Chapter to appear in O. N. Medvedev, C. U. Krägeloh, R. J. Siegert & N. N. Singh (Eds.) *Handbook of Assessment in Mindfulness Research.* New York: Springer.

Idionomic Assessment of Mindfulness

Cristóbal Hernández

Escuela de Psicología, Universidad Adolfo Ibáñez.

Avenida Diagonal Las Torres 2640, Peñalolén, Santiago, Chile.

Baljinder Kaur Sahdra

Institute for Positive Psychology and Education, Australian Catholic University.

33 Berry Street North Sydney NSW 2060 Australia

Joseph Ciarrochi

Institute for Positive Psychology and Education, Australian Catholic University.

33 Berry Street North Sydney NSW 2060 Australia

Steven C. Hayes (corresponding author)

University of Nevada, Reno.

1664 N. Virginia Street, Reno, NV 89557, United States of America.

stevenchayes@gmail.com

Abstract

The traditional approach to assessing mindfulness has largely focused on between-person variability, often overlooking the unique, idiographic patterns that emerge within individuals over time. Idionomic assessment emphasizes the importance of modeling processes and their dynamic interactions in particular people (or couples, families, and so one) before generating nomothetic extensions. By critically examining the assumptions underlying conventional psychometric methods, particularly the ergodic theorem, we highlight the limitations of applying group-level findings to particular people. We propose an idionomic approach that integrates intensive longitudinal data to capture the nuanced and individualized nature of mindfulness and its impact on well-being. This method allows for a more precise understanding of how mindfulness practices influence personal outcomes, acknowledging the significant heterogeneity in these effects. Through a series of empirical examples, we demonstrate the practical applications of idionomic assessment in potentially identifying personalized intervention strategies that are more effective than traditional "one size fits all" approaches. The data we provide show that mindfulness processes are highly individualized, and their effects on psychological outcomes varies significantly between persons. This has profound implications for the field of applied psychology, suggesting a shift towards more personalized, process-based, and contextually sensitive assessments and intervention strategies. Ultimately, the idionomic approach offers a promising avenue for enhancing the treatment utility of mindfulness assessments, paving the way for more tailored and effective therapeutic practices.

Keywords

Idionomic assessment, Psychometrics, Individualized interventions, Ergodicity, Longitudinal data, Process-based therapy, Psychological flexibility

Introduction

Mindfulness is a process. What is the purpose of assessing mindfulness in applied psychology and how can we best fulfill it? There are undoubtedly multiple purposes, of course, but Paul Meehl stated the bottom line for clinical assessment quite plainly: "in what way and to what extent does this . . . information help us in treating the patient" (1959, p. 117). He called this question "ultimately the practically significant one by which the contributions of our [assessment] techniques must be judged" (1959, p. 116). This core question is central to the applied assessment of mindfulness.

The question of purpose is a very different place to start in assessment work and theory. Usually, the question is assumed or deferred. We typically begin with a concept, define what “it” is, and then determine how to assess “it”. This produces a long list of supposed psychometrically demanded requirements. Arguments in the field break out about the proper definitions of “it” and a variety of “its” are soon being pursued and compared one to another. Competing groups argue about whether requirements have been met, hoping their definition of “it” will be accepted as the correct one. None of this, however, is yet touching the real purpose, as Meehl stated it.

The degree to which assessment is shown to help improve clinical outcomes is known as “treatment utility” (Hayes, et al., 1987). Although there continues to be wide agreement that this is generally the ultimate purpose of clinical assessment (Hunsley & Mash, 2007), little real progress has been made in consistently demonstrating that clinical assessment yields better intervention outcomes. Indeed, recent large and well-controlled studies with extensive follow-ups have continued to fail to show any benefit for extended assessment approaches above and beyond such low-level processes as an intake interview (Olsson & Friedel, 2018) and reviews that have focused on possible future directions in treatment utility research do not seem on their face to fully explain half a century of failure (e.g., Kamphuis, et al., 2021). This lack of progress applies to the assessment of mindfulness as well: so far as we can ascertain, there have been no experimental demonstrations of the treatment utility of any mindfulness measures.

Note that treatment utility is usually thought to stand at the end of the line of reliability and validity verification questions. Like a person running the gauntlet who arrives at the end exhausted, questions of the practical contributions of our measures are put off to another day. Most often, these questions are never asked, and when they are, the answers are not salutary.

This psychometric gauntlet is thought to be necessary because traditional psychometric theory rests on a key assumption. The various reliability and validity tests depend on the consistency of responses *between* people in a collection of respondents. People who score above average on one mindfulness item tend to score above average on other mindfulness items. That consistency allows a "true" score for, say, "mindfulness" as a concept, to be sought – hidden beneath an error-filled approximation by test or item scores provided by particular people. If a person tends to score about average on most mindfulness items but deviates and scores below average on some items, that deviation is assumed to be error. But what if that assumption itself is wrong?

The assumption that between-person variability in a collection of people is a good guide to within-person variability for particular people in that collective has been periodically called into question throughout the long history of assessment (Cattell, 1952). It was, for example, part of disagreement about when to use different forms of factor analysis, with one side looking for consistencies between people (in what came to be called the R-technique) and others for consistencies within people or the P-technique (e.g., Cattell, et al., 1947). Complaints about the "ecological fallacy" (Estes, 1956) of interpreting results for population samples as if they apply to particular people over time and situation have been made. Ways of analyzing within-person changes of particular people over time have been described (Lamiell, 1981). Some developmental psychologists (e.g., Nesselroade, 2001), behavioral psychologists (Sidman, 1953) and a smattering of others have objected. In the main, however, assessment research has continued to rely on between-person variability in constructing and assessing its concepts.

Recently, however, the challenges to this assumption have grown (e.g., Nesselroade, et al., 2007). Some of the factors involved are technological, such as the emergence of the smartphones as a data collection device, making longitudinal data collection more possible and practical (Burke, et al., 2017).

But in recent times perhaps one of the most profound challenges and changes has come from the recognition that the mathematical and scientific validity of applying assessment information from a collection of people to particular individuals is based on a mathematical assumption that is almost impossible to meet: ergodicity (Molenaar, 2004).

The ergodic theorem is a long-established concept from statistical physics. It was first introduced by Ludwig Boltzmann (1984) in the late 19th century to describe the behavior of systems over long periods and was mathematically proven in the early 1930s (Birkhoff, 1931). The theorem provides the conditions under which spatial and temporal samples converge. In behavioral science, this theorem has recently gained attention as researchers have recognized its implications for longitudinal data analysis and for psychometrics (Molenaar, 2004; 2015). As a result, the relevance of ergodicity to behavioral science has been discussed with increasing frequency (e.g., Bringmann et al., 2022; Wang & Maxwell, 2015) as its implications have become more fully understood.

Traditionally, psychologists and clinicians have relied on group-level data to infer the likelihood of individual outcomes over time. For example, a person’s trait mindfulness scores are examined to determine their health behaviors over time, and some predictive values are generated (Sala, et al., 2019). The ergodic theorem suggests, however, that such generalizations are only mathematically valid, even probabilistically, when applied to a particular person (your score is X and there your likely outcome over time is Y) if the phenomenon in question is ergodic.

The criteria for ergodicity are somewhat similar to the homogeneity assumption with which most students of statistics and measurement are familiar, but the requirements are more severe. Phenomena are ergodic if the measured processes are stationary, and the same dynamic model applies to all individuals (Gates, et al., 2023).

There are phenomena in nature that are ergodic, and if those conditions are met, the behavior of a collection of events is known to represent the behavior of the particular elements in that collection. Some noble gasses are like that for example, if you have a volume of the right kind of gas you can place sensors of gas activity anywhere in the volume (i.e., take spatial samples) or watch these sensors over time (i.e., take temporal samples) and the same means and standard deviations of activity will eventually result. From there, you are able to derive mathematically that the behavior of the collections and conditions that impact that behavior (pressure, heat, and so on) impact the behavior of individual molecules in a similar fashion. But what if that were *not* true? What if, say, heat permanently changed specific molecules, and these changed molecules but not others now reacted to pressure entirely differently over time, even interacting with other molecules in new or different ways. Heating up one side of the volume of gas, say, would produce rogue molecules mixed in with others because of that history. In that case, the empirical tests for ergodicity will fail and generalizations from the behavior of collectives to that of particular individual molecules will no longer be trustworthy.

But that is exactly the situation faced by all of psychology and behavioral science. History *does* matter – but it's poorly measured and modeled in traditional statistics. Imagine we are studying a specific mindfulness program's impact and mechanisms of action reducing anxiety. In a typical randomized trial, we might collect process (e.g., mindfulness) and outcome (e.g., anxiety) data from a large group of participants before and after the intervention and at follow-up, and do the same for a control group, calculating the average change in mindfulness and anxiety levels, and the relationship between them. The statistical significance of these data would be based on variability between people within conditions such as assessed by such things as a standard deviation. Such measures of variability are used like a rubber ruler to measure the difference between the central tendencies of the two conditions because that variability allows us to measure the importance of “chance” events once we assume that these extraneous events are truly randomly distributed. But what does that tell us about the histories and circumstances of individuals and how they responded to treatment? Perhaps some in the trial had a poor response to the intervention and decreased in mindfulness. These events are hardly “chance” (Britton, et al., 2021). What the practitioner needs to know about these events and how they interact to make good treatment decisions is not even measured, never mind modeled.

Note that to use the information collected at the level of groups, such as in policy or program evaluation decisions, participants need to be drawn randomly or representatively from a known population (that is, one in which the likelihood of selection is 1/N), that is isomorphic with the conceptual population (e.g., the collection of people indicated by the description of a population such as “persons with schizophrenia”). But even when this is done (e.g., Daly, et al., 2022), generalizations only mathematically apply to other random samples from that same population – not to a given clinic’s or practitioner’s unique sample of such persons. No applied agency applies its methods to representative samples. Said in another way, external validity is not a “mathematical must” in the traditional approach regardless of how well designed the group comparison study may be. At best, external validity is just a logical generalization, not a statistical one.

Typically, the shortcut used is to argue that effects seen at the level of the collection of individuals apply probabilistically to particular individuals over time if they carry the same diagnostic label. Such a strategy can be assumed to be legitimate only if the phenomena is ergodic and samples within a label are homogeneous, but diagnostic labels have been shown to be fuzzy and heterogeneous (e.g., Fried & Nesse, 2015). By definition, improvements or deterioration are not stationary phenomena, and the same statistical model virtually never applies to all people. In behavioral science, non-ergodicity is a nearly universal situation rather than an exception (Fisher, et al., 2018). People respond to intervention differently due to their unique circumstances, histories, and personalities. As we will show in this chapter, individual trajectories in mindfulness and its relationship to other events often diverge significantly from group averages. A new approach is needed to create more practical and theoretical progress.

An Idionomic Approach

A fresh and novel way to approach this problem is by modeling idiographic patterns—those unique to the individual—*before* generalizing to nomothetic (group-level) patterns, and then retaining only nomothetic generalizations that incrementally add to our idiographic understanding. We have labeled this approach “idionomic” analysis (Hayes & Hofmann, 2021; Ciarrochi, et al., 2022)

Note that unlike group comparison methods that must be “representative” to be generalized to other representative samples, idionomic methods apply to the particular person being modeled. The principles discovered may then be examined to see if they logically generalize based on the relevant theoretical and measured variables' precision, scope, and depth. For example, as mindfulness processes are more deeply understood person by person, the detected patterns can be tested directly for their theoretical and treatment utility.

No assumption of "representativeness" is needed a priori because the degree of generalization will be assessed. In other words, there is no gauntlet to run beyond successfully modeling process-to-outcome relationships for each person individually. After that, treatment utility can be directly tested. Does individual modeling improve outcomes?

This is precisely how functional analytic concepts such as “reinforcement” or “stimulus control” were vetted in the history of psychology, and these concepts have proven to be useful. In this chapter we will show both the need for idionomic analysis in the measurement of mindfulness, and how progress is unfolding rapidly as these cutting-edge methods are deployed to the measurement of mindfulness. Idionomic assessment may provide numerous benefits, such as personalizing mindfulness interventions by identifying which aspects of mindfulness need attention and determining the optimal sequence for addressing them. Additionally, this approach places mindfulness-relevant processes within the broader context of an individual's network of processes (e.g., self-efficacy, motivation), enabling a comprehensive personalization of the therapeutic process (Ciarrochi et al., 2024; Sanford, et al, 2022).

How is Ergodicity Assessed?

Ergodicity assumes that the statistical properties observed in a population are equivalent to those observed in individuals over time. This assumption is crucial in psychological research because it allows researchers to infer individual behaviors and processes from aggregated data collected from different individuals.

Mindfulness has been broadly defined as the awareness that arises through paying attention on purpose, in the present moment, non-judgmentally (Kabat-Zinn, 2013). There are a wide range of measures that link to these ideas, including single factor measures focused on attentional processes (Brown & Ryan, 2003) to multi-factor models focused on attention, acceptance, non-reactivity, describing, and observing experience (Baer, et al., 2006). Much research indicates that individuals with higher levels of mindfulness tend to experience greater well-being compared to those with lower levels (Baer et al., 2004). Experience sampling shows that, on average, mindfulness links to well-being within-person over time (Brown & Ryan, 2003). Intervention studies show that groups participating in mindfulness practices exhibit improvements in well-being when compared to control groups (e.g., Baer, 2003; Hofmann, et al., 2010) and sometimes mediational analysis show that these changes are due to change in mindfulness measures (e.g., Gu, et al., 2015). Such findings suggest a positive effect of mindfulness on well-being.

Whether we can conclude that mindfulness leads to enhanced well-being for each individual within these groups, however, depends on the applicability of the ergodic assumption. This assumption would allow us to apply group-based statistics to individual cases, assuming that time-averaged effects observed in a single individual align with the ensemble averages observed across the group. Only if this ergodic assumption holds true can we reliably use these group-based findings to make inferences about individual changes in well-being due to mindfulness. To evaluate the ergodic assumption, we need to collect intensive time-series data where we repeatedly measure mindfulness and important outcomes.

Nonstationarity violates ergodicity. A stationary time series has properties that do not depend on the time at which the series is observed (Hyndman & Athanasopoulos, 2018). This means that the statistical characteristics like mean, variance, and autocorrelation (the relationship between values at different times) are constant throughout the time series.

Now, let's consider the link between mindfulness and well-being. If mindfulness is not stationary, then knowing a person's mean, variance, and autocorrelation at one time does not necessarily let us know what it is at another time. Hypothetically, if there are times when mindfulness is chronically low and has high inertia (high autocorrelation), then mindfulness may have a little within-person effect on changes in well-being. At other times, when mindfulness varies substantially, perhaps due to an introduction to mindfulness practice, then changes in mindfulness may have substantial effects within a person.

Thus, if the variables involved in the process are non-stationary, the effects observed at one time may not be observed at another. We cannot conclude that at a particular time, mindfulness will have a specific benefit on the well-being for a particular individual. There are a number of statistical tests that can check for indicators of nonstationarity, including the Augmented Dickey-Fuller (ADF) test and the Kwaitkowski-Phillips+Schimdt-Shin (KPSS) test (Kwiatkowski et al., 1992), which are routinely used in time-series forecasting, whether manual or automatic (Hyndman & Khandakar, 2008).

Effect heterogeneity violates ergodicity. If the effects of mindfulness are not homogenous (vary across individuals), then mindfulness may be highly beneficial for some but inert or harmful to others. Testing for violations in homogeneity can take many forms. One method examines the importance of link between processes and outcomes like well-being or performance. Here, "importance" refers to the strength of a relationship. For example, we might ask if the average mindfulness link to well-being (e.g., $r = .30$) is consistent for each individual. Testing this assumption requires repeated assessment of processes and outcomes within individuals, enabling the modeling of both group and individual dynamic systems (Ciarrochi, Sahdra, Fraser, et al., 2024; Ciarrochi, Sahdra, Hayes, et al., 2024; Sahdra et al., 2024). One can then perform meta-analysis, treating each individual as a study, and examine if the group or "pooled" effect describes individuals well, or if there is a significant heterogeneity.

Thinking of this as a kind of meta-analysis is especially useful because it provides heterogeneity statistics, such as I^2, which represents the percentage of total variability across studies (or participants in this case) that is due to true heterogeneity (systematic differences among instances) rather than chance (Higgins et al. 2003). Generally, values over 50% reflect high inconsistency, and over 75% reflect very high inconsistency, and according to meta-analytic experts generally should prompt the person to search for subgroups and avoid reporting pooled effects (Ioannidis et al. 2007). So far, we have found that almost all process variables show high levels of heterogeneity in effects and the majority show very high level (Ciarrochi, Sahdra, Fraser, et al. 2024; Ciarrochi, Sahdra, Hayes, et al. 2024; Sahdra et al. 2023; 2024).

Structural Heterogeneity Violates Ergodicity

The concept of mindfulness is often assumed to be reflected in behaviors such as observing, judging, and labeling emotions. These might even be considered elements of mindfulness as in Jon Kabat-Zinn's well-known definition of mindfulness as "the awareness that arises from paying attention, on purpose, in the present moment and non-judgmentally" (2017, p. 1127). Elements such as these may not be uniformly applicable across individuals. Some may judge more when paying attention to the present moment, others may judge less – and such features may vary over time. This variability extends beyond traditional definitions of mindfulness, positioning it within a broader network that could include elements like self-efficacy and feelings of connection. These relations, too, may be idiosyncratic and may vary over time.

For mindfulness to represent the same phenomenon for everyone in ways that do not require a focus on particular individuals, it must maintain a consistent internal structure and exhibit uniform interactions with other aspects of the network of mindfulness related processes across people. Without such consistency, the same nomothetic measures of mindfulness reflect different forms of mindfulness person by person. For example, mindfulness may function differently amongst those who are self-compassionate or self-critical, amongst those who seek pleasure (hedonists) versus those who value self-control, fortitude, and wisdom over worldly pleasures (stoicism), or among those in collectivistic versus individualistic cultures.

The variability in how mindfulness is experienced poses significant measurement challenges and barriers to the treatment utility of assessment for mindfulness interventions. Research suggests that beginners and experts have different mindfulness experiences. For example, beginners might engage more in mind wandering (Kashkouli Nejad et al., 2014; Lu & Rodriguez-Larios, 2022), use more top-down processes (Chiesa et al., 2013), and find pain more unpleasant (Lutz et al., 2013). Thus, we almost by definition expect an intervention to change the "nature" of mindfulness over time.

A similar and prevalent assumption is the uniform applicability of the same mindfulness factor model across different participants. For example, we might assume that the same five latent factors drive people's responses to mindfulness items as assessed in measures such as the Five Factor Mindfulness Questionnaire (Baer et al., 2008), namely, observing, nonreactivity, non-judgment, acting with awareness, and describing. This assumption facilitates the application of a common statistical model to all participants, simplifying the analysis and interpretation of mindfulness' effects across various outcomes. However, embracing heterogeneity may provide a more accurate reflection of human experience, where each individual follows a unique behavioral trajectory, shaped by factors like baseline stress levels, psychological resilience, and personal mindfulness history. If every participant responds to mindfulness based on a distinct factor model, this diversity implies that standard factor analysis might fail to capture important effects, potentially leading to inaccurate conclusions about the effectiveness and impact of mindfulness practices. For example, the relationship between observing and nonjudging might be negative for some individuals who observe primarily to detect social threats, suggesting that promoting observing in these cases could be counterproductive. Latent profile analysis research on mindfulness does identify a distinct profile of "judgmentally observing" people who seem to live highly effective lives despite some cost to their mental health (Sahdra et al., 2017).

There are various methods to evaluate structural heterogeneity, including p-factor analysis (Nesselroade et al., 2007) and network analyses (Epskamp et al., 2018; Schmittmann et al., 2013). Rather than assuming that individuals all exhibit the same factor structure or mindfulness, p-factor analysis allows a separate measurement model to be estimated for each individual. P-factor still makes the assumption that latent variables drive observed mindfulness responses. For example, one person's responses may be driven by a global "mindfulness" factor, whereas another person's responses may be driven more by a latent "observing experience" and a "curious, open orientation to experience" factors. These p-factors may differ substantially from the factor structure observed at the group level, where responses are presumably driven by a global mindfulness factor and five sub-factors (describe, act with awareness, non-judgment, nonreactivity, and observing). Network analysis goes a step further by making no assumption about latent constructs and allowing processes to interrelate in different ways to each other. Such networks can be constructed at the individual or subgroup level (Ciarrochi, Sahdra, Fraser, et al., 2024; Sahdra et al., 2024). There are now tests that can evaluate if the structures of idiographic networks are significantly (in)equal (Hoekstra et al., 2024).

The Traditional Response to Ergodicity

Most people will readily agree that people have heterogeneous responses to mindfulness interventions. However, the common response to such heterogeneity is to presume it to be “error” or unexplained variance, the price of doing business in psychology. The most common statistical way of dealing with it is to utilize a method like multi-level modeling (MLM), where heterogeneity is measured and “controlled for”, and the pooled effect and standard error are unbiased. In this way of thinking, the average effect is the true signal; everything else is just noise. Collecting larger numbers of participants and having better designed studies is expected to reduce the noise and get us closer and closer to an estimate of the true signal.

If the ergodic assumption is violated, however, no amount of statistical controls or improved design will fully remedy the situation. The average effect will fail to describe many individuals, even if that average effect is based on five billion participants or the most perfectly controlled trial. There is no way to log transform the heterogeneity away. Nonergodicity implies that different causal dynamics are operating for different people.

The issue becomes particularly clear when we talk about what we have coined “equisyncratic” variables (Ciarrochi, Hernandez, et al., 2024). These are variables that have no significant average population-level effect but have substantial individual-level variability in effect. For example, research has found that some social variables like “being assertive” and “expressing feelings” have no average effect on loneliness (Ciarrochi, Sahdra, Hayes, et al., 2024). and yet they still have powerful but heterogeneous individual effects, such as assertiveness increasing loneliness for some and reducing loneliness for others. If the “average” is king, then these social behaviors would be considered unimportant. The tests of heterogeneity and what is seen clinically will tell another story.

Shrinkage for the “Greater Good”

Shrinkage in the context of within-person relationships between mindfulness and well-being refers to the bias of individual estimates towards the group average. A high positive relationship indicates that when a person is highly mindful relative to their own baseline, they experience high levels of well-being. However, shrinkage suggests that an individual's estimate might be biased towards the group average, underestimating their unique experience with mindfulness.

Within-person data can be examined in three ways: (1) “No-pooling” approach: each person’s within-person associations (e.g., between mindfulness and well-being) can be examined separately ignoring the group-level estimates; (2) “Complete pooling” approach: the association between mindfulness and the outcome can be calculated for the overall sample ignoring the fact that measurements from the same individuals are not independent; and (3) “Partial pooling” approach: a multilevel modeling compromise between the no-pooling and complete pooling approaches, where individual-level estimates are shrunk towards the group mean (Gelman & Hill, 2007). Multilevel models are commonly used to estimate individual trajectories in longitudinal studies both in high and low-density settings (Bolger & Laurenceau, 2013; Singer & Willet, 2003), as they allow both the estimation of a fixed effect (an average across individuals) and individual deviations (random effects). Said random effects are usually assumed to be an adequate estimation of individual trajectories, in particular when within-person variance decomposition is used (i.e., variations relative to a person’s average; Curran & Bauer, 2011). However, the main goal of multilevel models is not to precisely describe individual trajectories to represent each case faithfully but to estimate an average effect considering the existence of deviations. In fact, the common MLM model utilized in psychology has an objective function of maximizing the likelihood of the observed data given the fixed effects to make inferences about them (Raudenbush & Bryk, 2002), so the random effects are something to be controlled for.

In this context, random effects are also assumed to be normally distributed around the fixed effect (Singer & Willett, 2003). Consequently, and for the “greater good” of creating more stable estimates that are less affected by fluctuations at the individual level, random effects (e.g., the effect of mindfulness on depressive symptoms for each person) are adjusted towards the fixed effect (e.g., the weighted average effect across individuals). This adjustment, known as the "shrinkage factor", is based on the precision of the random effects estimates: the more precise the random effect, the less the adjustment; the less precise the random effect, the greater the adjustment. Based on a simulation study, Liu, et al. (2021) showed that more shrinkage will occur with larger within-person residual variance, larger random effect variance, smaller level 1 predictor variance, and smaller number of within-person observations. The algorithm is appropriate if the main goal is to get more stable estimates based on fixed effects. However, if the main goal is to explore heterogeneous effects, edge cases, or individual trajectories, then MLM random effects may obscure individual differences (Liu, et al., 2021). The higher the heterogeneity, the more pronounced the shrinkage will be.

How Hidden Can Someone Be?

The extent to which an estimate for an individual will be shrunk towards the average is hard to estimate beforehand. However, recent findings by Sahdra et al. (2024) have shown not only that shrinkage occurs and can severely decrease variance and thus model-implied heterogeneity when using MLM; but surprisingly this shrinkage can even flip the sign of individual estimates for many cases in which process → outcome relations can be moved from helpful to hurtful or vice versa for particular individuals. Sahdra et al. (2024) compared three methods for calculating individual-level estimates of within-person associations of valued action ('doing what matters') with joyfulness and sadness: raw within-person correlations, multilevel model implied estimates of individuals' associations, and estimates from idiographic autoregressive integrated moving average models with an exogenous variable (i-ARIMAX; Ciarrochi et al., 2024). The mean estimates from the three methods were identical (mean of estimates of valued action linked to joyfulness: 0.40), though the spread around the mean was much smaller for the individual estimates from multilevel models (SD = 0.05), compared to the raw within-person correlations (SD = 0.28) or i-ARIMAX estimates (SD = 0.29). The range of multilevel model implied estimates was much narrower (0.21 to 0.58) than the range of either raw within-person correlations (-0.58 to 0.98) or i-ARIMAX estimates (-0.58 to 1.00). Alarmingly, the range of the multilevel model implied estimates did not include any negative values, evident in raw within-person associations and i-ARIMAX estimates.

Using only multilevel model estimates would create the illusion that most people are affected the same way by valued action (e.g., by having a positive association with joyfulness). This would inadvertently hide dissonant voices and hinder our progress towards psychological interventions that precisely target processes that are personally relevant, not only based on average effects. In practical terms, increasing valued action for those counter-average cases would theoretically increase the likelihood of adverse events or negative consequences.

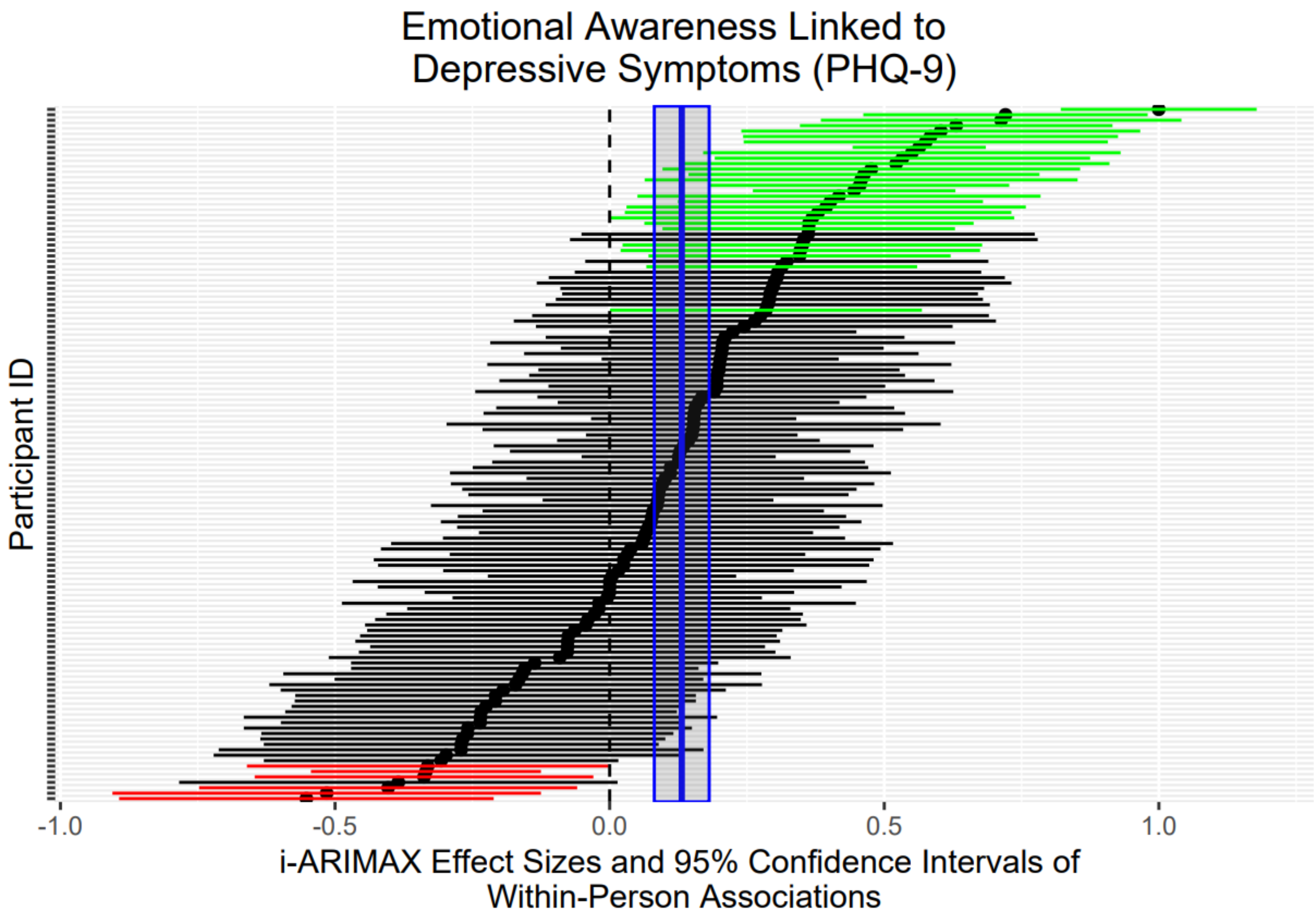
If the goal is to make population-level inferences, then multilevel modeling and other purely nomothetic methods are appropriate because such methods allow us to estimate group mean effects as precisely as possible. However, if the goal is to apply the insights from a nomothetic analysis to predict how a new individual, e.g., in a fresh sample or a new client in a therapist's office might behave, then the mean effects from nomothetic studies can be misleading.

Using an example from Sahdra et al. (2024), a multivariate random-effects meta-analysis of the i-ARIMAX estimates of valued action items predicting joyfulness yielded a pooled effect estimate of 0.35 with 95% CI of 0.32 to 0.39. The I^2 estimate of heterogeneity of the effects was 88.71%. According to Higgins et al. (2022), if the I^2 is above 50%, extreme caution is needed while reporting and interpreting the pooled effect because it can misrepresent the population effect. Applying the same logic to the I^2 of 88.71% reported by Sahdra et al. (2024), the pooled effect showing a positive association between valued action and joy needs to be interpreted with a big grain of salt. Sahdra et al. also reported the prediction interval of the pooled effect, which was -0.30 to 1.00. Even when the confidence intervals of the mean effects are relatively narrow, that does not guarantee that the prediction interval would be similarly narrow. Confidence intervals convey the precision of the estimate of the population parameter, whereas prediction intervals tell us where the point estimate of a new person in a fresh sample might lie (IntHout et al., 2016). When the prediction interval includes zero, even if the confidence interval does not, we need to be extremely cautious in using past knowledge of statistically significant effects to guide interventions for specific individuals from whom we have no prior data on mindfulness or the relevant variables of interest.

Illustration with Emotional Awareness

We return now to the implications of these ideas and findings for assessing one aspect of mindfulness, emotional awareness. In a recent study from our team, we selected a subsample of 128 adolescents and young adults aged 14 to 25 years who had at least 20 valid observations in an EMA study of 10 days, with 3 prompts per day. We evaluated the behavior emotional awareness (i.e., "Since the last beep/yesterday, I paid attention to how I was feeling") in a seven-point VAS scale with anchors "not at all" and "to a great extent". We also evaluated depressive symptoms using both PHQ2 (Kroenke, et al., 2003) items adapted to EMA format using the same VAS scale. Furthermore, we measured the same features at baseline with their original scales: Paying attention to emotions was measured with question 2 from the Difficulties in Emotion Regulation Scale (Gratz & Roemer, 2004), while depressive symptoms were measured with the PHQ2, so items were equivalent in content at baseline and EMA stages. Between-person baseline correlations showed the expected pattern of a small negative relationship ($r = -0.249$, $p < 0.001$). However, when analyzing within-person associations using both idionomic i-ARIMAX and MLM the relationship reversed and moments higher in paying attention to emotions were associated with higher depressive symptoms (i-ARIMAX pooled effect = .131, $p < .001$; MLM fixed effect = .148, $p < .001$). This reversal was not due to modeling strategy or within-person centering, as the distribution of between-person correlations at each timepoint was mostly positive. Moreover, I^2 was 69.75%, suggesting that the average effect was not a good representation due to high heterogeneity, as shown in Figure 1 with cases in the positive and negative effect range.

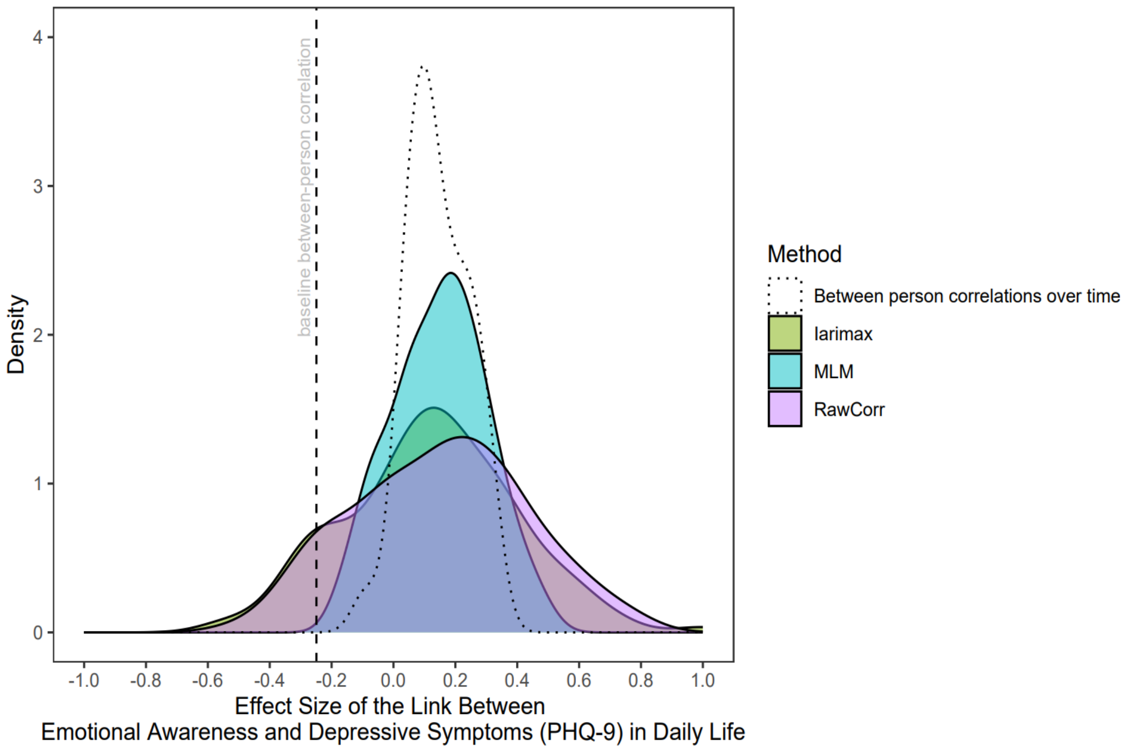
Figure 1



Note: Blue vertical lines represent the RE-MA pooled (nomothetic) effect in the middle, and the lower and upper bounds of the 95% CI of the pooled effect. Green horizontal lines indicate 95% CI of positive associations and red indicate negative.

Raw correlations showed that 32% of the participants had relationships that were opposite to the pooled effect, however this number decreased to 17% when MLM model implied random effects were used. I-ARIMAX individual relationships were more closely related to the original data with 31% of participants showing opposite patterns. As expected, MLM variance of individual estimates was reduced drastically (.023) when compared to raw correlation variance (.079), which was consistent with i-ARIMAX variance of individual effects (.077). See Figure 2 for a comparison of these different methods and between-person correlations.

Figure 2



Note: Density plots for I-ARIMAX estimates for the exogenous covariate (green), MLM estimates of individual slopes reconstructed from random effects (blue), and individual raw correlations (purple) are presented. The between-person correlation at baseline is represented with a vertical dashed line. Between-person correlations for each time point are presented with a dotted-line density plot.

Is paying attention to our emotions helpful or hurtful then? It depends, and that is the main point. More information is needed, but by focusing on group-level estimates, we may lose track of the human voices that are being obscured by the group or shrunk and flipped for the "greater good". Acknowledging their existence is a necessary starting point to explore and try to discover new principles that would help us understand why and when some things may be useful or harmful when helping particular people. It is common sense not to think of an average (e.g., average wage in a given country) as the representation of every unit but rather as a sometimes useful summary (e.g., not everyone earns the "average wage"), but we may forget that in our work if we fail to remember that our MLM estimates are weighted averages of relationships. Human beings are complex and processes such as paying attention to emotions, exist in relationship to a network of other related processes, that together help understand its function and meaning (Hayes et al., 2019). To understand mindfulness better, a useful strategy may be to lower the level of abstraction from the broad "mindfulness" experience to related processes that participate in the unfolding lives of those we serve.

EEMM Machine Analysis of all Mindfulness Measures

At the higher level of abstraction (i.e., latent construct), confirmatory factor analyses have been used as evidence that a certain measure indicates a latent construct in particular, such as mindfulness (e.g., Bear et al., 2008). However, a good fit of the model to the data is no guarantee that the proposed items measure a given "latent" construct in "reality" – it assesses whether the proposed structure is capable of reproducing the original variance-covariance matrix (Brown, 2015). Thus, model fit does not rule out other structures or directly say anything about what is measured: That is inferred based on theory and the content of items. In fact, sometimes, the same data can be equally represented by a network structure in which items are directly related between them (Kruis & Maris, 2016).

Consequently, assuming a latent and well-defined construct based on good model fit can be misleading. This assumption may create the false impression that different questionnaires labeled as "mindfulness" measure the same underlying construct, which is not always true. One example is depression, where different measures have been shown to lack substantive content overlap (Fried, 2017). Is it then likely that different mindfulness measures are heterogeneous because they measure different processes within and between questionnaires?

A useful background to answering this question is the Extended Evolutionary Meta-Model (EEMM), a common language and framework for discussing change processes (Hayes, et al., 2020). Here, processes of change are understood as theory-based, dynamic, progressive, contextually bound modifiable, and multilevel sequences of events that are known to alter or transform a person's functioning (e.g., well-being).

The EEMM posits that adaptation depends on the variation, selection, and retention of change processes in particular contexts (Ciarrochi, Hernández, et al., 2024; Hayes, et al., 2020). In the context of mindfulness, this implies that specific processes such as paying attention to the present moment, being aware of emotions, and noticing reactions to mistakes, are chosen from a wide range of possible experiences. These processes are maintained based on their function or significance. For some, their function might be experiencing joy, while for others it may be a way of taking away the attention from worries, developing a self-concept of being “strong” when facing adversities, or avoiding taking action based on guilt. As such, within mindfulness or any other human experience, different processes interacting may play a role.

To navigate this heterogeneity and foster common ground, the EEMM classifies processes of change into six contextually bound psychological dimensions: cognition, affect, self, motivation, attention, and overt behavior. These dimensions are nested within three levels of complexity: biophysiological (or "physical"), psychological, and social (Hayes, et al., 2020). Of course, ultimately the physical and social levels of analysis should also be dimensionalized, but that is an as yet undone project for the field at large interested in processes of change. With this structure in mind, we can then ask ourselves: what families of processes are we actually measuring when we measure mindfulness with existing measures?

Rather than do this in a way that is difficult to replicate, we examined 16 unique adult and adolescent measures of mindfulness that were described as chapters in the present volume. We categorized their items with two distillBERT language models previously trained to classify item questionnaires into the six EEMM psychological dimensions, context, and the three levels of analysis (see Ciarrochi, Hernández, et al., 2024 for a description, performance metrics, and scoring manual). Table 1 shows the percentage of appearance of each dimension and level within each questionnaire. As expected, the most represented category was “Attention”, which was addressed to some degree by all 16 measures. All other dimensions and levels were absent in at least some measures and different processes were more or less emphasized between questionnaires. For example, the Langer measure had a heavy focus on self and motivation and much less focus on attention, whereas the Mindfulness Attention and Awareness

Scale focused most strongly on attention and had little motivation or self content.

Table 1

Percentage of times items were inferred to contain each dimension and level by fine-tuned versions of DistilBERT, as described in Ciarrochi et al. (2024), by questionnaire.

measure	Cognition	Affect	Self	Motivation	Attention	Overt Behavior	Context	Social	Psychological	Physical
Adolescent and Adult Mindfulness Scale (AAMS)	37,5	37,5	4,17	0	54,17	16,67	0	0	100	16,67
Applied Mindfulness Process Scale (AMPS)	62,5	18,75	6,25	0	18,75	18,75	0	0	100	12,5
Comprehensive Inventory of Mindfulness Experiences (CHIME)	36,36	45,45	4,55	0	63,64	27,27	0	9,09	100	4,55
Freiburg Mindfulness Inventory	30	20	40	0	40	0	10	10	100	10
Mindfulness Attention and Awareness Scale	26,67	13,33	0	6,67	86,67	40	0	13,33	100	13,33
Child and Adolescent Mindfulness Scale	40	50	20	0	40	10	0	0	100	0
Cognitive and Affective Mindfulness Scale-Revised (CAMS-R)	30	30	10	0	70	0	0	0	100	0
Five Facet Mindfulness Questionnaire	51,28	30,77	7,69	0	58,97	15,38	0	0	100	12,82
Interpersonal Mindfulness Questionnaire	31,58	2,63	15,79	0	52,63	26,32	44,74	100	100	31,58
Kentucky Inventory of Mindfulness Skills	43,59	28,21	7,69	0	56,41	15,38	0	0	100	23,08
Langer Mindfulness Scale	38,1	0	33,33	23,81	9,52	23,81	4,76	9,52	100	0
PHLMS Philadelphia Mindfulness Scale	30	45	5	0	80	10	0	15	100	20
Relaxation Mindfulness Scale for Adolescents	27,78	22,22	5,56	0	55,56	33,33	0	5,56	100	27,78
Southampton Mindfulness Scale	52,94	23,53	17,65	0	58,82	0	0	0	100	0
State Mindfulness Scale	14,29	23,81	9,52	0	76,19	0	0	0	100	33,33
Toronto Mindfulness Scale	80	13,33	26,67	0	60	0	0	0	100	0
Average (all items across measures)	39,70	24,48	11,94	1,79	55,52	16,42	5,67	14,63	100,00	15,52

Note: Two DistilBERT models were fine-tuned based on human and AI ratings of a set of measures used in clinical mediational studies (Hayes et al., 2022). One model was used to classify dimensions (Cognition, Affect, Self, Motivation, Attention, Overt Behavior, and Context), while the other was used to classify levels (Social, Psychological, and Physical) using a multi-label task. For further details and performance metrics, see Ciarrochi et al. (2024). Both models were used for the current task of classifying mindfulness measure items

This visible difference in focus makes it more relevant to decrease the level of abstraction from overall latent constructs to those processes involved in their coming to existence. Even within each family of processes, different items can have different implications in the context of a complex network. For example, items categorized as “Self” by our AI model, such as “I disapprove of myself when I have irrational ideas,” and “I see my mistakes and difficulties without judging them” can have different implications for well-being, depending on their function and context. A seemingly negative process such as self-disapproval, may at times be useful if it fosters healthy actions and change when necessary, while a seemingly positive one such as diminishing self-judgment of mistakes and difficulties, may be harmful if it stagnates change and movement. By lowering the level of abstraction and exploring heterogeneity from a bottom-up perspective an idionomic approach allows us as a field to explore critical questions about the promotion of well-being that remain elusive when under the solely nomothetic umbrella.

Early Application of Idionomics to Mindfulness Measures

Although not the same as mindfulness, compassion is a mindfulness-related construct. We will briefly explore this work here because recent research has examined compassion through an idionomic lens in small studies (Ciarrochi, Sahdra, Fraser et al., 2024; Sahdra et al., 2023). Sahdra et al. (2023) observed substantial heterogeneity in the link between self-compassion and compassion towards others: although most people showed a positive link, some people showed a null effect, and a minority showed a negative link. The researchers compared raw within-associations of self- and other-compassion with multilevel model implied individual estimates and found multilevel based estimates to have a narrower range than the raw associations. They then examined the links between compassion variables and well-being and tested whether the within-person correlations between self- and other-directed compassion moderated how compassion affected well-being. Self- or other-compassion were positively linked to well-being but only among those who had positive within-person correlations of self- and other-compassion. Sahdra et al. named this pattern 'self-other harmony' in compassion. Individuals who lacked self-other harmony in compassion did not show the well-being related benefits of compassion in their day-to-day life. It was important to examine the data idionomically to identify such people in the data. This approach is significant from a clinical point of view because individuals who deviate from the nomothetic effect may not benefit from mindfulness-related interventions designed to boost compassion. It may be necessary to first examine the reasons for the lack of harmony of compassion in such individuals before trying to promote compassion.

In a related study, Ciarrochi, Sahdra, Fraser, et al. (2024) examined the links between self-compassion, other-compassion, and romantic attraction in couples. They first computed i-ARIMAX estimates of within-person associations between self-compassion, other-compassion, and attraction, and summarized the results using random-effects meta-analyses. These analyses showed substantial heterogeneity. Some couples had strong associations between compassion and attraction for the female partner but not the male, or vice versa. Some couples showed strong links of attraction with self-compassion but not other-directed compassion. The researchers then conducted a cluster analysis, which yielded two distinct groups of couples: 'synergistic' couples for whom compassion affected attraction and 'independent' couples for whom compassion was inert for attraction. Their study is the first to conduct an idionomic analysis at the dyad level and show violations of the ergodic assumption at that level. The nomothetic effects based on aggregation of data of all couples did not apply to specific couples. Efforts to use compassion to improve attraction in couples are not likely to be beneficial for couples lacking a synergy in compassion and attraction. It would be more productive for clinicians to look for other variables that may have a stronger impact on interpersonal attraction in such couples.

A more direct idionomic analysis of mindfulness can be found in Hayes et al. (2025), who examined the links of mindfulness with negative functioning. In a high-frequency within-person data, the researchers measured mindfulness using the item, “I was struggling to connect with the moments of my day-to-day life” (from the PBAT; Ciarrochi et al., 2022), and negative functioning using the Stop-D instrument (Young et al., 2007). They computed i-ARIMAX estimates of the mindfulness item predicting negative functioning for each person and then meta-analyzed the estimates. The pooled effect was 0.25 (95% CI: 0.16, 0.33), suggesting that lower levels of mindfulness were related with higher negative functioning nomothetically. When they examined the effect heterogeneity, the I^2 of 88% was alarmingly high. There were several individuals in the data for whom the within-person link between mindfulness and negative functioning was not reliably different from zero. For many other individuals, lower levels of mindfulness were associated with higher (not lower) levels of negative functioning. Such individuals who deviate from the expected overall effect often remain hidden in the mindfulness literature, which is dominated by nomothetic effects that shrink individual effects towards the group mean effect.

They then examined interconnections of mindfulness with several other variables in two groups of individuals based on the median split of i-ARIMAX estimates. The networks consisted of the following variables: mindfulness measured as struggling to connect with the present moment, feeling stuck and unable to change, struggling to keep doing something that was important, having no outlet for feelings, thinking helping life, and negative affect. Hayes et al. (2025) used multilevel vector autoregressive models to build within-person contemporaneous networks (Epskamp et al., n.d.), which show within-person conditional links between variables within the same time point accounting for previous time points. The networks showed that mindfulness was interconnected with every other variable in the network in the group with high i-ARIMAX estimates but with fewer variables in the group with low i-ARIMAX scores. That is, how mindfulness operated in the constellation of other relevant variables differed in the two groups. These effects became evident from the bottom-up approach of idionomic analysis, which started with idiographic analysis to create the two groups. They would have remained hidden if the starting point had been a purely nomothetic approach, such as multilevel modeling that is typically used to analyze experience sampling data.

Another recent study by Moritz et al. (2024) also provides a direct idionomic analysis of mindfulness. The researchers used an experience sampling design and administered five items of state mindfulness based on the MAAS (e.g., “I was finding it difficult to stay focused on what was happening in the present”), three items of experiential avoidance (“my emotions have got in the way of things which I wanted to do,” “I’ve tried to block negative thoughts out of my mind,” and “I’ve tried to avoid painful memories.”), and eight items of PANAS measuring affective states (joyful, content, relaxed, enthusiastic, anxious, sad, angry, and sluggish). Mindfulness items were reversed scored, such that higher scores indicated greater mindfulness. The researchers conducted separate i-ARIMAX models linking mindfulness with experiential avoidance items (15 effects nested per person) and a second set of i-ARIMAX models linking the mindfulness items to the affect items (40 effects nested per person). To summarize the overall pooled effects and heterogeneity estimates in a concise manner, they conducted multivariate random-effects meta-analysis (MV RE-MA) models in which the effects of mindfulness items on the outcome items—experiential avoidance or affect items—were nested within person. The estimates of negative affect items were reverse scored, which allowed all affect related estimates to be included in a single MV RE-MA. Log-likelihood ratio tests were run to compare the fit of a 3-level (“Full”) MV RE-MA model that allowed the effects at all levels to vary with the fit of a 2-level (“Reduced”) model where the person level was fixed. A 3-level model fits the data better than a 2-level model ($p<0.001$) in each case.

The results of the MV RE-MA showed that the pooled effect of mindfulness predicting experiential avoidance was -0.20 (95% CI: -0.26, -0.14). On average, mindfulness was linked with less avoidance. However, there was substantial heterogeneity of this effect: the 95% prediction interval was -0.78 to 0.38, and the I^2 was 82.62%. The MV RE-MA of mindfulness predicting hedonic well-being yielded a pooled effect of 0.18 (95% CI: 0.02, 0.14), showing an overall positive effect of mindfulness on well-being. However, the I^2 was 76.22% and the 95% prediction interval was -0.31 to 0.67.

These idionomic results replicate past nomothetic findings regarding the benefits of mindfulness for experiential avoidance and affect, but also importantly add to the literature by documenting the substantial heterogeneity in the observed idiographic effects. They combine the insights from idiographic and nomothetic analyses. They show violations of the ergodic assumption. They question the psychological homogeneity assumption that mindfulness' effect on well-being is similar for all people. At least for some people, mindfulness increases avoidance or reduces hedonic well-being.

An idionomic approach insists on acknowledging such deviations from the overall effect, which can prevent clinicians from applying a "one-size-fits-all" approach of promoting mindfulness in every client. Mindfulness is beneficial for many people—a claim supported by the data. But it is not beneficial for everyone. The alarmingly wide prediction intervals documented by Mortiz et al. (2024) tell us that a new person in a fresh sample may have negative or positive association between mindfulness and well-being; there is no way to be sure in the absence of data from that person. These findings should give a pause to clinicians while applying past nomothetic findings about mindfulness to new clients and encourage them to gather data from their clients to make data-driven clinical decisions to personalize treatment of their clients.

The Practicality of Particularity

We can return now to the ultimate practical question: Are mindfulness measures doing a good job of orienting practitioners and researchers toward how best to empower the development of particular human beings? If we do not yet know, how can we make rapid progress in answering that question?

From the limited data we have reviewed in this chapter, it is clear that mindfulness measures and models will fracture in the presence of the booming voices of particular people. Idionomic analysis amplifies the voices of those we serve and like an opera singer shattering a glass, in the well examined lives of many particular people existing models begin to crack and crumble. This can be frightening but it is also exciting because existing models can be reassembled in a new form and after decades of stagnation we can begin to examine the most practical question without further ado.

If the focus of reassembling our measures, analyses, models, and methods is on treatment utility and the theoretical utility of idionomic principles with high precision / scope / depth that can feed that outcome, the coming era is likely to be a progressive if iconoclastic exercise. Despite the conceptual anomie it is likely to produce, exploring this possibility could be rapid because, unlike the last 100 years of psychometric development, small research groups or even individual practitioners can be part of the idionomic research process, with well-examined particular lives as the focus.

Consider a hypothetical example. A research group believes that mindfulness is primarily a matter of attention, but adds items on emotional acceptance, non-judgment, and a spiritual sense of self vs. items examining a more egoic sense of self. Along with emotional well-being and some measure of social function, these are tracked daily for a month or two in two dozen persons. In our imaginary study, an idionomic analysis shows that "mindful attention" is negatively related within individuals to well-being over time for ⅓ of those in the study, positively for another ⅓, and uncertain for the final third. By relying on the larger network of assessment questions it begins to make sense. Some of those for whom attentional mindfulness is negatively related to well-being become emotionally overwhelmed because their acceptance or non-judgment skills are limited; those with better outcomes tend to have a broader and more complete set of skills. Curiously some of those with negative relationships of attentional mindfulness to well-being have all the skills, but on closer examination they have adopted a narcissistic narrative (a "guru syndrome"), which is creating social problems that in turn is interfering with well-being. If each analysis is followed by individually assessed treatment targeting the key processes in a given analysis, data on the treatment utility of more complex analyses such as these would rapidly accumulate.

Scores of teams, and hundreds of practitioners, could readily mount such an idionomic research program on mindfulness measures in an iterative inductive fashion. The field would need to be ready with a larger set of items and measures reflecting theoretical diversity to follow this strategy when idionomic analyses fractionates existing nomothetic models because real treatment utility is likely to require it. But the democratization of idionomic research has unique potential advantages in the speed of variations that can be tested. Generalization and its nomothetic statistical expression will require high precision/scope/depth ways of speaking about the individuals we analyze, but that in turn will be dependent on the networks of processes needed to understand each person. Whether the field is ready or is not, rapid feedback is likely, and with it the chance for rapid improvement.

Five conclusions from this line of thought can be derived as follows:

1. Idionomic analyses will need to lead to networked analyses of the "acts in context" of particular individuals (and particular couples; and particular families; etc)
2. For that reason, assessment must be broad, and thus mindfulness measurement must be seen in the context of a range of other relevant processes in order to be understood. That may sound overwhelming, but it is not since EMA approaches abandon large multi-item scales and can enormously democratize process-based research itself.
3. Clinically and practically speaking, individually relevant outcomes and goals need to matter in such networks because what the person cares about gives an anchor and purpose to the networks and to the assessment itself.
4. The production of clinical outcomes are best tested experimentally and thus the linkage to treatment and treatment utility is the ultimate arbiter of success. In the model just described no gauntlet need be run. The field can begin. Now.
5. In this idionomic strategy every voice matters -- and if you are not measuring and modeling at the level of the particular (the particular person / couple / family / organization and so on) you are not listening.

The idionomic approach is in its infancy but progressing very rapidly. It is a fundamentally new direction in assessment and intervention science. Time will tell if it is progressive, but at least it can meet its own analytic assumptions, which, as we have shown, the traditional approach simply cannot.

References

Baer, R. A. (2003). Mindfulness Training as a Clinical Intervention: A Conceptual and Empirical Review. Clinical Psychology: Science and Practice, 10(2), 125–143. https://doi.org/10.1093/clipsy/bpg015

Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Assessment of mindfulness by self-report: the Kentucky inventory of mindfulness skills. *Assessment, 11*(3), 191–206. https://doi.org/10.1037/t11612-000

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

Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences, 17*(12), 656–660. https://doi.org/10.1073/pnas.17.12.656

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford Press.

Boltzmann, L. (1884). Ueber eine von Hrn. Bartoli entdeckte Beziehung der Warmestrahlung zum zweiten Hauptsatze. *Annalen der Physik, 258*(5), 31–39. https://doi.org/10.1002/andp.18842580503

Bringmann, L. F., Albers, C., Bockting, C., Borsboom, D., Ceulemans, E., Cramer, A., ... & Wichers, M. (2022). Psychopathological networks: Theory, methods and practice. *Behaviour Research and Therapy, 149*, 104011. https://doi.org/10.1016/j.brat.2021.104011

Britton, W. B., Lindahl, J. R., Cooper, D. J., Canby, N. K., & Palitsky, R. (2021). Defining and measuring meditation-related adverse effects in mindfulness-based programs. *Clinical Psychological Science*, 9(6), 1185-1204. https://doi.org/10.1177/2167702621996340

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, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.

Burke, L. E., Shiffman, S., Music, E., Styn, M. A., Kriska, A., Smailagic, A., Siewiorek, D., Ewing, L. J., Chasens, E., French, B., Mancino, J., Mendez, D., Strollo, P., Rathbun, S. L. (2017). Ecological momentary assessment in behavioral research: Addressing technological and human participant challenges. *Journal of Medical Internet Research, 19*(3), e77. https://doi.org/10.2196/jmir.7138

Cattell, R.B. (1952). The three basic factor-analytic designs – Their interrelations and derivatives. *Psychological Bulletin, 49*, 499-520. https://doi.org/10.1037/h0054245

Cattell, R. B., Cattell, A. K., & Rhymer, R. M. (1947). P-technique demonstrated in determining psychophysiological source traits in a normal individual. *Psychometrika, 12*, 267-288. https://doi.org/10.1007/bf02288941

Chiesa, A., Serretti, A., & Jakobsen, J. C. (2013). Mindfulness: top-down or bottom-up emotion regulation strategy? *Clinical Psychology Review, 33*(1), 82–96. https://doi.org/10.1016/j.cpr.2012.10.006

Ciarrochi, J., Hernández, C., Hill, D., Ong, C., Gloster, A. T., Levin, M., Yap, K., Fraser, M. I., Sahdra, B. K., Hofmann, S. G., & Hayes, S. C. (2024). Process-based therapy: A common ground for understanding and utilizing therapeutic practices. Journal of Psychotherapy Integration. https://doi.org/10.1037/int0000348

Ciarrochi, J., Sahdra, B., Fraser, M. I., Hayes, S. C., Yap, K., & Gloster, A. T. (2024). The compassion connection: Experience sampling insights into romantic attraction. *Journal of Contextual Behavioral Science, 32*, 100749. https://doi.org/10.1016/j.jcbs.2024.100749

Ciarrochi, J., Sahdra, B., Hayes, S. C., Hofmann, S. G., Sanford, B., Stanton, C., Yap, K., Fraser, M. I., Gates, K., & Gloster, A. T. (2024). A personalised approach to identifying important determinants of well-being. *Cognitive Therapy and Research*. https://doi.org/10.1007/s10608-024-10486-w

Ciarrochi, J., Sahdra, B., Hofmann, S. G., & Hayes, S. C. (2022). Developing an item pool to assess processes of change in psychological interventions: The Process-Based Assessment Tool (PBAT). *Journal of Contextual Behavioral Science,* 23, 200–213. https://doi.org/10.1016/j.jcbs.2022.02.001

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology, 62*(1), 583–619. https://doi.org/10.1146/annurev.psych.093008.100356

Daly, M., Sutin, A. R., & Robinson, E. (2022). Longitudinal changes in mental health and the COVID-19 pandemic: Evidence from the UK Household Longitudinal Study. *Psychological Medicine, 52*(13), 2549-2558. https://doi.org/10.1017/S0033291720004432

Epskamp, S., Deserno, M., Bringmann, L., Veenman, M. (n.d.). mlVAR: Multi-Level Vector Autoregression (Version 0.5.1). https://CRAN.R-project.org/package=mlVAR

Epskamp, S., van Borkulo, C. D., van der Veen, D. C., Servaas, M. N., Isvoranu, A.-M., Riese, H., & Cramer, A. O. J. (2018). Personalized network modeling in psychopathology: The importance of contemporaneous and temporal connections. *Clinical Psychological Science, 6*(3), 416–427. https://doi.org/10.1177/2167702617744325

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin, 53*(2), 134140. https://doi.org/10.1037/h0045156

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, 115(27). https://doi.org/10.1073/pnas.1711978115

Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders, 208*, 191–197. https://doi.org/10.1016/j.jad.2016.10.019

Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders, 172*, 96–102. https://doi.org/10.1016/j.jad.2014.10.010

Gates, K. M., Chow, S. M., & Molenaar, P. C. (2023). *Intensive longitudinal analysis of human processes.* Chapman and Hall/CRC.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge University Press.

Gratz, K. L., & Roemer, L. (2004). Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure, and Initial Validation of the Difficulties in Emotion Regulation Scale. Journal of Psychopathology and Behavioral Assessment, 26(1), 41–54. https://doi.org/10.1023/b:joba.0000007455.08539.94

Gu, J., Strauss, C., Bond, R., & Cavanagh, K. (2015). How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. *Clinical Psychology Review, 37*, 1-12. https://doi.org/10.1016/j.cpr.2015.01.006

Hayes, S. C. & Hofmann, S. G. (2021). "Third-wave" cognitive and behavioral therapies and the emergence of a process-based approach to intervention in psychiatry. *World Psychiatry, 20*(3), 363-375. https://doi.org/10.1002/wps.20884

Hayes, S. C., Hofmann, S. G. & Ciarrochi, J. (2020). A process-based approach to psychological diagnosis and treatment: The conceptual and treatment utility of an extended evolutionary model. *Clinical Psychology Review, 82*, 101908. https://doi.org/10.1016/j.cpr.2020.101908

Hayes, S. C., Hofmann, S. G., Stanton, C. E., Carpenter, J. K., Sanford, B. T., Curtiss, J. E., & Ciarrochi, J. (2019). The role of the individual in the coming era of process-based therapy. *Behaviour Research and Therapy, 117*, 40-53. https://doi.org/10.1016/j.brat.2018.10.005

Hayes, S. C., Nelson, R. O. & Jarrett, R. (1987). Treatment utility of assessment: A functional approach to evaluating the quality of assessment. *American Psychologist, 42*, 963-974. https://doi.org/10.1037//0003-066X.42.11.963

Hayes, S. C., Sahdra, B. K., Ciarrochi, J., Hofmann, S. G., & Sanford, B. (2025). How a process-based idionomic approach changes our understanding of mindfulness as a method and process. Chapter to appear in K. W. Brown, J. D. Creswell, & R. M. Ryan (Eds.), *Handbook of Mindfulness* (2nd ed.): *Theory, research, and practice.* Guilford Press.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2022). *Cochrane handbook for systematic reviews of interventions* (ver 6.3; updated February 2022).

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ , 327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hoekstra, R. H. A., Epskamp, S., Nierenberg, A. A., Borsboom, D., & McNally, R. J. (2024). Testing similarity in longitudinal networks: The Individual Network Invariance Test. *Psychological Methods.* https://doi.org/10.1037/met0000638

Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*(2), 169–183. https://doi.org/10.1037/a0018555

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology, 3*, 29 –51. Doi: https://doi.org/10.1146/annurev.clinpsy.3.022806.091419

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice.* (2nd ed.) OTexts. https://otexts.org/fpp2/

Hyndman, R. J., & Khandakar, Y. (2008). *Automatic time series forecasting: The forecast Package for R. Journal of Statistical Software, 27*(3), 1–22. https://doi.org/10.18637/jss.v027.i03

IntHout, J., Ioannidis, J. P., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open, 6*(7), e010247. https://doi.org/10.1136/bmjopen-2015-010247

Ioannidis, J. P. A., Patsopoulos, N. A., & Evangelou, E. (2007). Uncertainty in heterogeneity estimates in meta-analyses. *BMJ , 335*(7626), 914–916. https://doi.org/10.1136/bmj.39343.408449.80

Kabat-Zinn J. (2017) 'Too early to tell: the potential impact and challenges—ethical and otherwise—inherent in the mainstreaming of dharma in an increasingly dystopian world', *Mindfulness 8*(5), 1125–1135. https://doi.org/10.1007/s12671-017-0758-2

Kabat-Zinn, J. (2013). *Full catastrophe living (Revised Edition): Using the wisdom of your body and mind to face stress, pain, and illness.* Random House.

Kamphuis, J. H., Noordhof, A., & Hopwood, C. J. (2021). When and how assessment matters: An update on the Treatment Utility of Clinical Assessment (TUCA). *Psychological Assessment, 33*(2), 122–132. https://doi.org/10.1037/pas0000966

Kashkouli Nejad, K., Sugiura, M., Thyreau, B., Nozawa, T., Kotozaki, Y., Furusawa, Y., Nishino, K., Nukiwa, T., & Kawashima, R. (2014). Spinal fMRI of interoceptive attention/awareness in experts and novices. *Neural Plasticity, 2014*, 679509. https://doi.org/10.1155/2014/679509

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The Patient Health Questionnaire-2. *Medical Care, 41*(11), 1284–1292. Doi: https://doi.org/10.1097/01.mlr.0000093487.78664.3c

Kruis, J., & Maris, G. (2016). Three representations of the Ising model. *Scientific Reports, 6*(1). https://doi.org/10.1038/srep34175

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics, 54*(1-3), 159–178). https://doi.org/10.1016/0304-4076(92)90104-y

Lamiell, J. T. (1981). Toward an idiothetic psychology of personality. *American Psychologist, 36*(3), 276-289. https://doi.org/10.1037/0003-066X.36.3.276

Liu, S., Kuppens, P., & Bringmann, L. (2021). On the use of empirical bayes estimates as measures of individual traits. *Assessment, 28*(3), 845–857. https://doi.org/10.1177/1073191119885019

Lu, Y., & Rodriguez-Larios, J. (2022). Nonlinear EEG signatures of mind wandering during breath focus meditation. *Current Research in Neurobiology, 3*, 100056. https://doi.org/10.1016/j.crneur.2022.100056

Lutz, A., McFarlin, D. R., Perlman, D. M., Salomons, T. V., & Davidson, R. J. (2013). Altered anterior insula activation during anticipation and experience of painful stimuli in expert meditators. *NeuroImage, 64*, 538–546. https://doi.org/10.1016/j.neuroimage.2012.09.030

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46(4)*, 806-834. https://doi.org/10.1037//0022-006x.46.4.806

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives,* 2(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Molenaar, P. C. M. (2015), On the relation between person-oriented and subject-specific approaches. *Journal for Person-Oriented Research, 1*(1-2), 34-41. https://doi.org/10.17505/jpor.2015.04

Moritz, S., Ciarrochi, J., Fraser, M., Krafft, J., Klimczak, K., Levin, M., Hernández, C. E., Hayes, S. C., Yap, K., & Sahdra, B. K. (2024). An idionomic analysis of mindfulness, experiential avoidance, and affect among individuals exhibiting high and low-hoarding behaviours. Manuscript in preparation.

Nesselroade, J. R. (2001). Intraindividual variability in development within and between individuals. *European Psychologist, 6(3)*, 187-193. https://doi.org/10.1027//1016-9040.6.3.187

Nesselroade, J. R., Gerstorf, D., Hardy, S. A., & Ram, N. (2007). Idiographic filters for psychological constructs. *Measurement: Interdisciplinary Research and Perspectives, 5*(4), 217–235. https://doi.org/10.1080/15366360701741807

Olsson, T. M., & Fridell, M. (2018). The five-year costs and benefits of extended psychological and psychiatric assessment versus standard intake interview for women with comorbid substance use disorders treated in compulsory care in Sweden. *BMC Health Services Research, 18,* 53. https://doi.org/10.1186/s12913-018-2854-y

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (2nd ed.). Sage Publications.

Sahdra, B. K., Ciarrochi, J., Fraser, M. I., Yap, K., Haller, E., Hayes, S. C., Hofmann, S. G., & Gloster, A. T. (2023). The compassion balance: Understanding the interrelation of self- and other-compassion for optimal well-being. *Mindfulness, 14*(8), 1997-2013. https://doi.org/10.1007/s12671-023-02187-4

Sahdra, B. K., Ciarrochi, J., Klimczak, K. S., Krafft, J., Hayes, S. C., & Levin, M. (2024). Testing the applicability of idionomic statistics in longitudinal studies: The example of "doing what matters." *Journal of Contextual Behavioral Science, 32*, 100728. https://doi.org/10.1016/j.jcbs.2024.100728

Sahdra, B. K., Ciarrochi, J., Parker, P. D., Basarkod, G., Bradshaw, E. L., & Baer, R. (2017). Are people mindful in different ways? Disentangling the quantity and quality of mindfulness in latent profiles and exploring their links to mental health and life effectiveness. *European Journal of Personality, 31*(4), 347-365. https://doi.org/10.1002/per.2108

Sala, M., Rochefort, C., Lui, P. P., & Baldwin, A. S. (2019). Trait mindfulness and health behaviours: A meta-analysis. *Health Psychology Review, 14*(3), 345–393. https://doi.org/10.1080/17437199.2019.1650290

Sanford, B. T., Ciarrochi, J., Hofmann, S. G., Chin, F., Gates, K. M., & Hayes, S. C. (2022). Toward empirical process-based case conceptualization: An idionomic network examination of the process-based assessment tool. *Journal of Contextual Behavioral Science, 25,* 10–25. https://doi.org/10.1016/j.jcbs.2022.05.006

Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology, 31*(1), 43–53. https://doi.org/10.1016/j.newideapsych.2011.02.007

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology.* Basic Books.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* Oxford University Press.

Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods, 20*(1), 63-83. https://doi.org/10.1037/met0000030

Young, Q. R., Ignaszewski, A., Fofonoff, D., & Kaan, A. (2007). Brief screen to identify 5 of the most common forms of psychosocial distress in cardiac patients: Validation of the screening tool for psychological distress. *Journal of Cardiovascular Nursing, 22*(6), 525-534. https://doi.org/10.1097/01.jcn.0000297383.29250.14