703 | .618 | .788 | .709 | .626 | .792 | .791 | .731 |
| Romania 2 | .792 | .746 | .839 | .795 | .749 | .840 | .850 | .800 |
| Australia | .815 | .772 | .858 | .817 | .774 | .859 | .866 | .825 |
| China | .629 | .552 | .706 | .634 | .559 | .710 | .703 | .667 |
| Poland | .583 | .536 | .631 | .587 | .541 | .634 | .639 | .617 |
| Spain | .819 | .803 | .835 | .821 | .805 | .836 | .864 | .833 |
| Hong Kong 2 | .722 | .657 | .786 | .735 | .674 | .795 | .818 | .771 |
| Sweden | .747 | .713 | .781 | .750 | .717 | .784 | .784 | .767 |
| Norway 2 | .709 | .653 | .764 | .714 | .660 | .769 | .737 | .772 |
| Average | .735 | .689 | .781 | .739 | .694 | .785 | .788 | .761 |

| Observing | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | α | α Low | α High | ω | ω Low | ω High | GLB | H |
| Norway 1 | .800 | .761 | .839 | .803 | .764 | .842 | .847 | .832 |
| India | .705 | .655 | .755 | .707 | .657 | .757 | .765 | .719 |
| Hong Kong 1 | .747 | .708 | .786 | .751 | .712 | .789 | .808 | .764 |
| Portugal | .817 | .783 | .851 | .818 | .784 | .852 | .858 | .827 |
| New Zealand | .791 | .760 | .822 | .794 | .764 | .825 | .838 | .812 |
| Germany | .741 | .709 | .774 | .745 | .712 | .777 | .781 | .776 |
| USA 1 | .883 | .862 | .904 | .884 | .864 | .905 | .891 | .893 |
| USA 2 | .696 | .648 | .744 | .697 | .649 | .746 | .768 | .728 |
| USA 3 | .737 | .688 | .786 | .742 | .693 | .791 | .740 | .774 |
| USA 4 | .760 | .724 | .795 | .762 | .727 | .797 | .818 | .777 |
| USA 5 | .879 | .850 | .907 | .879 | .850 | .908 | .910 | .891 |
| Austria | .766 | .744 | .788 | .769 | .747 | .791 | .767 | .789 |
| Croatia | .721 | .670 | .773 | .725 | .673 | .777 | .801 | .758 |
| Chile | .731 | .692 | .770 | .734 | .695 | .774 | .790 | .762 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Romania 1 | .722 | .643 | .801 | .731 | .656 | .806 | .816 | .863 |
| Romania 2 | .794 | .748 | .840 | .799 | .756 | .843 | .834 | .812 |
| Australia | .753 | .698 | .809 | .753 | .696 | .810 | .837 | .792 |
| China | .791 | .749 | .833 | .797 | .756 | .838 | .824 | .835 |
| Poland | .705 | .672 | .738 | .706 | .673 | .739 | .779 | .727 |
| Spain | .841 | .827 | .855 | .845 | .831 | .858 | .848 | .858 |
| Hong Kong 2 | .820 | .779 | .861 | .823 | .782 | .863 | .861 | .839 |
| Sweden | .748 | .715 | .781 | .754 | .721 | .787 | .760 | .790 |
| Norway 2 | .700 | .644 | .755 | .703 | .647 | .760 | .718 | .751 |
| Average | .767 | .727 | .807 | .770 | .731 | .810 | .811 | .799 |

*Notes.* α and ω are given with 95% confidence intervals. GLB = Greatest Lowest Bound, H = Coefficient H.

**Table 2**

*Fit of the correlated five-facet model in all samples (Model 1)*

| | $\chi^2$ | df | χ2/df | CFI | RMSEA | SRMR | BIC | $\hat{\gamma}$ | Fit |
|---|---|---|---|---|---|---|---|---|---|
| Australia | 1066.037 | 692 | 1.541 | .804 | .063[.055, .070] | .087 | 16239.310 | .896 | Poor |
| Austria | 2015.234 | 692 | 2.912 | .863 | .049[.046, .051] | .062 | 103318.107 | .935 | Poor |
| Chile | 1454.520 | 692 | 2.102 | .798 | .056[.052, .060] | .076 | 43284.173 | .911 | Poor |
| China | 1139.102 | 692 | 1.646 | .834 | .059[.053, .066] | .089 | 20785.331 | .904 | Poor |
| Croatia | 1139.410 | 692 | 1.647 | .826 | .057[.051, .063] | .085 | 23609.921 | .913 | Poor |
| Germany | 1617.572 | 692 | 2.338 | .862 | .054[.051, .057] | .072 | 50166.099 | .918 | Poor |
| Hong Kong 1 | 1495.797 | 692 | 2.162 | .760 | .061[.057, .065] | .102 | 34972.663 | .899 | Poor |
| Hong Kong 2 | 1271.140 | 692 | 1.837 | .723 | .079[.072, .086] | .127 | 16197.035 | .851 | Poor |
| India | 1281.745 | 692 | 1.852 | .768 | .056[.052, .061] | .080 | 34224.170 | .908 | Poor |
| New Zealand | 1166.261 | 692 | 1.685 | .925 | .045[.041, .005] | .062 | 38963.832 | .943 | Good |
| Norway 1 | 1262.450 | 692 | 1.824 | .834 | .066[.060, .071] | .102 | 23732.765 | .884 | Poor |
| Norway2 | 1011.725 | 692 | 1.462 | .882 | .047[.040, .053] | .076 | 22896.893 | .937 | Poor |
| Poland | 2078.016 | 692 | 3.003 | .706 | .058[.055, .061] | .069 | 79991.295 | .908 | Poor |
| Portugal | 956.373 | 692 | 1.382 | .915 | .044[.037, .050] | .065 | 24069.981 | .949 | Good |
| Romania 1 | 1031.248 | 692 | 1.490 | .761 | .072[.062, .081] | .120 | 11353.125 | .863 | Poor |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Romania 2 | 1145.370 | 692 | 1.655 | .814 | .066[.059, .072] | .105 | 2019.526 | .887 | Poor |
| Spain | 2548.066 | 692 | 3.682 | .894 | .055[.052, .057] | .061 | 106709.930 | .924 | Poor |
| Sweden | 1557.979 | 692 | 2.251 | .855 | .056[.052, .059] | .072 | 48203.188 | .918 | Poor |
| USA 1 | 138.223 | 692 | 1.995 | .870 | .070[.064, .075] | .078 | 26862.114 | .888 | Poor |
| USA 2 | 1353.844 | 692 | 1.956 | .842 | .057[.052, .061] | .073 | 34595.281 | .910 | Poor |
| USA 3 | 1156.401 | 692 | 1.671 | .848 | .058[.052, .064] | .083 | 24784.349 | .911 | Poor |
| USA 4 | 1601.731 | 692 | 2.315 | .819 | .061[.058, .065] | .078 | 40318.308 | .895 | Poor |
| USA 5 | 1027.263 | 692 | 1.484 | .886 | .064[.055, .072] | .089 | 15041.121 | .901 | Poor |

*Note.* Good fit is assessed as CFI > .90, SRMR <.08, RMSEA <.08, and $\hat{\gamma}$ > .90.

**Table 3**

*Fit of the five-facet model with higher-order factor in all samples (Model 2)*

| | $\chi^2$ | df | χ2/df | CFI | RMSEA | SRMR | BIC | $\hat{\gamma}$ | Fit |
|---|---|---|---|---|---|---|---|---|---|
| Australia | 1081.374 | 697 | 1.551 | .797 | .064[.056, .071] | .100 | 16239.403 | .893 | Poor |
| Austria | 2153.269 | 697 | 3.089 | .849 | .051[.049, .053] | .078 | 103454.430 | .929 | Poor |
| Chile | 1515.631 | 697 | 2.175 | .783 | .058[.054, .062] | .087 | 43322.426 | .905 | Poor |
| China | 120.681 | 697 | 1.723 | .814 | .063[.057, .069] | .111 | 20825.884 | .893 | Poor |
| Croatia | 1168.053 | 697 | 1.676 | .818 | .058[.052, .064] | .092 | 23612.141 | .909 | Poor |
| Germany | 1628.717 | 697 | 2.337 | .861 | .054[.051, .057] | .073 | 50147.725 | .917 | Poor |
| Hong Kong 1 | 156.543 | 697 | 2.239 | .742 | .063[.059, .067] | .110 | 35018.999 | .892 | Poor |
| Hong Kong 2 | 1288.426 | 697 | 1.849 | .717 | .080[.073, .087] | .132 | 16195.825 | .848 | Poor |
| India | | | | | | | | | |
| New Zealand | 1189.317 | 697 | 1.706 | .922 | .046[.041, .050] | .069 | 38958.876 | .940 | Good |
| Norway 1 | 1342.119 | 697 | 1.926 | .812 | .070[.064, .075] | .134 | 23797.827 | .871 | Poor |
| Norway2 | 1034.899 | 697 | 1.485 | .875 | .048[.042, .054] | .085 | 22897.068 | .933 | Poor |
| Poland | 209.282 | 697 | 2.999 | .704 | .058[.055, .060] | .070 | 79972.251 | .908 | Poor |
| Portugal | 976.863 | 697 | 1.402 | .910 | .045[.038, .052] | .078 | 24071.365 | .946 | Good |
| Romania 1 | 1051.356 | 697 | 1.508 | .750 | .073[.064, .082] | .130 | 11354.325 | .858 | Poor |
| Romania 2 | 1236.571 | 697 | 1.774 | .779 | .071[.065, .078] | .146 | 20272.028 | .869 | Poor |
| Spain | 2737.551 | 697 | 3.928 | .883 | .057[.055, .059] | .083 | 106922.635 | .917 | Poor |
| Sweden | 1642.605 | 697 | 2.357 | .841 | .058[.054, .062] | .089 | 48281.106 | .911 | Poor |
| USA 1 | 1444.098 | 697 | 2.072 | .858 | .072[.067, .078] | .111 | 26929.869 | .880 | Poor |
| USA 2 | 1398.729 | 697 | 2.007 | .832 | .058[.054, .063] | .084 | 3462.415 | .905 | Poor |
| USA 3 | 1172.349 | 697 | 1.682 | .844 | .059[.053, .065] | .090 | 24777.556 | .910 | Poor |
| USA 4 | 1695.075 | 697 | 2.432 | .801 | .064[.060, .068] | .098 | 40394.628 | .886 | Poor |
| USA 5 | 1057.980 | 697 | 1.518 | .875 | .066[.058, .074] | .124 | 15077.749 | .894 | Poor |

*Note.* Good fit CFI > .90, SRMR <.08, RMSEA <.08, and $\hat{\gamma}$ > .90.

**Table 4**

*Fit of the correlated five-facet model with uncorrelated methods factors in all samples (Model 3a)*

| | $\chi^2$ | df | χ2/df | CFI | RMSEA | SRMR | BIC | $\hat{\gamma}$ | Fit |
|---|---|---|---|---|---|---|---|---|---|
| Australia | 847.342 | 653 | 1.298 | .899 | .046[.037,.055] | .073 | 16166.657 | .943 | Poor |
| Austria | 1388.402 | 653 | 2.126 | .924 | .038[.035,.040] | .055 | 102833.227 | .963 | Good |
| Chile | 1073.780 | 653 | 1.644 | .888 | .043[.038,.047] | .064 | 43092.839 | .949 | Poor |
| China | 865.065 | 653 | 1.325 | .923 | .042[.034,.049] | .079 | 20652.210 | .952 | Good |
| Croatia | 840.536 | 653 | 1.287 | .929 | .038[.030,.045] | .076 | 23438.717 | .962 | Good |
| Germany | 1114.613 | 653 | 1.707 | .932 | .039[.035,.043] | .053 | 49807.239 | .957 | Good |
| Hong Kong 1 | 1111.212 | 653 | 1.702 | .862 | .048[.043,.052] | .082 | 34765.514 | .940 | Poor |
| Hong Kong 2 | 1002.905 | 653 | 1.536 | .836 | .063[.055,.070] | .108 | 16034.245 | .904 | Poor |
| India | 941.087 | 653 | 1.441 | .887 | .041[.035,.046] | .070 | 34062.651 | .953 | Poor |
| New Zealand | 897.697 | 653 | 1.375 | .961 | .033[.028,.039] | .054 | 38874.514 | .970 | Good |
| Norway 1 | 896.109 | 653 | 1.372 | .929 | .044[.037,.051] | .087 | 23511.392 | .947 | Poor |
| Norway2 | 773.658 | 653 | 1.185 | .955 | .029[.020,.037] | .072 | 22841.430 | .975 | Good |
| Poland | 1183.308 | 653 | 1.812 | .886 | .037[.034,.040] | .045 | 79220.804 | .963 | Poor |
| Portugal | 742.429 | 653 | 1.137 | .972 | .026[.014,.035] | .062 | 23977.315 | .982 | Good |
| Romania 1 | 886.570 | 653 | 1.358 | .840 | .060[.050,.070] | .114 | 11340.670 | .902 | Poor |
| Romania 2 | 889.100 | 653 | 1.362 | .903 | .049[.040,.057] | .079 | 20093.797 | .938 | Good |
| Spain | 1613.793 | 653 | 2.471 | .946 | .040[.038,.042] | .049 | 105733.782 | .959 | Good |
| Sweden | 1094.279 | 653 | 1.676 | .926 | .041[.037,.045] | .057 | 47882.094 | .956 | Good |
| USA 1 | 902.207 | 653 | 1.382 | .954 | .043[.036,.049] | .057 | 26408.132 | .956 | Good |
| USA 2 | 938.635 | 653 | 1.437 | .932 | .038[.033,.043] | .062 | 34335.032 | .959 | Good |
| USA 3 | 833.406 | 653 | 1.276 | .943 | .037[.029,.044] | .064 | 24568.835 | .964 | Good |
| USA 4 | 992.079 | 653 | 1.519 | .932 | .039[.034,.043] | .059 | 39859.413 | .958 | Good |
| USA 5 | 827.018 | 653 | 1.266 | .942 | .047[.036,.056] | .074 | 14953.421 | .946 | Good |

*Note.* Good fit is assessed as CFI > .90, SRMR <.08, RMSEA <.08, and $\hat{\gamma}$ > .90.

**Table 5**

*Fit of the five-facet model with higher-order factor and uncorrelated methods factors in all samples (Model 3b)*

| | $\chi^2$ | df | χ2/df | CFI | RMSEA | SRMR | BIC | $\hat{\gamma}$ | Fit |
|---|---|---|---|---|---|---|---|---|---|
| Australia | 863.401 | 658 | 1.312 | .893 | .048[.038, .056] | .077 | 16161.693 | .940 | Poor |
| Austria | 1404.632 | 658 | 2.135 | .922 | .038[.035, .040] | .058 | 102817.776 | .962 | Good |
| Chile | 1078.240 | 658 | 1.639 | .888 | .043[.038, .047] | .065 | 43067.739 | .949 | Poor |
| China | 876.412 | 658 | 1.332 | .921 | .042[.034, .049] | .084 | 20637.827 | .950 | Poor |
| Croatia | 85.675 | 658 | 1.293 | .927 | .038[.030, .045] | .075 | 23421.280 | .961 | Good |
| Germany | 1126.162 | 658 | 1.711 | .931 | .039[.035, .043] | .053 | 49789.072 | .957 | Good |
| Hong Kong 1 | 1122.898 | 658 | 1.707 | .861 | .048[.043, .052] | .079 | 3474.750 | .939 | Poor |
| Hong Kong 2 | | | | | | | | | |
| India | | | | | | | | | |
| New Zealand | 91.841 | 658 | 1.384 | .960 | .034[.028, .039] | .058 | 38859.691 | .969 | Good |
| Norway 1 | 898.340 | 658 | 1.365 | .930 | .044[.036, .051] | .087 | 23486.444 | .948 | Poor |
| Norway2 | 784.256 | 658 | 1.192 | .953 | .030[.021, .038] | .074 | 22825.098 | .974 | Good |
| Poland | 1194.081 | 658 | 1.815 | .885 | .037[.034, .040] | .046 | 79198.709 | .962 | Poor |
| Portugal | 754.523 | 658 | 1.147 | .970 | .027[.016, .035] | .063 | 23964.162 | .981 | Good |
| Romania 1 | 892.290 | 658 | 1.356 | .837 | .061[.050, .070] | .109 | 11337.296 | .902 | Poor |
| Romania 2 | 905.491 | 658 | 1.376 | .898 | .050[.042, .058] | .082 | 20086.420 | .935 | Poor |
| Spain | 1681.194 | 658 | 2.555 | .943 | .041[.039, .044] | .056 | 105783.298 | .957 | Good |
| Sweden | 1112.486 | 658 | 1.691 | .924 | .041[.037, .045] | .060 | 47872.465 | .955 | Good |
| USA 1 | 92.545 | 658 | 1.399 | .951 | .044[.037, .050] | .065 | 26402.898 | .954 | Good |
| USA 2 | | | | | | | | | |
| USA 3 | 843.467 | 658 | 1.282 | .941 | .037[.029, .044] | .066 | 24552.970 | .963 | Good |
| USA 4 | 1022.460 | 658 | 1.554 | .927 | .040[.035, .045] | .064 | 39864.212 | .955 | Good |
| USA 5 | 837.260 | 658 | 1.272 | .940 | .047[.037, .056] | .075 | 14938.844 | .944 | Good |

*Note.* Good fit CFI > .90, SRMR <.08, RMSEA <.08, and $\hat{\gamma}$ > .90.

**Table 6**

*Fit of the five-facet model with higher-order factor and correlated methods factors in all samples (Model 4a)*

| | $\chi^2$ | df | χ2/df | CFI | RMSEA | SRMR | BIC | $\hat{\gamma}$ | Fit |
|---|---|---|---|---|---|---|---|---|---|
| Australia | 862.152 | 657 | 1.312 | .893 | .048[.038, .056] | .077 | 16166.066 | .940 | Poor |
| Austria | 138.550 | 657 | 2.101 | .925 | .037[.034, .040] | .053 | 102796.410 | .963 | Good |
| Chile | 1071.225 | 657 | 1.630 | .89 | .042[.038, .047] | .064 | 43065.210 | .949 | Poor |
| China | | | | | | | | | |
| Croatia | 849.929 | 657 | 1.294 | .927 | .038[.030, .045] | .076 | 23425.610 | .961 | Good |
| Germany | 1124.103 | 657 | 1.711 | .932 | .039[.035, .043] | .053 | 49793.511 | .957 | Good |
| Hong Kong 1 | 1032.957 | 657 | 1.572 | .887 | .043[.038, .048] | .061 | 34648.941 | .950 | Poor |
| Hong Kong 2 | | | | | | | | | |
| India | | | | | | | | | |
| New Zealand | | | | | | | | | |
| Norway 1 | 898.477 | 657 | 1.368 | .930 | .044[.036, .051] | .087 | 23491.818 | .948 | Poor |
| Norway2 | 78.572 | 657 | 1.188 | .954 | .030[.020, .038] | .073 | 22827.432 | .974 | Good |
| Poland | 1193.827 | 657 | 1.817 | .885 | .037[.034, .040] | .046 | 79205.114 | .962 | Poor |
| Portugal | 747.052 | 657 | 1.137 | .972 | .026[.014, .035] | .059 | 23961.743 | .982 | Good |
| Romania 1 | | | | | | | | | |
| Romania 2 | 904.734 | 657 | 1.377 | .898 | .050[.042, .058] | .063 | 20095.623 | .935 | Poor |
| Spain | 1676.866 | 657 | 2.552 | .943 | .041[.039, .044] | .056 | 105788.159 | .957 | Good |
| Sweden | 1112.413 | 657 | 1.693 | .924 | .041[.037, .046] | .060 | 47878.465 | .955 | Good |
| USA 1 | 902.648 | 657 | 1.374 | .954 | .042[.035, .049] | .057 | 26385.420 | .957 | Good |
| USA 2 | | | | | | | | | |
| USA 3 | 836.839 | 657 | 1.274 | .943 | .037[.029, .044] | .066 | 24551.420 | .964 | Good |
| USA 4 | 1022.408 | 657 | 1.556 | .927 | .040[.035, .045] | .065 | 39869.647 | .955 | Good |
| USA 5 | 829.959 | 657 | 1.263 | .942 | .046[.036, .056] | .070 | 14938.478 | .946 | Good |

*Note.* Good fit CFI > .90, SRMR <.08, RMSEA <.08, and $\hat{\gamma}$ > .9.

**Table 7**

*Fit of the five-facet model with higher-order factor and correlated methods factors and correlation with the higher-order factor in all samples (Model 4b)*

| | $\chi^2$ | df | χ2/df | CFI | RMSEA | SRMR | BIC | $\hat{\gamma}$ | Fit |
|---|---|---|---|---|---|---|---|---|---|
| Australia | 859.449 | 655 | 1.312 | .894 | .048[.038, .056] | .072 | 1617.071 | .940 | Poor |
| Austria | 1374.974 | 655 | 2.099 | .925 | .037[.034, .040] | .055 | 102802.007 | .963 | Good |
| Chile | | | | | | | | | |
| China | 836.509 | 655 | 1.277 | .934 | .039[.030, .046] | .063 | 20606.615 | .959 | Good |
| Croatia | 819.351 | 655 | 1.251 | .938 | .035[.027, .043] | .061 | 23399.769 | .966 | Good |
| Germany | | | | | | | | | |
| Hong Kong 1 | 99.480 | 655 | 1.512 | .900 | .041[.035, .046] | .059 | 34608.679 | .955 | Poor |
| Hong Kong 2 | | | | | | | | | |
| India | | | | | | | | | |
| New Zealand | 91.498 | 655 | 1.390 | .960 | .034[.029, .039] | .056 | 38873.585 | .968 | Good |
| Norway 1 | 878.051 | 655 | 1.341 | .935 | .042[.035, .049] | .069 | 2348.800 | .951 | Good |
| Norway2 | 757.962 | 655 | 1.157 | .962 | .027[.017, .036] | .061 | 22817.984 | .979 | Good |
| Poland | 1187.497 | 655 | 1.813 | .886 | .037[.034, .040] | .045 | 79209.213 | .963 | Poor |
| Portugal | 746.272 | 655 | 1.139 | .972 | .026[.014, .035] | .058 | 23971.099 | .982 | Good |
| Romania 1 | 858.459 | 655 | 1.311 | .860 | .056[.045, .066] | .077 | 11302.090 | .913 | Poor |
| Romania 2 | | | | | | | | | |
| Spain | 1617.994 | 655 | 2.470 | .946 | .040[.038, .043] | .047 | 105736.160 | .959 | Good |
| Sweden | 1099.970 | 655 | 1.679 | .925 | .041[.037, .045] | .059 | 47881.443 | .956 | Good |
| USA 1 | | | | | | | | | |
| USA 2 | | | | | | | | | |
| USA 3 | 83.891 | 655 | 1.269 | .944 | .036[.028, .044] | .061 | 24553.485 | .964 | Good |
| USA 4 | | | | | | | | | |
| USA 5 | 821.172 | 655 | 1.254 | .945 | .045[.035, .055] | .070 | 14937.002 | .948 | Good |

*Note.* Good fit CFI > .90, SRMR <.08, RMSEA <.08, and $\hat{\gamma}$ > .9.

**Table 8**
Comparison of Models using pre-registered criteria, supplemented by Vuong's test of non-nested models.

| sample | ΔCFI | $\Delta\hat{\gamma}$ | ΔBIC | Fit Improved? | $\omega^2$ | *p* | Disting-uishable? | LRT | Model1 *p* | Model2 *p* | Which Model fits best (Vuong's Test) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Comparison in Model fit between Model 2 and Model 1 | | | | | | | | | | | |
| Australia | -.007 | -.003 | .093 | No | .289 | .000 | Yes | 1.857 | .032 | .968 | Model1 |
| Austria | -.014 | -.006 | 136.323 | No | .292 | .000 | Yes | 5.064 | .000 | 1.000 | Model1 |
| Chile | -.015 | -.006 | 38.253 | No | .377 | .000 | Yes | 2.783 | .003 | .997 | Model1 |
| China | -.020 | -.011 | 40.553 | No | .387 | .000 | Yes | 3.693 | .000 | 1.000 | Model1 |
| Croatia | -.008 | -.004 | 2.220 | No | .169 | .000 | Yes | 2.317 | .010 | .990 | Model1 |
| Germany | -.001 | -.001 | -18.374 | No | .026 | .070 | No | 1.760 | .039 | .961 | Model1 |
| Hong Kong 1 | -.018 | -.007 | 46.336 | No | .371 | .000 | Yes | 3.252 | .001 | .999 | Model1 |
| Hong Kong 2 | -.006 | -.003 | -1.210 | No | .293 | .043 | Yes | 1.737 | .041 | .959 | Model1 |
| India | Did not converge | | | | | | | | | | |
| New Zealand | -.003 | -.003 | -4.956 | No | .087 | .000 | Yes | 2.120 | .017 | .983 | Model1 |
| Norway 1 | -.022 | -.013 | 65.062 | No | .641 | .000 | Yes | 3.845 | .000 | 1.000 | Model1 |
| Norway 2 | -.007 | -.004 | .175 | No | .153 | .000 | Yes | 2.272 | .012 | .988 | Model1 |
| Poland | -.002 | .000 | -19.044 | No | .031 | .037 | Yes | 1.483 | .069 | .931 | Undetermined |
| Portugal | -.005 | -.003 | 1.384 | No | .172 | .000 | Yes | 2.204 | .014 | .986 | Model1 |
| Romania 1 | -.011 | -.005 | 1.200 | No | .341 | .017 | Yes | 2.016 | .022 | .978 | Model1 |
| Romania 2 | -.035 | -.018 | 81.502 | No | .725 | .000 | Yes | 4.669 | .000 | 1.000 | Model1 |
| Spain | -.011 | -.007 | 212.705 | No | .309 | .000 | Yes | 6.558 | .000 | 1.000 | Model1 |
| Sweden | -.014 | -.007 | 77.918 | No | .379 | .000 | Yes | 3.975 | .000 | 1.000 | Model1 |
| USA 1 | -.012 | -.008 | 67.755 | No | .630 | .000 | Yes | 3.610 | .000 | 1.000 | Model1 |
| USA 2 | -.010 | -.005 | 25.134 | No | .205 | .000 | Yes | 3.236 | .001 | .999 | Model1 |
| USA 3 | -.004 | -.001 | -6.793 | No | .165 | .015 | Yes | 1.627 | .052 | .948 | Undetermined |
| USA 4 | -.018 | -.009 | 76.320 | No | .489 | .000 | Yes | 3.809 | .000 | 1.000 | Model1 |
| USA 5 | -.011 | -.007 | 36.628 | No | 1.797 | .000 | Yes | 1.848 | .032 | .968 | Model1 |
| Comparison in Model fit between Model 3a and Model 2 | | | | | | | | | | | |
| Australia | .102 | .050 | -72.746 | Yes | 3.867 | 0 | Yes | -5.887 | 1 | 0 | Model3a |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Austria | .075 | .034 | -621.203 | Yes | 1.756 | 0 | Yes | -11.175 | 1 | 0 | Model3a |
| Chile | .105 | .044 | -229.587 | Yes | 2.042 | 0 | Yes | -8.646 | 1 | 0 | Model3a |
| China | .109 | .059 | -173.674 | Yes | 4.39 | 0 | Yes | -6.672 | 1 | 0 | Model3a |
| Croatia | .111 | .053 | -173.424 | Yes | 4.884 | 0 | Yes | -6.035 | 1 | 0 | Model3a |
| Germany | .071 | .040 | -340.486 | Yes | 2.591 | 0 | Yes | -8.324 | 1 | 0 | Model3a |
| Hong Kong 1 | .120 | .048 | -253.485 | Yes | 2.935 | 0 | Yes | -7.82 | 1 | 0 | Model3a |
| Hong Kong 2 | .119 | .056 | -161.580 | Yes | 5.603 | 0 | Yes | -6.293 | 1 | 0 | Model3a |
| India | | | | | | Did not converge | | | | | |
| New Zealand | .039 | .030 | -84.362 | Yes | 1.798 | 0 | Yes | -6.493 | 1 | 0 | Model3a |
| Norway 1 | .117 | .076 | -286.435 | Yes | 5.642 | 0 | Yes | -7.377 | 1 | 0 | Model3a |
| Norway 2 | .080 | .042 | -55.638 | Yes | 1.943 | 0 | Yes | -6.852 | 1 | 0 | Model3a |
| Poland | .182 | .055 | -751.447 | Yes | 2.947 | 0 | Yes | -11.43 | 1 | 0 | Model3a |
| Portugal | .062 | .036 | -94.050 | Yes | 4.907 | 0 | Yes | -4.803 | 1 | 0 | Model3a |
| Romania 1 | .090 | .044 | -13.655 | Yes | 6.493 | 0 | Yes | -4.125 | 1 | 0 | Model3a |
| Romania 2 | .124 | .069 | -178.231 | Yes | 4.48 | 0 | Yes | -7.115 | 1 | 0 | Model3a |
| Spain | .063 | .042 | -1188.853 | Yes | 3.614 | 0 | Yes | -11.602 | 1 | 0 | Model3a |
| Sweden | .085 | .045 | -399.012 | Yes | 3.004 | 0 | Yes | -8.713 | 1 | 0 | Model3a |
| USA 1 | .096 | .076 | -521.737 | Yes | 13.058 | 0 | Yes | -6.364 | 1 | 0 | Model3a |
| USA 2 | .100 | .054 | -285.383 | Yes | 3.695 | 0 | Yes | -7.615 | 1 | 0 | Model3a |
| USA 3 | .099 | .054 | -208.721 | Yes | 8.098 | 0 | Yes | -5.06 | 1 | 0 | Model3a |
| USA 4 | .131 | .072 | -535.215 | Yes | 4.477 | 0 | Yes | -9.46 | 1 | 0 | Model3a |
| USA 5 | .067 | .052 | -124.328 | Yes | 9.495 | 0 | Yes | -4.502 | 1 | 0 | Model3a |
| Comparison in Model fit between Model 3a and Model 3b | | | | | | | | | | | |
| Australia | -.006 | -.003 | -4.964 | No | .251 | .006 | Yes | 1.598 | .055 | .945 | Undetermined |
| Austria | -.002 | -.001 | -15.451 | No | .038 | .000 | Yes | 1.560 | .059 | .941 | Undetermined |
| Chile | .000 | .000 | -25.100 | No | .021 | .346 | No | .845 | .199 | .801 | Undetermined |
| China | -.002 | -.002 | -14.383 | No | .132 | .020 | Yes | 1.173 | .120 | .880 | Undetermined |
| Croatia | -.002 | -.001 | -17.437 | No | .046 | .230 | No | 1.500 | .067 | .933 | Undetermined |
| Germany | -.001 | .000 | -18.167 | No | .148 | .000 | Yes | .746 | .228 | .772 | Undetermined |
| Hong Kong 1 | -.001 | -.001 | -24.764 | No | 1.289 | .000 | Yes | .109 | .456 | .544 | Undetermined |
| Hong Kong 2 | | | | | | Did not converge | | | | | |
| India | | | | | | Did not converge | | | | | |
| New Zealand | -.001 | -.001 | -14.823 | No | .045 | .056 | No | 1.789 | .037 | .963 | Model3a |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Norway 1 | .001 | .001 | -24.948 | No | .013 | .849 | No | .617 | .269 | .731 | Undetermined |
| Norway 2 | -.002 | -.001 | -16.332 | No | .055 | .044 | Yes | 1.520 | .064 | .936 | Undetermined |
| Poland | -.001 | -.001 | -22.095 | No | .019 | .075 | No | 1.462 | .072 | .928 | Undetermined |
| Portugal | -.002 | -.001 | -13.153 | No | .077 | .018 | Yes | 1.642 | .050 | .950 | Undetermined |
| Romania 1 | -.003 | .000 | -3.374 | No | 2.293 | .002 | Yes | .634 | .263 | .737 | Undetermined |
| Romania 2 | -.005 | -.003 | -7.377 | No | 1.272 | .013 | Yes | .612 | .270 | .730 | Undetermined |
| Spain | -.003 | -.002 | 49.516 | No | .180 | .000 | Yes | 2.940 | .002 | .998 | Model3a |
| Sweden | -.002 | -.001 | -9.629 | No | .054 | .009 | Yes | 2.075 | .019 | .981 | Model3a |
| USA 1 | -.003 | -.002 | -5.234 | No | .138 | .003 | Yes | 1.845 | .032 | .968 | Model3a |
| USA 2 | Did not converge | | | | | | | | | | |
| USA 3 | -.002 | -.001 | -15.865 | No | .075 | .075 | No | 1.361 | .087 | .913 | Undetermined |
| USA 4 | -.005 | -.003 | 4.799 | No | .124 | .000 | Yes | 2.468 | .007 | .993 | Model3a |
| USA 5 | -.002 | -.002 | -14.577 | No | .179 | .294 | No | 1.010 | .156 | .844 | Undetermined |
| Comparison in Model fit between Model 3b and Model 1 | | | | | | | | | | | |
| Australia | .089 | .044 | -77.617 | Yes | 3.640 | .000 | Yes | -5.648 | 1.000 | .000 | Model3b |
| Austria | .059 | .027 | -500.331 | Yes | 1.708 | .000 | Yes | -11.100 | 1.000 | .000 | Model3b |
| Chile | .090 | .038 | -216.434 | Yes | 2.049 | .000 | Yes | -8.547 | 1.000 | .000 | Model3b |
| China | .087 | .046 | -147.504 | Yes | 4.717 | .000 | Yes | -6.241 | 1.000 | .000 | Model3b |
| Croatia | .101 | .048 | -188.641 | Yes | 4.725 | .000 | Yes | -5.987 | 1.000 | .000 | Model3b |
| Germany | .069 | .039 | -377.027 | Yes | 2.587 | .000 | Yes | -8.154 | 1.000 | .000 | Model3b |
| Hong Kong 1 | .101 | .040 | -231.913 | Yes | 3.227 | .000 | Yes | -7.388 | 1.000 | .000 | Model3b |
| Hong Kong 2 | Did not converge | | | | | | | | | | |
| India | Did not converge | | | | | | | | | | |
| New Zealand | .035 | .026 | -104.141 | Yes | 1.723 | .000 | Yes | -6.345 | 1.000 | .000 | Model3b |
| Norway 1 | .096 | .064 | -246.321 | Yes | 5.663 | .000 | Yes | -7.334 | 1.000 | .000 | Model3b |
| Norway 2 | .071 | .037 | -71.795 | Yes | 1.871 | .000 | Yes | -6.722 | 1.000 | .000 | Model3b |
| Poland | .179 | .054 | -792.586 | Yes | 2.875 | .000 | Yes | -11.454 | 1.000 | .000 | Model3b |
| Portugal | .055 | .032 | -105.819 | Yes | 4.671 | .000 | Yes | -4.712 | 1.000 | .000 | Model3b |
| Romania 1 | .076 | .039 | -15.829 | Yes | 5.439 | .000 | Yes | -4.096 | 1.000 | .000 | Model3b |
| Romania 2 | .084 | .048 | -104.106 | Yes | 4.810 | .000 | Yes | -6.552 | 1.000 | .000 | Model3b |
| Spain | .049 | .033 | -926.632 | Yes | 3.261 | .000 | Yes | -11.524 | 1.000 | .000 | Model3b |
| Sweden | .069 | .037 | -330.723 | Yes | 3.028 | .000 | Yes | -8.403 | 1.000 | .000 | Model3b |
| USA 1 | .081 | .066 | -459.216 | Yes | 12.973 | .000 | Yes | -6.195 | 1.000 | .000 | Model3b |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| USA 2 | | | | | Did not converge | | | | | | |
| USA 3 | .093 | .052 | -231.379 | Yes | 8.100 | .000 | Yes | -4.929 | 1.000 | .000 | Model3b |
| USA 4 | .108 | .060 | -454.096 | Yes | 4.448 | .000 | Yes | -9.078 | 1.000 | .000 | Model3b |
| USA 5 | .054 | .043 | -102.277 | Yes | 9.638 | .000 | Yes | -4.331 | 1.000 | .000 | Model3b |
| Comparison in Model fit between Model 3b and Model 4a | | | | | | | | | | | |
| Australia | .000 | .000 | 4.373 | No | .010 | .349 | No | -.280 | .610 | .390 | Undetermined |
| Austria | .003 | .001 | -21.366 | No | .373 | .000 | Yes | -.742 | .771 | .229 | Undetermined |
| Chile | .002 | .000 | -2.529 | No | .036 | .002 | Yes | -1.122 | .869 | .131 | Undetermined |
| China | | | | | Did not converge | | | | | | |
| Croatia | .000 | .000 | 4.330 | No | .009 | .287 | No | -.384 | .650 | .350 | Undetermined |
| Germany | .001 | .000 | 4.439 | No | .121 | .000 | Yes | -.114 | .546 | .454 | Undetermined |
| Hong Kong 1 | .026 | .011 | -91.809 | Yes | 1.654 | .000 | Yes | -1.983 | .976 | .024 | Model4a |
| Hong Kong 2 | | | | | Did not converge | | | | | | |
| India | | | | | Did not converge | | | | | | |
| New Zealand | | | | | Did not converge | | | | | | |
| Norway 1 | .000 | .000 | 5.374 | No | .000 | .835 | No | -.068 | .527 | .473 | Undetermined |
| Norway 2 | .001 | .000 | 2.334 | No | .031 | .050 | No | -.573 | .717 | .283 | Undetermined |
| Poland | .000 | .000 | 6.405 | No | .000 | .693 | No | -.156 | .562 | .438 | Undetermined |
| Portugal | .002 | .001 | -2.419 | No | .109 | .011 | Yes | -.759 | .776 | .224 | Undetermined |
| Romania 1 | | | | | Did not converge | | | | | | |
| Romania 2 | .000 | .000 | 9.203 | No | 1.564 | .000 | Yes | .118 | .453 | .547 | Undetermined |
| Spain | .000 | .000 | 4.861 | No | .011 | .006 | Yes | -.312 | .622 | .378 | Undetermined |
| Sweden | .000 | .000 | 6.000 | No | .001 | .656 | No | -.149 | .559 | .441 | Undetermined |
| USA 1 | .003 | .003 | -17.478 | No | .144 | .011 | Yes | -1.822 | .966 | .034 | Model4a |
| USA 2 | | | | | Did not converge | | | | | | |
| USA 3 | .002 | .001 | -1.550 | No | .066 | .072 | No | -.875 | .809 | .191 | Undetermined |
| USA 4 | .000 | .000 | 5.435 | No | .003 | .499 | No | -.264 | .604 | .396 | Undetermined |
| USA 5 | .002 | .002 | -.366 | No | .158 | .015 | Yes | -.545 | .707 | .293 | Undetermined |
| Comparison in Model fit between Model 4a and Model 4b | | | | | | | | | | | |
| Australia | .001 | .000 | 4.005 | No | .225 | .048 | Yes | -.510 | .695 | .305 | Undetermined |
| Austria | .000 | .000 | 5.597 | No | .299 | .000 | Yes | -.239 | .595 | .405 | Undetermined |
| Chile | | | | | Did not converge | | | | | | |
| China | | | | | Did not converge | | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Croatia | .011 | .005 | -25.841 | Yes | .394 | .000 | Yes | -1.885 | .970 | .030 | Model4b |
| Germany | | | | | | Did not converge | | | | | |
| Hong Kong 1 | .012 | .005 | -40.262 | Improved | 1.257 | .000 | Yes | -1.212 | .887 | .113 | Undetermined |
| Hong Kong 2 | | | | | | Did not converge | | | | | |
| India | | | | | | Did not converge | | | | | |
| New Zealand | | | | | | Did not converge | | | | | |
| Norway 1 | .005 | .003 | -11.018 | No | .231 | .004 | Yes | -1.519 | .936 | .064 | Undetermined |
| Norway 2 | .008 | .005 | -9.448 | No | .284 | .003 | Yes | -1.232 | .891 | .109 | Undetermined |
| Poland | .001 | .001 | 4.099 | No | .023 | .027 | Yes | -1.133 | .871 | .129 | Undetermined |
| Portugal | .000 | .000 | 9.356 | No | .023 | .460 | No | -.352 | .638 | .362 | Undetermined |
| Romania 1 | | | | | | Did not converge | | | | | |
| Romania 2 | | | | | | Did not converge | | | | | |
| Spain | .003 | .002 | -51.999 | No | .119 | .000 | Yes | -2.814 | .998 | .002 | Model4b |
| Sweden | .001 | .001 | 2.978 | No | .133 | .000 | Yes | -.581 | .719 | .281 | Undetermined |
| USA 1 | | | | | | Did not converge | | | | | |
| USA 2 | | | | | | Did not converge | | | | | |
| USA 3 | .001 | .000 | 2.065 | No | .106 | .026 | Yes | -.877 | .810 | .190 | Undetermined |
| USA 4 | | | | | | Did not converge | | | | | |
| USA 5 | .003 | .002 | -1.476 | No | .937 | .043 | Yes | -.479 | .684 | .316 | Undetermined |

**Table 9**

*Equivalence of the individual FFMQ facets across all countries.*

| | $\chi^2$ | df | χ2/df | CFI | RMSEA | SRMR | $\hat{\gamma}$ | ΔCFI | $\Delta\hat{\gamma}$ | Equivalence |
|---|---|---|---|---|---|---|---|---|---|---|
| Non-Judging | | | | | | | | | | |
| Structural | 359.322 | 180 | 1.996 | .976 | .061[.051, .070] | .033 | .986 | | | Yes |
| Metric | 573.700 | 244 | 2.351 | .960 | .068[.061, .075] | .113 | .975 | -.017 | -.11 | No |
| Scalar | 1333.076 | 300 | 4.444 | .886 | .103[.097, .109] | .133 | .926 | -.074 | -.049 | No |
| Non-Reacting | | | | | | | | | | |
| Structural | 221.483 | 112 | 1.978 | .968 | .058[.047, .069] | .035 | .990 | | | Yes |
| Metric | 326.392 | 161 | 2.027 | .954 | .058[.049, .067] | .076 | .985 | -.014 | -.005 | No |
| Scalar | 873.322 | 203 | 4.302 | .831 | .098[.092, .105] | .102 | .943 | -.123 | -.042 | No |
| Observing | | | | | | | | | | |
| Structural | 682.084 | 280 | 2.436 | .959 | .058[.052, .063] | .036 | .986 | | | Yes |
| Metric | 975.959 | 384 | 2.542 | .943 | .058[.054, .063] | .069 | .980 | -.016 | -.006 | No |
| Scalar | 2549.695 | 475 | 5.368 | .817 | .094[.090, .098] | .096 | .934 | -.126 | -.046 | No |

*Note.* We did not test the equivalence of the Acting with Awareness Facet and the Describing facet, as they only showed good fit in a single sample.

**Table 10**

*Results of cross-cultural principal component analysis (varimax rotation).*

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| I tell myself that I shouldn't be thinking the way I'm thinking. | .78 | | | | | |
| I think some of my emotions are poor or inappropriate and I shouldn't feel them. | .78 | | | | | |
| I believe some of my thoughts are abnormal or poor and I shouldn't think that way. | .72 | | | | | |
| I make judgments about whether my thoughts are good or poor. | .69 | | | | | |
| I disapprove of myself when I have irrational ideas. | .69 | | | | | |
| I tell myself I shouldn't be feeling the way I'm feeling. | .69 | | | | | |
| When I have distressing thoughts or images, I judge myself as good or poor, depending what the thought/image is about. | .68 | | | | | |
| I criticize myself for having irrational or inappropriate emotions. | .68 | | | | | |
| I'm good at finding words to describe my feelings. | | .77 | | | | |
| I can usually describe how I feel at the moment in considerable detail. | | .74 | | | | |
| I can easily put my beliefs, opinions, and expectations into words. | | .72 | | | | |
| It's hard for me to find the words to describe what I'm thinking. | | .72 | | | | |

| | | | |
|---|---|---|---|
| I have trouble thinking of the right words to express how I feel about things | .71 | | |
| Even when I’m feeling terribly upset, I can find a way to put it into words. | .68 | | |
| My natural tendency is to put my experiences into words. | .68 | | |
| When I have a sensation in my body, it’s difficult for me to describe it because I can’t find the right words. | .60 | | |
| I pay attention to sensations, such as the wind in my hair or sun on my face. | | .74 | |
| I pay attention to sounds, such as clocks ticking, birds chirping, or cars passing. | | .66 | |
| I notice visual elements in art or nature, such as colors, shapes, textures, or patterns of light and shadow. | | .64 | |
| I notice the smells and aromas of things. | | .63 | |
| When I take a shower or bath, I stay alert to the sensations of water on my body. | | .62 | |
| When I’m walking, I deliberately notice the sensations of my body moving. | | .58 | |
| I notice how foods and drinks affect my thoughts, bodily sensations, and emotions. | | .54 | |
| I pay attention to how my emotions affect my thoughts and behavior. | | .43 | |
| When I have distressing thoughts or images I am able just to notice them without reacting. | | | .70 |

| | | | |
|---|---|---|---|
| When I have distressing thoughts or images, I just notice them and let them go. | .67 | | |
| When I have distressing thoughts or images, I "step back" and am aware of the thought or image without getting taken over by it. | .64 | | |
| When I have distressing thoughts or images, I feel calm soon after. | .63 | | |
| I watch my feelings without getting lost in them. | .56 | | |
| I perceive my feelings and emotions without having to react to them. | .55 | | |
| In difficult situations, I can pause without immediately reacting. | .53 | | |
| I do jobs or tasks automatically without being aware of what I'm doing. | | .78 | |
| It seems I am "running on automatic" without much awareness of what I'm doing. | | .72 | |
| I find myself doing things without paying attention. | | .70 | |
| I rush through activities without being really attentive to them. | | .66 | |
| When I do things, my mind wanders off and I'm easily distracted. | | | .83 |
| I am easily distracted. | | | .82 |
| I don't pay attention to what I'm doing because I'm daydreaming, worrying, or otherwise distracted. | | | .67 |
| I find it difficult to stay focused on what's happening in the present. | | | .60 |

*Note:* We report component loadings higher than .40.

**Table 11**

Congruence coefficients (Tucker's φ) of the individual countries towards the perfect five-factor matrix

| | Describing | Observing | Acting with Awareness | Non-Judging | Non-Reacting | Average |
|---|---|---|---|---|---|---|
| Australia | .910 | .880 | .930 | .940 | .900 | .912 |
| Austria | .960 | .910 | .930 | .950 | .930 | .936 |
| Chile | .930 | .890 | .940 | .930 | .890 | .916 |
| China | .930 | .890 | .910 | .910 | .880 | .904 |
| Germany | .950 | .940 | .890 | .920 | .910 | .922 |
| Spain | .950 | .930 | .940 | .940 | .920 | .936 |
| Hong Kong | .920 | .890 | .910 | .880 | .810 | .882 |
| Croatia | .940 | .830 | .910 | .920 | .850 | .890 |
| India | .900 | .870 | .820 | .860 | .800 | .850 |
| Norway | .950 | .880 | .930 | .930 | .910 | .920 |
| New Zealand | .950 | .950 | .950 | .960 | .960 | .954 |
| Poland | .900 | .940 | .820 | .900 | .870 | .886 |
| Portugal | .930 | .920 | .950 | .950 | .890 | .928 |
| Romania | .900 | .880 | .870 | .910 | .890 | .890 |
| Sweden | .950 | .910 | .940 | .940 | .940 | .936 |
| USA | .950 | .920 | .930 | .940 | .950 | .938 |
| Average | .933 | .902 | .911 | .924 | .894 | .913 |

**Table 12**
Congruence coefficients (Tucker's φ) of the individual countries towards the pooled six-factor matrix

| | 1 | 2 | 3 | 4 | 5 | 6 | Average |
|---|---|---|---|---|---|---|---|
| Australia | .980 | .960 | .950 | .960 | .890 | .920 | .943 |
| Austria | .990 | .990 | .980 | .980 | .970 | .970 | .980 |
| Chile | .970 | .980 | .970 | .950 | .890 | .870 | .938 |
| China | .970 | .970 | .970 | .950 | .840 | .890 | .932 |
| Germany | .980 | .990 | .970 | .970 | .950 | .960 | .970 |
| Spain | .990 | 1 | .980 | .980 | .920 | .940 | .968 |
| Hong Kong | .940 | .970 | .970 | .870 | .770 | .780 | .883 |
| Croatia | .980 | .970 | .980 | .960 | .970 | .940 | .967 |
| India | .950 | .970 | .920 | .860 | .790 | .820 | .885 |
| Norway | .990 | .990 | .980 | .980 | .970 | .970 | .980 |
| New Zealand | .980 | .990 | .980 | .980 | .940 | .950 | .970 |
| Poland | .970 | .970 | .970 | .920 | .850 | .910 | .932 |
| Portugal | .970 | .980 | .970 | .950 | .860 | .770 | .917 |
| Romania | .970 | .970 | .970 | .970 | .840 | .840 | .927 |
| Sweden | .990 | .990 | .980 | .980 | .930 | .900 | .962 |
| USA | 1 | 1 | .990 | .990 | .980 | .990 | .992 |
| Average | .976 | .981 | .971 | .953 | .898 | .901 | .947 |