



BEVI Manual:
Beliefs, Events, and Values Inventory (BEVI)
<https://thebevi.com/>

This BEVI Manual provides relevant information for individuals who are trying to understand what the BEVI is and/or are considering using or taking the BEVI for various purposes. Because this manual is written for a wide range of audiences, individuals may be interested in specific topics, or all topics, across the following ten hyperlinked areas, which are organized in terms of complexity (from relatively basic to increasingly complex):¹

- I. [About the BEVI](#)
- II. [BEVI Usage](#)
- III. [BEVI Rationale](#)
- IV. [BEVI Structure, Reliability, and Validity](#)
- V. [Reviewing the BEVI](#)
- VI. [Taking the BEVI](#)
- VII. [Frequently Asked Questions About the BEVI](#)
- VIII. [Guiding Principles and Best Practices of Assessment](#)
- IX. [Using the BEVI in Research and Practice](#)
- X. [Further Information About the BEVI](#)

I. About the BEVI

The BEVI is an accessible, adaptable, and powerful analytic tool that may be used across multiple populations and settings – from education and research to leadership and mental health – to evaluate, understand, and facilitate processes and outcomes of learning, growth, and transformation. Based upon over 30 years of research and practice in the U.S. and internationally – and with excellent psychometric properties – the BEVI asks respondents a series of questions about beliefs, values, and life events which seek to illuminate “who learns what and why, and under what circumstances.” A grounded theory and mixed methods measure, including both quantitative and qualitative items, the BEVI was developed on the basis of hundreds of actual “belief” statements from individuals all over the world.

The most recent version of this measure – BEVI-3 – has been normed upon an international sample of over 10,000 individuals representing over 100 different countries. The BEVI uses a secure, web-based system that takes about 30 minutes to complete, while generating an array of reports for individuals, couples, families, groups, organizations, and institutions. In addition to BEVI-based coaching, consulting, and counseling, results may also be discussed with "[Being Bevi,](#)" the BEVI's AI / VR entity that lives within the “Beviverse.”

¹ Aspects of this document are excerpted and/or adapted from Shealy, C.N. (Ed.). (in press). *Cultivating the globally sustainable self: How the human species might fulfill its potential*. Oxford University Press.

II. BEVI Usage

There have been over 150,000 administrations of the BEVI across a wide range of countries and cultures. From government and policy agencies, to educational, healthcare, and community organizations, to research and technology systems, to corporate leadership teams and non-profit boards, the BEVI has been used for an array of purposes by individuals, couples, families, groups, communities, organizations, businesses, and institutions. For example, the BEVI and its reporting / AI / VR systems have been used to:

- Evaluate and facilitate learning, growth, and development in educational contexts
- Show how leaders and organizations can become more effective
- Assess and enhance mental health and wellbeing
- Address professional development, team building, and strategic planning objectives
- Advance programmatic research across a wide range of disciplines
- Empirically illuminate the forces and factors that “cause” change to occur
- Pursue personal and professional goals such as greater self-awareness, emotional intelligence, critical thinking, sociocultural understanding, collaborative problem solving, and interpersonal skills
- Help resolve conflict among individuals and groups
- Demonstrate and improve the quality of programs and interventions
- Comply with accreditation, certification, or regulatory requirements
- And much more

III. BEVI Rationale

Individuals, groups, institutions, and organizations use the BEVI for many reasons, such as the following:

First, it is comprehensive, using a mixed methods format (e.g., quantitative and qualitative items) measuring multiple aspects of learning, growth, and development including, but not limited to, basic openness; receptivity to different cultures, religions, and social practices; the tendency (or not) to stereotype in particular ways; self and emotional awareness; and preferred but implicit strategies for making sense of why “other” people and cultures “do what they do.”

Second, it is empirically validated and theoretically grounded, with a demonstrable track record for over 30 years of international development and usage across a diverse array of countries and populations, with excellent psychometric properties (e.g., reliability, validity, scale structure) and an integrative and interdisciplinary theoretical framework. Along with other relevant information and materials, peer reviewed articles and chapters are available via the Admin section of the BEVI website at <https://thebevi.com/>.

Third, it is highly accessible and adaptable, with a secure and on-line system of administration, which requires about 30 minutes to complete and provides individual, group, institutional, and cross-institutional report and analysis options, and the ability to add items and/or be used in conjunction with other assessment measures.

Fourth, it includes both evaluative and practical applications, which can not only examine processes of change within and between individuals, groups, organizations, and institutions, but also facilitate learning, growth, and development through a sophisticated and personally tailored report system as well as the “Beviverse,” an AI-mediated system that allows individuals to “walk into their own self” via their own BEVI profiles within an engaging VR (virtual reality) world (see <https://researchfeatures.com/into-beviverse-why-beliefs-values-matter/>).

IV. BEVI Structure, Reliability, and Validity

The BEVI is a statistically reliable, well validated, and standardized measure that is used in a wide array of settings and contexts around the world. The BEVI consists of four interrelated components: 1) a comprehensive set of demographic/background items that may be modified for particular projects; 2) a life history questionnaire, which is built into the measure; 3) two validity and seventeen "process scales," comprised of 185 quantitative items answered on a four-point Likert-type scale (Strongly Agree, Agree, Disagree, Strongly Disagree); and 4) three qualitative "experiential reflection" items (please see <https://thebevi.com/about/> for a complete listing of BEVI scales).

Over the past 30 years, there have been three major iterations of the BEVI: 1) the original version of the BEVI, which included over 400 items; 2) the BEVI-2 long and short versions; and 3) the BEVI-3, which is standardized upon an N of over 10,000 individuals representing over 100 countries. When standardizing a measure, the overarching goal is to draw a sample from the population as a whole in as representative a manner as possible. Since the BEVI is used globally, the standardization sample for the BEVI deliberately drew upon a very wide and diverse range of respondents, a process that included randomized selection of the normative sample. The BEVI also has been translated from English into multiple languages, an intensive process that not only involves “back translation,” but procedures to ensure “equivalency of meaning” across languages and cultures. To read more about relevant processes in this regard, excellent resources include Geisinger's multi-volume [APA Handbook of Testing and Assessment in Psychology](#) as well as guidelines provided by the [International Test Commission](#).

To ensure data integrity, the BEVI deliberately includes two separate scales (Consistency and Congruency) as well as a variety of other indices that are built into the measure. As is discussed during administration / training processes with the BEVI, the issue is not whether someone can try to "fake out" the BEVI or any other measure, but rather whether or not such "faking" is detectable by that measure. Based upon past analytic work, the vast majority of respondents do not try to "fake out" the BEVI, and if they do, their report is automatically not included in any group report. It is best practice to include at least some type of validity measure in instruments

of this nature, and the BEVI includes multiple such indices. Ultimately, validity should be understood not just for its own sake (e.g., does an instrument measure what it says it measures), but also in the context of real world applications, which address matters of usefulness, etc. (e.g., [Hanson et al., 2023](#)). Please see <https://thebevi.com/about/validity/> for more information.

V. **Reviewing the BEVI**

One of the more interesting aspects of the BEVI is that people may believe this measure can be understood and/or be evaluated by reviewing items at a face valid level (i.e., thinking the BEVI can or should be understood simply by looking at each BEVI item “on the face of it”). By design, the BEVI as a whole is deliberately *not* face valid, a design feature that helps mitigate response set confounds such as social desirability (these and other psychometric and instrument design issues are discussed in depth during BEVI training and certification). Among other problems with thinking that the BEVI can be understood simply by reviewing BEVI items, individuals may project themselves into the heads of others (e.g., students, participants, colleagues, clients), imagining what they will feel or think as they take the BEVI, thereby assuming their own experience will be identical, universal, or sufficiently similar to the experience that others will have when they take this measure. *All sorts of attributions may arise during such a process, which is the very point of the measure.*

In other words, one's reaction to the BEVI is a direct result of one's own beliefs, values, schemas, attitudes, and worldview, which arise – from the perspective of the “equilintegration” or EI theoretical model underlying the BEVI (see <https://thebevi.com/about/eitheory/>) – out of one's own formative variables (e.g., life history, culture, gender, educational level, religious / political background, etc.) interacting with one's own core needs (e.g., affective, acknowledgement, affiliative), which are then filtered further through the identities / roles we assume (e.g., as researchers, educators, students, clinicians, clients, administrators, leaders, consultants, coaches, etc.). The fact that individuals may experience the BEVI differently (e.g., Person A “Strongly Agrees” with BEVI item X whereas Person B “Strongly Disagrees” with BEVI item X), says more about one’s own self structure than it does about the measure itself, which is officially neutral. In fact, the BEVI takes no position regarding the “rightness” or “wrongness” of any BEVI item. *Again, the individual / differential nature of such projections onto BEVI items is the very point and purpose of this measure.* That is because those experiences tell and show us (through individual / group reports and the AI / VR systems) how we are structured at these core levels of identity and self from the standpoint of the BEVI and its underlying EI theoretical framework.

As noted previously, BEVI items emerged from belief statements from real people all over the world, and have been extensively reviewed, not only statistically (e.g., through factor analyses, SEM, IRT), but by multiple groups of Subject Matter Experts (SMEs) (e.g., to ensure that items tend to fall within optimal experiential bounds, aren't inherently biased, etc.). If we somehow tried to neutralize or hygienize the subjective / affective / attributional / historical dimensionality of BEVI items, we would essentially eliminate the very purpose of the BEVI itself, which is to

"tap into" how and why we experience self, others, and the larger world as we do. It is essential that reviewers, administrators, or users of the BEVI understand these fundamental features of the BEVI *before* reviewing or taking the measure and/or evaluating its potential suitability for various projects (e.g., studies, courses, programs, grants, workshops, interventions, etc.).

VI. Taking the BEVI

The BEVI is an objective measure that functions in a projective manner. That is because people who take the measure are projecting their own meaning onto stimuli – 185 belief statements – to which they may strongly agree, agree, disagree, or strongly disagree (for reasons that are explained in BEVI training, there is no “undecided” option). After completing a series of background questions, participants are asked to respond to items covering a very wide range of issues and topics. People have different reactions to different items. Some of these items may seem straightforward whereas others may seem ambiguous. Because the BEVI is assessing many complex and interrelated factors, all of these items as well as the background questions are relevant to understanding how people experience self, others, and the larger world.

Although the BEVI is officially neutral regarding whether specific beliefs are true or false, or good or bad, human beings are not. As such, when encountering items on the BEVI, all sorts of thoughts and feelings may arise for the person experiencing them. That is why BEVI Administrators explicitly provide an orientation to prospective takers of the BEVI to 1) explain more about the measure (e.g., how long it takes; there are no “right” or “wrong” answers), 2) describe administration processes (e.g., the voluntary nature of completing the BEVI; informed consent; how they access the BEVI), 3) provide relevant context (e.g., what is the purpose of the BEVI), and 4) explain what participants may gain by completing the measure (e.g., access to their individual report; the opportunity to engage BEVI-AI).

For some people, the very idea of an instrument that measures “beliefs and values” can elicit reactions, from “that’s really cool” to “should such a measure even exist.” That is because [humans hold a range of beliefs about beliefs](#). Oftentimes, it is not even clear to people that they have such beliefs until they are faced with a measure that asks us to reflect upon them. It is only at that point (e.g., when they are introduced to the idea of the BEVI or are actually taking the measure) that they often begin to wonder about their own beliefs and values or the prospect of measuring them in the first place.

Because people may not rationally know or be able to acknowledge that they are having such an experience of their own beliefs and values, or those of others, all sorts of questions and reactions may emerge when taking the BEVI, a process that may or may not be consciously accessible. This experience can be a bit disconcerting, at least initially, but also can be highly compelling. For example, in the context of discussing BEVI theory, findings, or applications – or engaging the BEVI’s reporting and AI / VR systems – it is very common for people to say that the experience is “fascinating” or “very interesting.” Before people get to that point, however, they also may have other questions related to the BEVI, which are discussed next.

VII. Frequently Asked Questions About the BEVI

From the standpoint of the BEVI Team – <https://thebevi.com/about/team/> – based upon our experience over many years, here are ten fundamental questions we often receive from people who are wondering about, or have taken, the BEVI along with the below responses and explanations we offer. These and other questions are discussed during processes of BEVI training and certification. The materials referenced in this manual (e.g., various links) are also likely to provide answers to other questions you may have. If you have further questions, or need additional information, please contact us directly by [clicking here](#).

BEVI Question 1

“Shouldn’t I be able to understand the BEVI by taking it or reviewing the items on it?”

No. If you can tell what a measure is measuring simply by taking it – or by reviewing the items on it – it may not be a very good measure, particularly if what you are doing (e.g., in research, education, practice, etc.) is subject to response set confounds like [social desirability](#) (e.g., answering survey questions in a way that is perceived to be positive by others). Such measures (i.e., those that can be understood by looking at them or taking them) are often “face valid.”

By design, and based upon extensive research and development, the BEVI is NOT face valid (i.e., it typically cannot be understood just by reviewing the items on it or by taking it). This issue is one of the most complex and sometimes vexing about assessment in general, including but not limited to the BEVI. That is because it defies what seems to be a “common sense” notion that one should be able to understand what a measure is measuring by looking at it or taking it. In fact, the opposite is true, and here is why.

Basically, if a test taker can tell what a measure is measuring by taking it or reviewing items on it, the measure may have serious problems with validity. That is because the test may not be measuring the phenomenon it purports to measure, but rather the ability and/or desire of the individual test taker to give responses they believe are better or more desirable from the standpoint of the test. This is a serious [“confound”](#) – an unintended and unaccounted source of measurement error – that too often is not understood in assessment research and practice, even by people who are otherwise trained in statistics or research design, as we have often experienced in the real world by reviewers of the BEVI (i.e., people who are in positions to review measures for various projects – or even in charge of research that uses assessment – may not understand these fundamental aspects of psychometric theory, research, and practice).

There may be good reasons for a measure to be comprised of “face valid” items, especially when there is no evidence that test-takers might prefer a particular result from a measure. It really depends upon the purpose of the test as well as the underlying motivations of the test-taker. Consider these issues in the context of mental health assessment. For all sorts of reasons, someone may want to receive – or not receive – a particular diagnosis (e.g., to gain access to a

desired medication; to be perceived in a particular way; to receive benefits; to eliminate obstacles; to gain privileges, access, or status). If such an individual can easily tell what the measure is asking (e.g., an instrument or interview) – because the items or questions are face valid – the “measure” may not be measuring whether someone actually meets criteria for a putative diagnosis, but rather the ability of the individual to present him or herself in a way that is more likely to receive, or not receive, the diagnosis.

To take another real world example, consider what happens when someone’s worth or performance is being evaluated vis-à-vis 1) how open or sensitive they are perceived to be regarding differences among people (e.g., ethnicity, gender) or 2) their degree of fidelity to particular values or ways of thinking, feeling, or being (e.g., religious, political). If the individual knows or suspects that particular expectations are operative in this regard, they may be more likely to speak or behave in ways that comport with *what they think the “better” or “correct” answers are* rather than what they *actually think and feel*. The issue here is not that we may have very good reasons for wanting people to believe, feel, or act in a particular manner. We may. Rather, the issue is whether we can 1) directly measure such thoughts, feelings, and actions accurately *or* if 2) what we are measuring is the relative capacity and inclination of people to tell or show us what they think we want to hear or see rather than what they actually believe and value.

[Social desirability is not bad. It is integral to being human.](#) However, and in summary, if an approach to measurement (e.g., an instrument, interview) is subject to response set confounds like social desirability – because its items or questions are face valid – then the measure may not be measuring what it says it measures, but rather the very human desire to be experienced by others in ways that are perceived to be good. Within the field and profession of assessment (e.g., test development, psychometrics, research design, data analysis, test interpretation), such matters are not abstractions, but rather at the very core of whether instruments are valid and results are useful (see research on [malingering](#), for example; see also [Hanson et al., 2023](#)).

As noted above, based upon considerable research and development, the BEVI is not face valid, an important characteristic of this measure. This core feature of the BEVI may be challenging to understand, which is why training in the BEVI is required for all administrators.

BEVI Question 2

“Does the BEVI take a position on whether beliefs are true or false or good or bad?”

No. There are four response possibilities to each BEVI item: strongly agree, agree, disagree, and strongly disagree. So no matter how a specific item is worded – and remember, these are actual statements that were made by a very diverse group of people from many different countries and cultures – people can agree or disagree as they wish. It can be perplexing to encounter items that one believes are, or should be, obvious to, or agreed upon, by everyone (i.e., “Surely everyone must know that belief X is true and belief Y is false”). In reality, [humans can believe](#)

[anything about anything](#). As such, the goal of the BEVI is to illuminate the unique organizational structure of each person's belief / value system so that they can make better sense of it for themselves and use such information to identify and pursue goals that matter most to them.

BEVI Question 3

“Because humans are so complex and diverse, is it really possible for a single measure to tell you anything about different individuals or groups from around the world?”

Yes. BEVI items emerged from statements made by a very diverse group of people from many different countries and cultures over a long period of time regarding all manner of experiences and perceptions about, self, others, and the larger world (i.e., items were identified by actual belief statements from real people, rather than an idea or theory regarding what beliefs and values are or should be). Through an array of statistical analyses (e.g., EFA, CFA, SEM, IRT) using different samples of BEVI data, only the most powerful items (i.e., predictive / explanatory) were retained on the BEVI. Some of these items are deceptively simple whereas others reflect interactions among multiple dimensions that are embedded in a single item. Because BEVI items are comprised of real statements made by real people having real experiences in the world, BEVI items typically have a multifaceted character, including attributional, cultural, developmental, emotional, nonconscious, and subjective properties.

Over the years, following 1) [best practices and guiding principles of test development](#), 2) from a normative sample of over 10,000 individuals representing over 100 countries around the world, and 3) based upon hundreds of analyses and thousands of reports, we know that the BEVI is able to understand, distinguish between, and predict all manner of phenomena regarding individuals and groups in a valid and reliable manner from very different backgrounds and contexts. Relevant findings and applications in this regard have also been [published and presented in hundreds of refereed forums and other professional settings](#). No measure is perfect, including the BEVI, and we are committed to ongoing research, development, and refinement of this measure over time. But, abundant data and applications – all over the world – demonstrate that the BEVI is in fact able to illuminate a whole lot about who we are, what makes us similar and different, and why that matters in our work, relationships, and lives.

BEVI Question 4

“What is it like to take the BEVI?”

Because the BEVI is asking about what you believe – while also taking into consideration your background and life history – the experience of taking the BEVI is an experience of encountering yourself. As a result, when you take the BEVI, the identity commitments we all have are activated in an optimally calibrated manner, based upon multiple review processes over the years. Although this process may seem rational (e.g., “I am engaging in an objective analysis”), in reality, it is very difficult, if not impossible, to eliminate one's subjective and

emotional reactions to various BEVI items. That is because such reactions are co-determined to no small degree by one's own background and life history. Which items evoke what reactions – and how and why we are different or similar in that regard – is precisely the point of the measure.

The BEVI is able to evaluate and facilitate change because of its origins and design. However, when taking the BEVI, these important features – from a psychometric standpoint – also may activate an experience of one's underlying identity and self structure, which may not otherwise be accessible on a logical or conscious basis. That is why the idea of the BEVI, much less the BEVI itself, may evoke “resistance,” by evoking self structures and identity commitments, which may be projected onto particular BEVI items or the very concept of the BEVI itself.

Of course, there are many good and legitimate questions to ask about the BEVI, or any measure, and we always welcome these sorts of conversations. But, it's important to understand that one's reaction to the BEVI may have nothing at all to do with the BEVI per se, but rather the concept of the BEVI and/or one's beliefs about beliefs (e.g., how beliefs can, or whether they should, be measured in the first place).

BEVI Question 5

“Can I take the BEVI?”

Yes. Basically, there are two ways you can take the BEVI.

First, if you are interested in taking the BEVI yourself (e.g., for learning, growth, or development purposes; in the context of a situation you want to understand better in your work, relationships, or life; out of curiosity), a qualified member of the BEVI Team will want to have a brief and confidential meeting with you in order to understand your interests and goals, to answer questions about the BEVI and its reporting / AI / VR systems. In addition to follow-up debriefing after you have completed the BEVI, we also provide BEVI-based coaching, consulting, and counseling with certified coaches, experienced consultants, and licensed clinicians.

Second, if you are interested in the possibility of using the BEVI for a particular purpose / in a given setting, that's great. In fact, we encourage potential users to review and/or take the BEVI, but only after they receive training in the basics of test development and assessment research and practice as well as the underlying theory of the BEVI in order to understand why this measure is designed as it is. Without such understanding (e.g., of depth-based and mixed methods assessment), the BEVI may be misunderstood or misused. However, even without training, potential users of the BEVI can still gain a good understanding of this measure by reviewing its design and structure, including [BEVI scales and sample items](#). Other sources of information about the BEVI include numerous publications as well as publicly available resources and initiatives (e.g., see the many links in this BEVI Manual).

If you are interested in taking the BEVI, or participating in BEVI-based coaching, consulting, or counseling, please contact us directly by [clicking here](#).

BEVI Question 6

“Why does the BEVI ask about my life history and past experiences?”

The BEVI asks these questions because we are trying to understand not only what you believe and value, but why. Through a wide array of statistical procedures (e.g., analysis of variance, regression, structural equation modeling), we have demonstrated that [there often is a strong relationship between one's life history, past experiences, and various scales on the BEVI](#). That is why we include both a background information / demographics section on the BEVI as well as a scale that explicitly assesses life history and experiences. That is also why the BEVI is called the “Beliefs, *Events*, and Values Inventory” not the “Beliefs and Values Inventory.” If we eliminated either the life history portion of the BEVI or the demographic / background section of the BEVI, we would only be able to describe what you believe and value. That is important information for sure. But where the BEVI becomes really interesting – and meaningful – is when it is able to offer explanations and interpretations regarding 1) why you believe and value what you do, 2) how those beliefs and values may impact your work, relationships, and life, and 3) what might be useful for you to consider in terms of getting your needs met and pursuing your goals. From the standpoint of research and practice, it also should be noted that it is possible to add or remove specific demographic / background items if needed for a particular purpose, project, or population.

BEVI Question 7

“Why does the BEVI ask me to answer questions about ‘controversial’ topics?”

Think about the conversations we have with family, friends, and colleagues where we find ourselves listening to, or expressing, various points of view. Topics could be weighty – from religion and politics to culture and gender – or mundane, regarding basic matters of the day. During these dialogues, we might offer our thoughts on why we, or others, do what we do; why events occur or circumstances are as they are; or how things should or should not be. For all sorts of reasons, we may or may not say what we “really think” out loud. Likewise, we might feel very strongly about various points and perspectives, or have no opinion at all.

Oftentimes, these verbal exchanges manifest as “belief / value” statements about the nature of the world and the people in it, including ourselves. These [“versions of reality”](#) are exactly what the BEVI measures. Spoken by a wide range of individuals from all over the world, these belief / value statements became questions – what we call “BEVI items” – which were then subjected to multiple statistical analyses and review processes over many years to determine which items were most predictive and useful. When we take the BEVI, we are assembling these items – like pieces of a puzzle – into a unique portrait of ourselves in terms of what we believe is most

important, real, or true. We know from [considerable evidence](#) that the answers we provide to BEVI items illuminate a great deal about who we say we are, and why we do what we do.

Since all BEVI items are statistically predictive, representing interrelated parts of a larger whole, upon what basis would we eliminate one or more of them? Remember that takers of the BEVI can strongly agree, agree, disagree, and strongly disagree with every BEVI item. An item that might feel “controversial” to one person is accepted as “a matter of fact” by another. That is the point of the measure, to 1) understand why we believe what we believe, 2) why that matters in our work, relationships, and lives, and 3) how we can use such information to help us better pursue our potential, interests, and goals.

In the final analysis, remember that the BEVI asks the questions it does because these are real statements made by real people from very diverse countries and cultures. From a statistical standpoint, and in the real world, it turns out that these “belief / value statements” – our “versions of reality” – actually matter. A lot. So, take the BEVI or don’t take the BEVI. Use the BEVI or don’t use the BEVI. It’s completely up to you. But that’s why the BEVI asks the questions it does. If you’re interested in the meaning of our answers – and the implications and applications that follow – then the BEVI may be meaningful for you.

BEVI Question 8

“Can I get access to BEVI-AI, Being Bevi, and the Beviverse?”

Yes. Depending on your arrangement with the BEVI, and what type of engagement you want, the BEVI’s AI / VR systems are accessible to anyone who takes the BEVI. That means you are able to have one or more confidential conversations with “Being Bevi” – the AI entity that lives within the “Beviverse,” the BEVI’s virtual world – on the basis of your unique BEVI profile.

Individuals, couples, and groups ask Being Bevi all manner of questions about themselves, others, and the larger world, from the personal to the professional, including, but by no means limited to:

- Why do I do what I do?
- How can I improve my relationships?
- Why do others perceive me as they do?
- How can I find the most compelling educational or career path?
- How can I pursue my potential and become better at what I do?
- Could I learn to resolve conflicts or solve problems more effectively?
- How might I enhance my wellbeing and mental health?
- Can I find greater meaning and purpose in my life?
- How does all of this relate to the challenges we face in the larger world?

To learn more about this fascinating process and experience, please see ["Into the Beviverse."](#)

BEVI Question 9

“How do I know if the BEVI is the right measure for me?”

Depending upon your interests and goals, a good question often is not “*What is the right measure?*,” but rather, “*What do I want to measure?*” If you are interested in 1) “going deep” into the underlying factors that make us who we are, 2) how such factors impact interventions or experiences that are designed to create change, and/or 3) why this awareness can help make things better in our work, relationships, and lives, the BEVI may be a good instrument for you. Because it accounts for these complex and interacting variables in an ecologically valid (i.e., real world) manner – through sophisticated analysis, reporting, and AI / VR systems – the BEVI is able to evaluate and facilitate depth-based change across a highly diverse array of settings and populations. Moreover, please know that we are able to juxtapose data from other assessment measures as well. In both research and practice, we welcome those sorts of collaborations.

BEVI Question 10

“Is there anything else I should know about the BEVI?”

Yes. The BEVI is a comprehensive and mixed methods measure that is comprised of 17 scales, multiple indexes, and a wide array of analysis and engagement components, including reporting and AI / VR systems. All of these aspects of the BEVI can be used to evaluate and facilitate numerous processes and outcomes that are directly related to our work, relationships, and lives. To learn more, the BEVI website – <https://thebevi.com/> – provides a lot of good information. In addition, individuals who receive training on the BEVI also are provided access to the Admin section of the BEVI – <https://thebevi.com/login/> – which includes a wide range of materials, including peer reviewed publications on the BEVI, explanatory videos, administration guidelines, and other resources. Finally, please feel free to contact us directly by [clicking here](#), and a BEVI Team member will be in touch.

VIII. Guiding Principles and Best Practices of Assessment

As of this writing, there are over 150,000 administrations of the BEVI all over the world. Since the 1990s, this measure also has been reviewed by numerous individuals and systems, across many different countries and contexts, from all sorts of perspectives and for a wide range of projects (e.g., publications, presentations, IRBs, ethics boards, grant reviews, dissertation research, etc.). Along the way, we have learned a great deal about [guiding principles and best practices of applied assessment research](#). From our experience, assessment intentions and goals are typically good and worthy. However, at the level of design and delivery, much less analysis and interpretation, assessment efforts may fall short, squandering time, effort, and resources.

One of the main reasons assessment initiatives go wrong is a lack of understanding about one or more fundamental aspects of assessment research and practice, from the kinds of questions that

can actually be asked and answered, to what is required in terms of preparation and training, to the integration of assessment measures into project design and implementation, to data analysis and interpretation, to how assessment can be used not only for purposes of evaluation, but also to facilitate various processes and outcomes of learning, growth, and change. As such, it may be helpful to back up for a moment and address what we actually mean by “assessment” as well as what is required to “do assessment” in an appropriate manner, both locally and globally, and across different cultures and contexts (e.g., see [Hanson et al., 2023](#)).

At the most basic level, assessment is a technology, a methodology, and a paradigm for understanding and illuminating who human beings are, how they function, and whether – and under what circumstances – they change. To enter this world credibly, it is essential that those who use assessment measures understand the nature of assessment – what makes a “good” or “bad” measure – and how measures should, and should not, be used to impact real people in the real world. These considerations should be informed by factors that are inextricably linked to the design, functioning, and usage of any test, such as its rationale or purpose; its etiological, theoretical, and epistemological underpinnings; psychometrics and design characteristics; how data are reported or displayed; and appropriate usage or application of the measure.

Assessment is a well-established academic and professional field of inquiry and practice. You shouldn’t just “do assessment,” since it requires supervised training, sophisticated knowledge, and demonstrable skill. That said, it is neither necessary nor reasonable to expect individuals who are interested in assessment – or responsible for initiatives or activities where assessment occurs – to possess assessment knowledge or skill themselves, although that may well be the case. What is essential – by anyone involved with assessment coordination, consultation, or oversight – is to appreciate that *access to comprehensive assessment expertise is required* in order to engage in competent and effective assessment research and practice.

Among other areas of expertise, individuals or teams that assume assessment coordination, consultation, or oversight roles must be deeply familiar with multiple assessment-related domains, including, but by no means limited to, the fundamentals of assessment theory, research, and practice (e.g., ethics; psychometrics; statistical analysis; mixed methods; test usage); research design (e.g., treatment outcome studies; quasi-experimental design, “small N” research designs); mediators and moderators (e.g., underlying drivers / predictors of learning, growth, and development); professional fields that purport to facilitate change (e.g., transformative, engaged, multicultural, international, high impact); core concepts and practices that are relevant to group-based measurement (e.g., response set confounds; within-group variability); research review (e.g., IRBs; ethics; grants; guidelines; standards); and international usage / application of measurement (e.g., back translation; equivalency of meaning; data storage and protection), to mention only some of the core competencies that are integral to assessment research and practice, both locally and globally. These interrelated approaches and perspectives bear directly on matters of test development, usage, and interpretation and are addressed further through the various resources and links in this document (see [Assessment: The Power and Potential of Psychology Testing, Educational Measurement, and Program Evaluation Around the World](#)).

Oftentimes, people who are in charge of such projects may be content experts in a particular domain that is central to the initiative, but have little expertise in assessment research and practice. Not knowing what one doesn't know can compound the problem since the project leader may think there is "nothing to it," and/or have negative attitudes about assessment on the basis of their own past experiences or disciplinary / professional inclinations. Again, it is *not necessary* for the project leader to have this expertise personally, but it *is necessary* for the project leader to recognize that such expertise is essential, and therefore integral, to the project.

No less problematic, project leaders or participants may have very strong predispositions *toward* or *against* particular measures – for reasons that are not always well informed – which means they may not be open to additional or alternative instruments or approaches to measurement. These sorts of partialities manifest all of the time, with unfortunate implications for assessment research and practice (e.g., problems with project design and/or delivery; instrument selection and/or development; insufficient time set-aside for training or project preparation, such as IRB or ethics review; biases for or against quantitative or qualitative assessment, etc.).

It also should be recognized that there are multiple types of "assessment experts," so even when such individuals are brought into an assessment-based project, they may or may not have the competencies necessary to address issues, questions, or requirements like those noted above. For example, an assessment consultant may have expertise in organizational-level assessment, but no idea how to measure learning mediators at an individual or group level. As another real world example, someone may have deep expertise in statistics or research design, but little knowledge of how to actually use assessment instruments in the real world to facilitate change. Of course, if the project coordinator has little background in assessment, they will not be in a position to evaluate whether an assessment "expert" or "consultant" actually has the competencies necessary to provide guidance of the level and scope necessary for the project. The best solution to this dilemma is to expose one's assessment plan to a range of assessment experts, from different methodological, disciplinary, and epistemological backgrounds, to ensure that best practices and guiding principles are in fact understood and agreed upon, and subsequently integrated into project design, implementation, evaluation, interpretation, and reporting.

The overarching desire by project leadership is, understandably, to make the whole process as accessible and straightforward as possible. That is a worthy goal, but may be unrealistic when it comes to assessing the complexities that are inherent to the project, whether such complexities are recognized or not. In this regard, if a project coordinator isn't able or willing to account for human complexity vis-à-vis the assessment-based research and practice, the "results" that emerge are likely to be far less useful than they should be if not irrelevant or misleading. For example, if a project is designed to evaluate or facilitate learning, growth, or development, from the standpoint of measurement, it likely will be necessary to take human complexity into account as well as how that complexity interacts with the design and delivery of "interventions" that are ostensibly intended to catalyze change. In our work, we refer to these variables – what "causes" change to occur – as the "8Ds": duration, difference, depth, drive, determine, design, delivery, debrief. These variables interact, which should be accounted for in project measurement.

Empirically, theoretically, and in practice – based upon such applied assessment research for over 30 years – it is very clear that interactions among these variables are etiologically relevant to “who learns what and why, and under what circumstances,” which is why [we must account for these interacting variables if we truly want to create curricula, programs, or interventions that are demonstrably transformative](#). The processes by which we become structured as we are – and the degree to which we are subject to change – are often mediated by complex emotional, historical, and nonconscious processes that may be unknown to us (e.g., we may not be aware of how these factors and forces profoundly shaped our experience of self, others, and the larger world). As such, in many ways, the most important competency for good assessment research and practice is theoretical depth grounded in real world experience.

Good measurement flows from good theory which flows from good practice. If our theory is simple or unidimensional, because our practice is inadequate, the measures we develop or select will be simple or unidimensional. That is not a problem if the phenomenon we’re studying or attempting to change can be understood by a simple or unidimensional theory. In fact, simple and unidimensional might be exactly what is needed in that case. However, if the phenomena we are examining, or attempting to influence through various interventions, involve depth-based processes within and between human beings (e.g., learning, growth, change) – which are inextricably linked to, or emergent from, multifaceted phenomena (e.g., affect, attribution, context, culture, development, history, motivation) – our theoretical understanding, and measurement approach, must be able to account for this interacting and depth-based complexity, which means it must be seen – acknowledged, respected, understood – for what it is. In principle and practice, denial does not diminish complexity. Alternatively, [by embracing the ineluctable reality of human complexity, we are able to evaluate the promise and facilitate the potential of cultivating globally sustainable selves](#).

From the standpoint of designing and delivering various assessment-based initiatives, [it is quite extraordinary how persistent the funders, developers, and implementers of various change-oriented interventions can be in denying human complexity](#), particularly when vast expenditures of time, money, and resources are devoted to policies, practices, and initiatives that are destined to fail from the outset. We have experienced *all* of the above processes and dynamics, and many more. Hanson et al. (2023) offer a good summary of the relevant issues in this regard:

It is essential that competent developers and users of psychological tests and measures understand theories of assessment and evaluation, from different types of validity, reliability, and attendant statistical procedures to what is meant by concepts such as ‘true score’ and ‘item response’ theory, ‘operationalization of constructs,’ and ‘response set’ and ‘confounding variables’ such as ‘social desirability,’ as well as countervailing perspectives such as ‘dustbowl empiricism.’ Without sufficient theoretical knowledge, test developers and users may inadvertently engage in assessment research and practice that is conceptually inadequate and inappropriately conducted (e.g., AERA, APA, & NCME, n.d.; Akin, 2012; Geisinger, 2013; IOCE; ITC; Wandschneider et al., 2015) (p. 52 – see the “Assessment” chapter at <https://www.apa.org/pubs/books/going-global>).

That is why we strive to follow [guiding principles and best practices of applied assessment research](#). We also have a strong basis for believing that more information is better than less, which is why we conclude this manual with a primer regarding three interrelated areas of BEVI usage in research and practice: 1) Practitioner-Scholar Assessment; 2) Small N Assessment; and 3) Experimental and Quasi-Experimental Research.

IX. Using the BEVI in Research and Practice

As may be clear by now, assessment occurs at multiple levels of analysis, for different purposes, and by scholars / practitioners with very different educational backgrounds and professional orientations. Overall, for purposes of the BEVI, these levels of analysis may be divided into three types of activity: 1) practitioner-scholar assessment, 2) small N assessment, and 3) experimental and quasi-experimental research. It is important to understand such matters in regards to the design, delivery, analysis, and interpretation of assessment research and practice with the BEVI.

1. Practitioner-Scholar Assessment

Regarding the first level of analysis (practitioner-scholar), it may be helpful to provide an example by considering how another widely used measure, the MMPI ([Minnesota Multiphasic Personality Inventory](#)), is analyzed and interpreted in the real world as there are many parallels to BEVI usage. When mental health practitioners review MMPI results from an individual client, they are mindful of, but not particularly focused upon, big picture data from the standpoint of inferential statistics and mean group differences. Rather, since the MMPI is a standardized measure, just like the BEVI (i.e., with means and standard deviations based upon sample characteristics drawn from the larger population), a clinician examines the meaning of individual scores across the 10 main scales of the MMPI as well as a wide range of subscale analyses. These scores and indexes help the clinician make sense of the individual's psychological functioning across multiple domains (e.g., depression, anxiety, etc.). On the MMPI, scale score interpretation typically involves ["at least 5 T-score points between \[the\] lowest scale in code type and next highest clinical scale in profile."](#) In a mental health context, findings like these may be prominently featured in the context of a psychological evaluation and/or other diagnostic and treatment recommendations for individual clients.

Although it is not a measure of psychopathology per se, the BEVI shares a number of design, usage, and interpretative features of the MMPI, given the BEVI's focus on processes and structures of the human self and identity, which are related to, but different from, our understanding of "personality" (e.g., see chapters 2, 3, and 4 in [Making Sense of Beliefs and Values: Theory, Research, and Practice](#)). In particular, when reviewing a range of BEVI profiles and indexes – for example, Individual Scales, Aggregate Profile, Decile Profile, Profile Contrast – or various report types from the BEVI report system – Individual, Group, Within, Between, Longitudinal – we are interpreting aspects of human functioning and self / identity structure through the lens of what it means to "score" in particular ways on the BEVI.

As such, from a practitioner-scholar perspective, we often say that a “five point difference between time 1 and time 2 administration of the BEVI may be meaningful or worth interpreting.” In offering that perspective, we are *not* making claims about significance testing, an area of statistics where we also have published and presented extensively (to see refereed articles, chapters, and dissertations in this regard, please see <https://thebevi.com/>). Rather, from an interpretive standpoint, we are talking about real world usage of the BEVI, a process that occurs thousands of times with different tests and measures every day all over the world, across multiple settings and a wide array of populations (i.e., at any moment in time, using all manner of instruments, practitioners are engaged internationally in interpretive processes like these).

Consider another example from a Within Group Report on the BEVI for a married couple. If substantial differences emerge between spouses on Negative Life Events, Emotional Attunement, Gender Traditionalism, Self Awareness, and Religious Traditionalism (e.g., 20 point differences or more), we know – as marriage and family therapists – that there may be an empirical basis for the concerns that the couple is having about why they are so often at odds about the nature of reality, which can affect everything: intimacy, child-rearing, finances, communication, political / religious commitments, and so on (e.g., see chapters 11 and 15 in [*Making Sense of Beliefs and Values: Theory, Research, and Practice*](#)).

Likewise, when juxtaposing Individual Reports and reviewing Within Group Report differences via coaching or consulting in an organizational or leadership context – when substantial variability emerges on the BEVI among leaders and their subordinates in how emotions are experienced, reality is described, or expectations are expressed – we are not surprised when major conflicts are reported by team members who feel misunderstood, and in some cases, mistreated (e.g., negated, ignored, dismissed) by their supervisors and vice versa (e.g., supervisors who feel like their authority or judgment is being unduly questioned) (see chapter 14 in [*Making Sense of Beliefs and Values: Theory, Research, and Practice*](#)).

As a final example, in an educational setting, when we review a Group Report of students who are about to do study abroad, participate in Collaborative Online International Learning (COIL), or engage in a course that is designed to be “transformative” or “high impact” (e.g., a multicultural, gender, justice, political, religious, or environmental studies course) – and we observe major BEVI differences within the cohort on scales like Sociocultural Openness, Emotional Attunement, Gender Traditionalism, or Religious Traditionalism – we also are seeing the potential for great conflict as well as great growth. Educators must be able to interpret and use such findings to evaluate and facilitate learning objectives, by designing and delivering evidence-based pedagogies and processes in a sensitive and skillful manner (e.g., see chapters 5, 7, 8, 9, 10, 12, and 13 in [*Making Sense of Beliefs and Values: Theory, Research, and Practice*](#)).

2. Small N Assessment

With the BEVI, there is a second level of analysis, which is informed by methodologies that are often subsumed under “small N,” “single subject” or “single case” research, and are designed to

assess changes in small groups of people (e.g., although sizes vary, N sizes may vary from 1 to 10 or more, depending upon the nature of the research questions). Such small Ns typically are not analyzable via parametric inferential statistics (e.g., analysis of variance, regression) because we cannot necessarily assume that the small N sample data we are analyzing from a larger population are normally distributed.

There are a number of related approaches / types of single subject methods, which also guide our applied assessment research, such as "time series" and "repeated measures" designs. There also are many guidelines that apply to this level of analysis. For accessible resources, see

- 1) <https://academy.pubs.asha.org/2014/12/single-subject-experimental-design-an-overview/>;
- 2) <https://opentextbc.ca/researchmethods/chapter/overview-of-single-subject-research/>; and
- 3) http://courses.phhp.ufl.edu/rcs6740/single_subject.pdf. For a more in-depth understanding of such models and methods, the following textbook by Alan Kazdin may be especially helpful: <https://www.amazon.com/Single-Case-Research-Designs-Clinical-Settings/dp/0190079975>.

Both of these levels of analysis – "practitioner-scholar" and "small N" – are fundamental to matters of BEVI usage and interpretation. To take a few more examples from the BEVI's reporting / AI / VR systems, we rightly pay deep attention regarding 1) within group differences at T1 or when T1 / T2 trends move in the same directions within and between groups on Aggregate Profile and Decile Profile; 2) when the Full Scale score shows the same trends on Profile Contrast between high, medium, and low optimal respondents; 3) when Within Group Report scores correspond with real world events and dynamics by the individuals who receive those scores and/or those individuals who interact with them; 4) when Longitudinal Reports show the same profile patterns within and between institutions across successive years; and 5) when demographic analyses reveal particularly striking and consistent trends by gender, ethnicity, educational background, and so forth.

As such, it is not possible to understand how to "interpret" the BEVI if one is only looking at statistically significant and mean-based differences at a group level, fascinating as that level of analysis is, and as we've seen through much of our own empirical research of this nature over the years. In fact, we often apply a range of statistical analyses to BEVI data, when we have sufficient Ns to do so, and can be confident that the parameters for inferential statistics apply (e.g., data are drawn from what we have reason to believe is a normal distribution of such data). We also have analyzed multiple phenomena using relevant statistical procedures (e.g., EFA / CFA / IRT) to help us understand the factor structure of the BEVI or through model / criterion-based studies that help us examine what various indices of the BEVI predict (e.g., using SEM, regression, ANOVA, etc.) (e.g., see chapters 4, 5, 6, 7, 8, 9, 10 in [*Making Sense of Beliefs and Values: Theory, Research, and Practice*](#)).

So again, and to be clear, when we say a "five point difference is potentially meaningful," we are not talking about significance testing across large groups, but rather, practices and processes that are fundamental to what it means to interpret profiles across relatively small groups of individuals from a "practitioner-scholar" or "small N" perspective. In other words, we are

looking at such data like any evaluator or interventionist would (e.g., clinician, educator, facilitator, consultant, interviewer, coach) – or how a researcher does, who is engaged in "visual analysis" of data from a single-subject / single-case / time series / repeated measures design. That is why we never say that a "five point difference" is statistically significant unless we've done significance testing, just as we always say we need to be careful about confusing correlation and causation. But those caveats do not refute what legitimately happens every day in practice and research with assessment instruments and data all over the world.

In short, from a "practitioner-scholar" and "small N" vantage point – given the structure, design, and usage characteristics of the BEVI, and all we have learned over the years through research and practice – it is not unreasonable to say that a "five point difference" is potentially meaningful, particularly if it is juxtaposed with additional analyses. That is the whole point and purpose of integrative, comprehensive, and depth-based interpretation of assessment data.

3. Experimental and Quasi-Experimental Research

As a final consideration regarding BEVI usage in research and practice, it also is important to understand the [difference between experimental and quasi-experimental research design as well as attendant implications and applications](#). Much of applied assessment research falls into the latter camp (e.g., quasi-experimental), as it relies upon psychometrically sound measures to ascertain the degree of change that may be occurring across a wide range of populations, settings, and contexts (e.g., [here is an example of this type of research with the BEVI](#); to see how such research and practice occurs internationally, please see Chapter 2, Assessment, in [Going Global: How Psychologists Can Meet a World of Need](#)).

Although [we have learned a great deal from quasi-experimental research with the BEVI \(e.g., see the research chapters in Making Sense of Beliefs and Values\)](#), it should be noted that one of the problems of between-group analyses – which focus on significance testing to determine if "change" occurs from T1 to T2 (pre / post) – is that within group variability is underemphasized from a measurement standpoint. In other words, we may well see aggregated changes from Time 1 to Time 2, but is that change always or typically uniform for all subgroups within the larger group? No, it is not. As was demonstrated through the six-year [Forum BEVI Project](#), and as we see through a wide range of indices on the BEVI (e.g., Decile Profile), we have to examine the interacting mediators and moderators of subgroup change in order to understand what is really happening within the group (e.g., within-group variability is often larger – and more interpretively meaningful – than between-group variability). From the standpoint of the BEVI method and EI model, such conceptual and methodological considerations are informed by [developmental psychopathology](#) as well as complementary fields of inquiry and practice.

Given the nature of these constructs – and the corresponding need to operationalize relevant scales on the BEVI – a great deal of complexity among variables must be accounted for through ecologically valid measurement. In this regard, true experimental research – randomized assignment to different intervention groups and the deliberate inclusion of a control group – is a

commendable objective that we wholly endorse, but from a real world usage perspective (e.g., using convenience samples), there are many complexities of engaging in such work across different countries and contexts. Relatedly, although an essential methodology, there also have been many [criticisms of randomized controlled trials or RCTs](#). For example, they tend to underemphasize ecological validity in favor of empirical precision, which doesn't always map well to the real world where numerous confounds are operative (e.g., intervention error that is inadvertently introduced).

On the other hand, given the demographic section of the BEVI, we have been able to conduct a number of statistical analyses by comparing and contrasting different groups within very large datasets (e.g., we can create comparison groups by juxtaposing matched groups within the BEVI and then conducting a wide array of statistical analyses). In other words, [we've been able to ascertain the differential impact of various predictors on a wide range of outcome variables](#), including scale scores and other criteria (e.g., grade point averages, course satisfaction; etc.).

In the final analysis, the BEVI has in fact been used in a wide range of quasi-experimental research activities over the years. Of particular note, the BEVI was utilized with around 20 institutions in the U.S. and internationally through the [Forum BEVI Project](#), a multi-year assessment-to-practice initiative which resulted in “over 20 publications (e.g., articles, chapters, dissertations), 50 presentations (e.g., symposia, papers, posters), and hundreds of separate analyses” (Wandschneider et al., 2015, p. 418). Using related analytic methodologies, the BEVI also is integral to [Collaborative Online International Learning or COIL and the COIL BEVI Project](#), a multi-institution, multi-country initiative that is evaluating and facilitating the effectiveness of online learning, including Virtual Exchange (VE) and Collaborative Online International Learning (COIL). [Click here](#) for a brief video about the COIL BEVI Project. For examples of other such research, please see <https://thebevi.com/about/projects/>. Likewise, if you would like to collaborate with us on any of these forms of research and practice – from practitioner-scholar, to small N, to experimental / quasi-experimental – please [contact us](#).

X. Further Information About the BEVI

We have covered a lot of content in this manual, but you may still have further questions or need additional information. Good resources in that regard include the BEVI website (<https://thebevi.com/>) as well as [Making Sense of Beliefs and Values: Theory, Research, and Practice](#). Individuals who are trained on the BEVI also may have access to the “Admin” system on the BEVI website (<https://thebevi.com/login/>), which includes a wide range of resources (e.g., peer reviewed articles on the BEVI; information about accessing the AI / VR “Beviverse”; training videos; administrator / IRB materials, etc.). Finally, please feel free to contact us directly at info@thebevi.com.

With hopes that the BEVI Manual has been helpful, we look forward to working with you. Thank you for your interest in the BEVI.