

DR. SCOTT KIMBALL

WHITE PAPER · 2026

The Summit Map

A science-backed model of how people grow: think, act and be in alignment, and climb from structure to stability to direction.

Dr. Scott Kimball, PhD
drscottkimball.com

Science-backed. **Life-tested.**

Contents

1. 1. Executive summary
2. 2. The problem: wanting success, and stalling anyway
3. 3. Historical and theoretical foundations
4. 4. The Summit Map model
5. 5. Measurement design
6. 6. The validity program and honest claims
7. 7. Across cultures and walks of life
8. 8. How people use it
9. 9. Ethics and limitations
10. 10. Conclusion: Now Go and Do
11. References

1. Executive summary

Most people who want a better life already know a great deal about what they should do. They have read the books, taken the personality tests and set the goals. What they often lack is a clear answer to a narrower question: *given who I am, where is my life thinly built right now, and what is the one thing to work on next?*

The Summit Map is a 70-item, roughly ten-minute self-report assessment built to answer that question. It rests on three ideas from established psychology.

- **Who you are and how your life is built are different things.** Personality science separates relatively stable traits from changing states (Fleeson, 2001; Steyer, Schmitt & Eid, 1999). The Summit Map measures both, separately. Your **gifts** (18 of them) describe how you are wired. Your **foundation** (nine squares) describes how well your life is built over the last 30 days.
- **A good life needs thought, action and identity to line up.** The oldest division of the mind in Western psychology splits it into cognition, conation and affect (Hilgard, 1980). We call these **Think · Act · Be**. We do not rank them. Research on cognitive dissonance, self-discrepancy and congruence (Festinger, 1957; Higgins, 1987; Rogers, 1959) suggests that strain arises when they tell different stories. The Summit Map therefore reports an **Alignment** score at each level of the climb.
- **Lives are built in layers.** Self-regulation research clusters into three well-replicated families: the scaffolding of habits and plans, regulation under pressure, and goals and meaning. We call these **Structure · Stability · Direction** and arrange them as levels of a climb. **Base Camp** is the whole Structure row. **The Summit** is the whole Direction row: a purpose you can see (Vision), pursue (Aimed Effort) and live (Purpose & Values).

From those scores the report derives four things a person can use on Monday: a **Summit Code** (their top two gifts, for example "Visionary · Spark"), a **Shadow** (how their leading gift tends to misfire when it runs on their thinnest level), an **Alignment** reading (where their thoughts, actions and identity diverge), and a **Next Camp** (the one square to build first, with a Think · Act · Be plan sized to their week).

The gifts are measured with graded forced-choice pairs, which make it harder to rate oneself highly on everything. The foundation is measured with concrete, time-bound Likert items that include reverse-keyed checks. The current scoring is transparent and deliberately simple.

Thurstonian item response theory (Brown & Maydeu-Olivares, 2011, 2018) is planned once enough data exist.

We want to be plain about what is and is not established. **The model is built on published research, but the instrument itself has not yet been validated.** Its reliability, factor structure, relationship to established measures, predictive power, equivalence across languages and resistance to the Barnum effect are all the subject of a pre-specified validity program described in Section 6. We will make public claims only as the evidence supports them.

The Summit Map is a tool for personal growth. It is not a diagnosis, and it must never be used for hiring or selection.

2. The problem: wanting success, and stalling anyway

2.1 The intention–behavior gap

Psychology has known for decades that intending to do something and doing it are only loosely connected. In a meta-analysis of experiments that successfully changed people's intentions, Webb and Sheeran (2006) found that a medium-to-large change in intention produced only a small-to-medium change in behavior. Reviews of the "intention–behavior gap" reach the same conclusion: a large share of people who form a sincere intention fail to act on it (Sheeran, 2002; Sheeran & Webb, 2016).

The gap is not mainly a failure of desire. It tends to come from predictable places: no plan for *when* and *where* to act, no routine to carry the behavior when motivation fades, poor regulation when stress arrives, and goals that were never really one's own. Each of these is a well-studied problem with well-studied remedies. The difficulty for an ordinary person is knowing which one is theirs.

2.2 Insight without a next step

The self-assessment market is very large and growing. Popular instruments sell an experience of recognition: *that's me*. Recognition has value. It gives people language for themselves and for others. But it rarely tells a person what to do differently, and a label that describes who you are, permanently, can quietly suggest that nothing needs to change.

There is also a risk in the opposite direction. Feedback does not automatically help. In a large meta-analysis, Kluger and DeNisi (1996) found that feedback improved performance on average, but in over a third of cases it *reduced* performance, especially when it drew attention to the self rather than to the task. A report that says "you are a certain kind of person" without saying "here is the next task" sits on the wrong side of that finding.

The research on goal monitoring is more encouraging. Monitoring progress toward a goal reliably improves the chance of reaching it, and the effect is larger when progress is recorded and revisited (Harkin et al., 2016). This is why the Summit Map is built around a concrete next step and a 90-day retake rather than a fixed label.

2.3 Strengths are not the whole story

A generation of strengths-based tools taught people to stop fixing weaknesses and invest in talents (Buckingham & Clifton, 2001). That advice has merit, and we keep its spirit. But it can miss something. Many people who stall are not short of talent. Their gifts are running on a foundation

that cannot carry them: no routine, no footing under pressure, no clear aim. The visionary who keeps starting over and the hard worker who stays busy without getting anywhere are not lacking strengths. They are lacking *ground*.

The Summit Map is designed around that observation. It asks two questions in parallel: *what are your natural gifts?* and *how well built is the ground they stand on, right now?* Then it asks a third: *do your thoughts, actions and identity line up?*

The one question the Summit Map answers: "What do I need to think, do and be to bring more structure, stability and direction into my life?" Insight without a next move is entertainment. A next move without insight is generic advice. The value is the two together.

3. Historical and theoretical foundations

The Summit Map does not introduce a new psychology. It arranges well-established findings into a structure that ordinary people can use. This section sets out those foundations, including where the evidence is mixed.

3.1 The trilogy of mind: from Plato to Hilgard, and Think · Act · Be

The idea that the mind has three parts is old. In the *Republic*, Plato divided the soul into reason, spirit and appetite. His division is not the same as the modern one, and we do not claim it is. But it began a long tradition of treating mental life as several distinct capacities that must be brought into harmony. Plato's own picture of a well-ordered person is one in which the parts work together rather than against each other, an early version of what we call alignment.

The modern form of the idea comes from eighteenth-century German faculty psychology, which Kant made central to his philosophy. In the *Critique of Judgment* (1790), Kant organized the mind's powers into three: the faculty of cognition, the feeling of pleasure and displeasure, and the faculty of desire. In 1980, Ernest Hilgard traced how this "trilogy of mind," under the names **cognition, affection and conation**, persisted through the history of psychology, and argued that the field had neglected conation, the striving, willing side of mental life (Hilgard, 1980).

The Summit Map's columns follow this trilogy:

Summit Map	Classical term	What it covers in the model
Think	Cognition	How you see: priorities, appraisal under pressure, vision
Act	Conation	How you move: routines, follow-through, where effort actually goes
Be	Affect, extended to identity and relationships	Who you are with others: boundaries and bonds, emotional steadiness, purpose and values

The **Be** column is our extension. The classical trilogy places feeling in the third position. We broaden it to include identity, values and relationships, because the research on emotional steadiness (Campbell et al., 1996), meaning (Steger et al., 2006) and social connection (Holt-Lunstad, Smith & Layton, 2010) consistently ties feeling to *who one takes oneself to be* and *with whom one lives*. Skeptical readers should treat this broadening as a design choice to be tested, not a historical given.

Two features of how we use the trilogy matter for everything that follows. First, **Think, Act and Be are not ranked**. Thinking is not "lower" than being, and being is not the goal that thinking serves. Second, the trilogy implies a standard: the parts of the mind should work *together*. That is the root of the Alignment score.

"Think anew, and act anew." The spirit of Think · Act · Be has a well-known American ancestor. In his Annual Message to Congress on December 1, 1862, in the middle of the Civil War, Abraham Lincoln wrote: *"The dogmas of the quiet past, are inadequate to the stormy present. The occasion is piled high with difficulty, and we must rise with the occasion. As our case is new, so we must think anew, and act anew. We must disenthrall ourselves, and then we shall save our country"* (Lincoln, 1862/1953). Lincoln's sequence anticipates ours: new thinking, then new action, then a change in who we are ("disenthrall ourselves"). Dr. Kimball credits this passage as an inspiration for the Think · Act · Be framework. The Summit Map adds the third term explicitly, **be**, and asks whether all three move together.

3.2 The strengths tradition and its critiques

Donald Clifton spent decades asking what would happen if psychology studied what people do well. His work, popularized through the Clifton StrengthsFinder (Buckingham & Clifton, 2001), proposed that talents are naturally recurring patterns of thought, feeling and behavior, and that talent becomes strength only through *investment*: practice, skill and knowledge (Clifton & Harter, 2003). The Summit Map borrows this distinction directly. Gifts are the talents. The foundation is a measure of the conditions those talents are invested in.

In 2000, Martin Seligman and Mihaly Csikszentmihalyi called for a **positive psychology** that would study strengths, well-being and the conditions that let people flourish, alongside the traditional focus on disorder (Seligman & Csikszentmihalyi, 2000). The movement produced a classification of character strengths (Peterson & Seligman, 2004) and a large literature on interventions.

The strengths tradition has drawn serious criticism, and we take it seriously:

- **Overstated effects.** Early meta-analyses of positive psychology interventions found modest benefits (Bolier et al., 2013). A later re-analysis that corrected for small-sample bias found the effects were considerably smaller than first reported (White, Utzl & Holder, 2019).
- **Thin theory and a one-sided view of life.** Lazarus (2003) argued that the movement separated positive and negative experience too sharply and leaned on weak methods. Wong (2011) called for a "second wave" that treats suffering and meaning together.
- **Strengths can be overused.** A strength pushed too far becomes a liability. Grant and Schwartz (2011) documented inverted-U relationships across many positive traits, and Kaiser and Overfield (2011) showed that leaders often overuse their strengths. Biswas-Diener, Kashdan and Minhas (2011) argued that strengths should be understood in context, not as fixed assets.
- **Publisher-generated evidence.** Much of the evidence for commercial strengths instruments comes from the publishers themselves.

The Summit Map responds to these critiques in its design. Every gift has a **shadow**, and the model predicts when the shadow will appear. Gifts are never reported without the foundation they run on. And our evidence program (Section 6) is pre-specified and will be published whatever it shows.

3.3 Self-regulation: habits, plans and self-control

The **Structure** and **Stability** levels draw on the self-regulation literature, one of the most productive areas of modern psychology.

Habits. A large share of daily behavior is repeated in stable contexts and runs with little deliberation. Wood and Neal (2007) described habits as context–response associations that, once formed, are triggered by cues rather than goals (see also Wood & Rünger, 2016). The practical lesson is that routines carry behavior when motivation does not. In a real-world study of habit formation, Lally, van Jaarsveld, Potts and Wardle (2010) found that the time for a new behavior to become automatic varied widely, from 18 to 254 days, with a median of 66. The point for our purposes is not the number but the variability: building a routine takes longer than people expect, and missing a single day did not derail the process.

Implementation intentions. Gollwitzer (1999) showed that simple "if–then" plans ("When situation X arises, I will do Y") substantially increase the chance that intentions turn into action. A meta-analysis found a medium-to-large effect across many domains (Gollwitzer & Sheeran, 2006). The Summit Map uses if–then plans throughout its Next Camp actions.

Planning and time management. A meta-analysis of time-management research found that it is associated with better performance and, more strongly, with well-being and lower distress (Aeon, Faber & Panaccio, 2021).

Self-control. Trait self-control predicts better outcomes across school, work, health and relationships (Tangney, Baumeister & Boone, 2004). We measure the behavior, following through, rather than the trait. Two cautions apply. First, the popular "ego depletion" account of self-control as a limited resource has not held up well: a large preregistered multi-site replication found an effect close to zero (Hagger et al., 2016). We therefore do not build on the "willpower is a muscle that tires" metaphor. Second, the related concept of **grit** (Duckworth, Peterson, Matthews & Kelly, 2007) has been widely marketed, but a meta-analysis found that grit overlaps very heavily with conscientiousness and adds little predictive power beyond it, and that its perseverance facet carries most of whatever predictive value it has (Credé, Tynan & Harms, 2017). We avoid the word, measure the concrete behavior, and plan to test whether our squares add anything beyond the Big Five (Section 6).

3.4 Emotion regulation and resilience

The **Stability** level asks whether a person keeps their footing when the ground moves.

Gross (1998) organized the field of emotion regulation around a process model: people can influence emotions by selecting and changing situations, by shifting attention, by changing how they appraise what is happening, and by modulating their response. Cognitive reappraisal, changing how one reads a situation, is generally associated with healthier outcomes than suppression in Western samples (Gross & John, 2003), though the costs of suppression appear to

vary by culture (see Section 7). Stress also tends to shift decision-making toward faster, more habitual and short-term choices (Starcke & Brand, 2012). These findings underpin the **Clear Head** square.

Bonanno (2004) challenged the assumption that most people are badly damaged by loss or trauma. Across studies, the most common response to potentially traumatic events was a stable trajectory of healthy functioning, which he called **resilience**, and it was achieved by a wider range of people and routes than expected. This finding matters for how we read scores: hardship does not imply a low foundation, and a steady score in a hard season can be genuine.

Self-concept clarity, the extent to which a person's beliefs about themselves are clear, confident and consistent, is associated with higher self-esteem and lower neuroticism (Campbell et al., 1996). Together with emotional stability, it underpins the **Steady Self** square.

3.5 Goals, hope and meaning

The **Direction** level, the Summit, draws on research about where effort is aimed and why it matters.

- **Goal-setting theory.** Specific, challenging goals, with feedback and commitment, lead to higher performance than vague "do your best" goals (Locke & Latham, 2002).
- **Hope theory.** Snyder (2002) defined hope as the perceived capacity to find routes to desired goals (*pathways*) and to motivate oneself to use them (*agency*). A clear picture of the future, and more than one way to reach it, is the core of the **Vision** square. Mental contrasting, pairing a desired future with the obstacles in the way, makes action more likely than positive fantasy alone (Oettingen, 2012). Related work on *possible selves* shows how imagined future selves guide present behavior (Markus & Nurius, 1986).
- **Self-concordance.** Goals that express a person's own interests and values attract more sustained effort, are more likely to be attained, and produce larger gains in well-being when attained than goals pursued from pressure or obligation (Sheldon & Elliot, 1999). This underpins the **Aimed Effort** square: effort that actually goes where one's own goals are.
- **Meaning and values.** The Meaning in Life Questionnaire distinguishes the *presence* of meaning from the *search* for it (Steger, Frazier, Oishi & Kaler, 2006). Schwartz (1992) found a common structure of basic human values across many countries, with people differing in how they prioritize them. Damon, Menon and Bronk (2003) described purpose as a stable intention to accomplish something that is meaningful to the self and consequential beyond it. Purpose in life also predicts longevity across adulthood (Hill & Turiano, 2014). These works ground the **Purpose & Values** square.

3.6 Congruence: why alignment matters

The Summit Map's most distinctive claim is that a person's thoughts, actions and identity should *tell the same story*, and that when they do not, people feel it. This idea has deep roots.

- **Cognitive dissonance.** Festinger (1957) proposed that holding inconsistent cognitions, including a belief and a contrary behavior, produces discomfort that people are motivated to reduce, often by changing their attitudes rather than their actions.
- **Self-discrepancy theory.** Higgins (1987) proposed that gaps between the *actual* self and the *ideal* self are associated with dejection-related emotions, and gaps between the actual and the *ought* self with agitation-related emotions. Discrepancy itself, not any one level, is the source of strain.
- **Rogers' congruence.** Rogers (1959) placed congruence between self-concept and experience at the center of psychological health. Incongruence, in his account, produces vulnerability and anxiety.
- **Self-concordance,** discussed above, makes a related point about goals: effort that does not express the self is harder to sustain (Sheldon & Elliot, 1999).
- **Identity-based behavior change.** Self-perception theory proposed that people infer their attitudes partly from watching their own behavior (Bem, 1972). Identity-based motivation theory proposes that people act in ways that feel congruent with salient identities (Oyserman, 2009). In one well-known set of studies, framing voting as an identity ("be a voter") rather than an action ("vote") increased turnout (Bryan, Walton, Rogers & Dweck, 2011), though a later large field experiment did not find the same effect (Gerber, Huber, Biggers & Hendry, 2016). The popular advice to build "identity-based habits" (Clear, 2018) draws on this literature. The evidence is promising but mixed, and we treat it that way.

The Summit Map draws a modest, testable inference from this literature: **at each level of the climb, large gaps between Think, Act and Be scores should be associated with strain and with stalled progress**, and closing a gap should be associated with gains. This is a hypothesis, and testing it is part of our predictive study.

3.7 Trait and state, person and situation

Psychology long argued over whether behavior is better explained by stable traits or by situations. The modern consensus is that both matter. Lewin (1936) summarized it in a formula: behavior is a function of the person and the environment, $B = f(P, E)$. Mischel and Shoda (1995) showed that people have stable *if-then* signatures: consistent patterns of responding differently to different situations.

Fleeson (2001) showed, using repeated daily measurement, that individuals vary as much *within* themselves from moment to moment as people vary from one another, and that traits are best understood as the *average* of many states. Latent state-trait theory provides formal models for separating the stable component of a measurement from the occasion-specific component (Steyer, Schmitt & Eid, 1999). McAdams (1995) proposed that personality has several levels: broad traits, characteristic adaptations (goals, strategies, values) and life narratives.

The Summit Map's two layers follow this logic:

- **Gifts** are measured as *trait-like* preferences, without a time frame, and should be relatively stable over months.
- **Foundation** is measured as a *state*, over the last 30 days, and should change as a person's circumstances and habits change.

This is why we expect, and will test, that gift rankings hold steady across a 90-day retake while foundation scores move. It is also why we never say a person "is" a low foundation. A foundation is built, and it can be rebuilt.

3.8 Frankl and purpose in hardship

Viktor Frankl, writing from his experience of the Nazi concentration camps, argued that the search for meaning is a primary human motivation and that meaning can be found even in unavoidable suffering, in the attitude one takes toward it (Frankl, 1946/2006). His work is testimony rather than experiment, but later research supports several of its claims. Meaning is associated with well-being even when happiness is low, and meaningfulness is linked to stress and struggle in ways that happiness is not (Baumeister, Vohs, Aaker & Garbinsky, 2013). Making meaning of stressful events is associated with adjustment (Park, 2010). In a large multinational survey, people in poorer nations reported greater meaning in life than people in wealthier nations, even as they reported lower life satisfaction (Oishi & Diener, 2014).

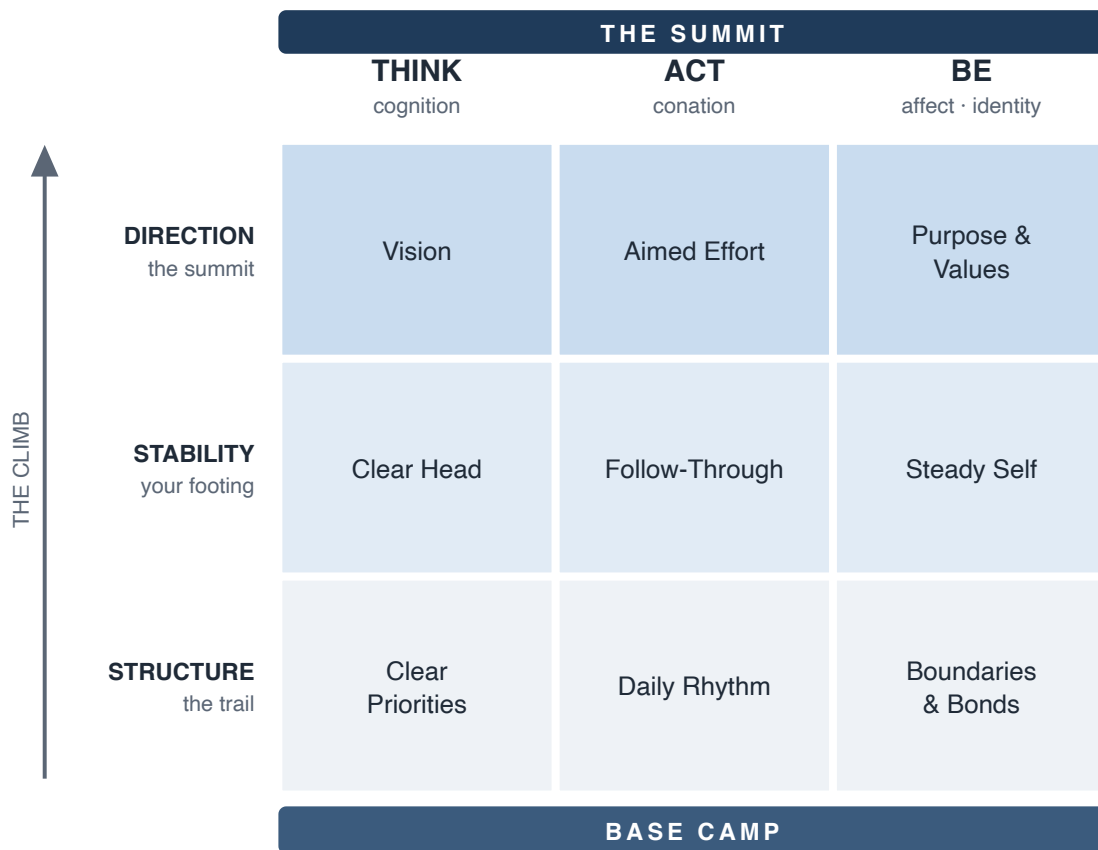
For the Summit Map, Frankl's lesson is direct: **purpose is not the same as success**. A person in a hard season can genuinely stand on the Summit level. Section 7 explains how the model distinguishes that case from a rosy self-report.

4. The Summit Map model

4.1 The grid

The Summit Map arranges nine squares in a three-by-three grid.

- **Across** the grid run **Think · Act · Be**. These are not ranked. Reading across a row asks whether your thoughts, actions and identity line up at that level. **Alignment is the goal.**
- **Up** the grid run **Structure → Stability → Direction**. These are levels of the climb. Structure is the trail, Stability is your footing, and Direction is the Summit.



Read across for alignment. Read up for the climb.

Figure 1. The Summit Map. Think · Act · Be run across and are not ranked; alignment across a row is the goal. Structure, Stability and Direction run up the side as the levels of the climb. Base Camp is the whole bottom row; The Summit is the whole top row.

Two clarifications prevent common misreadings. First, the grid is a map of nine separately measured squares, not a pair of continuous axes. A person does not have a single "position" on it. Second, the upward ordering of Structure, Stability and Direction is a **developmental hypothesis**, not a scoring rule. We present it because it is useful and plausible, and we will test it (Section 6).

4.2 The nine squares

Each square has a plain-language question, a definition, its research basis, and a Next Camp plan written in Think · Act · Be form.

Base Camp: the Structure level (the trail)

Square	Life question	Definition	Research basis
Clear Priorities (Think)	Do I know my top three things this week?	You know what matters most right now and have a simple plan for it.	Time-management meta-analysis (Aeon et al., 2021); goal specificity (Locke & Latham, 2002)
Daily Rhythm (Act)	What happens on a day I don't feel like it?	Routines carry you when motivation doesn't show up.	Habit automaticity (Wood & Neal, 2007; Lally et al., 2010); implementation intentions (Gollwitzer, 1999)
Boundaries & Bonds (Be)	Who gets my best, and who gets whatever is left?	You protect your time and energy and invest on purpose in the relationships that matter most.	Boundary management (Ashforth, Kreiner & Fugate, 2000); social relationships and mortality (Holt-Lunstad et al., 2010)

Structure is the trail: the scaffolding that turns intention into regular action. It is the least glamorous level and, we hypothesize, the one most often skipped.

Footings: the Stability level

Square	Life question	Definition	Research basis
Clear Head (Think)	When pressure hits, do I get sharper or foggy?	You think clearly under pressure and look for options instead of spiraling into worst cases.	Cognitive reappraisal (Gross, 1998); decision-making under stress (Starcke & Brand, 2012)
Follow-Through (Act)	How many unfinished things am I carrying?	You keep going when the work gets hard, boring or behind schedule, and you finish what you start.	Self-control (Tangney et al., 2004); persistence, with the grit critique noted (Credé et al., 2017)
Steady Self (Be)	Do my circumstances decide who I am today?	You stay emotionally steady and keep a clear sense of who you are when life shakes.	Resilience (Bonanno, 2004); self-concept clarity (Campbell et al., 1996); emotional stability as the inverse of neuroticism

Stability is footing: the capacity to keep going, and keep one's head and sense of self, when conditions turn.

The Summit: the Direction level

Square	Life question	Definition	Research basis
Vision (Think)	Can I describe my life three years from now?	You can picture a future worth climbing toward and see more than one way to get there.	Hope theory (Snyder, 2002); mental contrasting (Oettingen, 2012); possible selves (Markus & Nurius, 1986)
Aimed Effort (Act)	If someone watched my calendar, could they guess my goals?	Your time and effort actually go toward the goals and people that matter most to you.	Goal-setting theory (Locke & Latham, 2002); self-concordance (Sheldon & Elliot, 1999)
Purpose & Values (Be)	Would I be proud of how I spent this year?	You know what you stand for and why it matters, and your life reflects it.	Meaning in life (Steger et al., 2006); basic values (Schwartz, 1992); purpose (Damon et al., 2003)

4.3 The Summit and Base Camp

An earlier version of the model treated a single square, Purpose & Values, as the summit. We revised that. **The Summit is the whole top row:** a purpose you can *see* (Vision), *pursue* (Aimed Effort) and *live* (Purpose & Values). The change follows from the alignment principle. A purpose that is felt but not pursued, or pursued but not owned, is not yet a purpose-driven life. The report's Summit score is the mean of the three Direction squares.

Base Camp is the whole bottom row: priorities you can name, routines that hold, and boundaries and bonds that protect your energy and your people. Every climb starts there.

The model's purpose is simple to state. It aims to help people build a life in which a purpose they have chosen is visible in their plans, their days and their sense of who they are, on footing that holds when conditions get hard.

4.4 The 18 gifts

The second layer of the model is the **Signature:** 18 gifts, two homed in each square. A gift is a natural tendency that makes building that square easier. It is *not* the same as the square's current condition. A person can have strong Structure gifts and a thinly built Structure level, which is often the most useful insight in a report.

Level	Think	Act	Be
Direction	Visionary (sees the mountain before anyone else does) · Inventor (turns problems into new ideas that work)	Spark (turns ideas into motion today) · Trailblazer (goes first where others won't)	Champion (fights for the people and causes that matter) · Coach (sees who people could become and helps them grow)
Stability	Truth-Seeker (digs until they find what's actually true) · Realist (sees things as they are, without panic or hype)	Finisher (gets things over the finish line) · Fixer (at their best when something breaks)	Cornerstone (the steady one others lean on) · Bridge Builder (turns conflict into understanding)
Structure	Designer (sees the system that makes the mess make sense) · Navigator (always knows which step comes next)	Builder (turns plans into things people can touch) · Grinder (shows up every day, whether they feel like it or not)	Guardian (protects the people and promises in their care) · Rope Team (connects people so no one climbs alone)

Several design rules shaped the gifts:

- **Clear before catchy.** Every name must be instantly understandable to an ordinary reader with no knowledge of mountaineering or psychology. Names were also chosen to avoid the theme names of widely used commercial instruments.
- **Gifts are a communication layer on dimensional scores.** Every person has a score on all 18 gifts. The names describe the top of a continuous profile. They are not types, and no one is placed in a box.
- **Each gift has a light side and a shadow.** The report describes the gift at its best, what fuels it, what drains it, and how it misfires.

4.5 The Shadow principle

A gift running on a weak foundation tends to turn into its shadow.

This is the model's signature prediction. It follows from the research on strength overuse (Grant & Schwartz, 2011; Kaiser & Overfield, 2011) and from the person × situation principle: the same disposition expresses itself differently depending on the conditions it operates in (Lewin, 1936; Mischel & Shoda, 1995).

In practice, the report pairs a person's #1 gift with their thinnest level, the lowest of the Structure, Stability and Direction level scores, and selects one of 54 written shadow descriptions (18 gifts × 3 levels). Two examples, paraphrased:

- A **Visionary** on thin Structure sees the mountain clearly and sprints at it, but without a trail keeps starting over at Base Camp. The next move is not more vision. It is one structure.
- A **Grinder** on thin Direction shows up every day and works hard, but without an aim, motion can be mistaken for progress.

The Shadow principle is stated as a prediction because it can be tested. If shadows are real, people whose #1 gift sits on a thin level should endorse the matching shadow description more than other shadow descriptions, and more than people whose gift sits on solid ground. We plan to test this as part of the blind-report study (Section 6.7).

4.6 The Alignment score

At each level, the report asks whether Think, Act and Be tell the same story. The calculation is deliberately transparent:

- At each level, take the three square scores (Think, Act, Be) and find the **spread** between the highest and the lowest.
- The level's **Alignment** is 100 minus that spread. A level where all three squares score the same has an Alignment of 100.
- A spread of 20 points or more is reported as a **gap** and named by direction, for example "Think ahead of Act."
- **Overall Alignment** is the mean of the three level scores. The level with the largest spread is highlighted.

There are six possible gap directions, and each has a plain description and a remedy:

Gap	What it sounds like	First remedy
Think ahead of Act	"You know it. You aren't doing it yet."	One small, scheduled action, not more planning
Think ahead of Be	"You see it, but it isn't you yet."	Say it as an identity and back it with one action
Act ahead of Think	"Busy hands, unclear plan."	Fifteen minutes to decide what the doing is for
Act ahead of Be	"Doing the work, but it isn't you."	Connect one daily task to why it matters
Be ahead of Think	"You feel it, but haven't mapped it."	Write the next three steps that person would take
Be ahead of Act	"Your values are clear. Your calendar doesn't show them yet."	Put one value on the calendar as a real appointment

Alignment is reported alongside altitude, never instead of it. A level can be aligned but low, which the report treats as good news: because the three move together, one deliberate step can lift all of them.

We acknowledge a technical limitation. A difference between two noisy scores is noisier than either score alone, and a range across three short scales inherits their measurement error.

Alignment should be read as a prompt for reflection, not a precise measurement, until its reliability

has been estimated (Section 6.2).

4.7 Next Camp

The report recommends one square to build first, the **Next Camp**. The rule is simple:

1. Find the lowest-scoring square.
2. If several squares are within 5 points of the lowest, choose by build order: Structure before Stability before Direction, then Think before Act before Be.

The build-order tie-break expresses our working hypothesis that footing is easier to establish on a trail, and that aim is easier to sustain on footing. It is only a tie-break. A clearly lower square is always chosen first, wherever it sits.

The Next Camp section then does three things that make the report personal rather than generic. It quotes back the person's own answers that pulled the square down most. It selects a **leverage gift**: the highest of the person's top three gifts that lives in the Next Camp square or shares its level or column, so the plan is built *their* way. And it sizes a Think · Act · Be plan to the time the person says they have each week, applied to the area of life they want to focus on, with a workaround for the obstacle they named. A seven-day starter plan and a 30-day measure of success follow.

4.8 The Summit Code

Every person receives a **Summit Code**: their top two gifts, for example "Visionary · Spark" or "Builder · Coach." The code is short enough to remember and say aloud. It gives people a shared vocabulary for how they and others are wired, and a natural way to talk about the model. The full report shows the top three gifts and the order of all 18. Because gift order is within-person (Section 5.4), the Summit Code describes what a person leans on most. It does not say they have more of that gift than someone else does.

4.9 Reality checks built into the model

The grid is designed to check itself. Several profile patterns are named in the report:

- **Summit without a trail**: a high Direction level over a thin Structure level.
- **Trail without a summit**: a well-built Structure level under a thin Direction level.
- **Footing without aim**: strong Stability with thin Direction.
- **Vision without motion**: a high Vision score with a low Aimed Effort score.
- **Even, all built, all bare**: profiles that are flat, uniformly high or uniformly low.

When the Direction level is high while the Structure level is thin, the report does not simply congratulate the person on their purpose. It says that their sense of purpose is strong, that the trail under it is still thin, and that purpose becomes a life when ordinary days carry it. Section 7.6 explains why this matters.

5. Measurement design

5.1 Overview

The Summit Map has 70 items and takes about ten minutes.

Section	Items	Format	Measures
About you	7	Single choice	Tailoring only, never scored: age, life stage, focus area, weekly time, main obstacle, accountability style, past success lever
How you're wired	36	Graded forced-choice pairs, 5 points	The 18 gifts (trait-like)
Your last 30 days	27	5-point Likert, labeled "Not at all" to "Completely"	The 9 squares (state)

The person never sees construct names, items from different constructs are interleaved, and the scoring key is held on the server. The item wording and the item-to-construct key are confidential to protect the instrument's integrity. Qualified reviewers can access them under agreement.

5.2 Gifts: graded forced-choice pairs

Each gift item presents two statements, each describing a different gift, and asks which sounds more like the respondent on a five-point scale: much more like the left, somewhat more like the left, equally, somewhat more like the right, much more like the right.

Illustrative example only. This pair was written for this paper. It is not an item from the Summit Map.

Which sounds more like you?

Left: "I would rather map out the steps before anyone starts."

Right: "I would rather get people moving and work out the steps as we go."

The design follows several rules:

- Each of the 18 gifts appears in four pairs, twice on the left and twice on the right, so position effects balance out.
- The two statements in a pair always come from **different squares**, so a pair never asks a person to choose between two expressions of the same area.
- Pairings were generated from a balanced graph design so that every gift is compared against a spread of others.
- In translation, the two statements are kept equally attractive (Section 7.4).

Why forced choice? When people rate statements one at a time, they can rate themselves highly on everything that sounds good. This is the classic problem of socially desirable responding (Edwards, 1957; Paulhus, 1984). Forced choice requires a trade-off: favoring one gift necessarily means favoring the other less. Meta-analytic evidence suggests that forced-choice formats can reduce faking in high-stakes settings, though they do not eliminate it, and results depend on how well the paired statements are matched for desirability (Cao & Drasgow, 2019). We claim only that the format makes uniform self-flattery harder.

Why graded rather than binary? A five-point graded preference captures *how much* a person prefers one statement, not just which, which yields more information per item. Brown and Maydeu-Olivares (2018) extended Thurstonian modeling to graded-preference data of this kind.

5.3 Current scoring and planned Thurstonian IRT

The current (version 1) gift scoring is simple and fully transparent. Each response is recoded from -2 to +2. The gift the person leans toward gains that amount and the other gift loses the same amount. Over four appearances, each gift's raw score ranges from -8 to +8, rescaled to 0-100 for display. Gifts are ranked by raw score, with ties broken by the number of pairs the gift "won."

This classical scoring has a known limitation. Because every point gained by one gift is lost by another, the scores sum to a constant across the 18 gifts. Such scores are **ipsative**: they describe the relative order of gifts within a person, but they are not suitable for comparing people with one another, and they distort correlations and reliability estimates (Hicks, 1970).

The planned solution is **Thurstonian item response theory** (Brown & Maydeu-Olivares, 2011, 2018). Thurstonian IRT models each forced-choice response as the outcome of a comparison between two latent utilities, each driven by the respondent's standing on an underlying trait. Under suitable conditions, the model recovers *normative* trait scores from forced-choice data. Open-source software exists to fit these models (Bürkner, 2019). We will move to Thurstonian scoring once roughly 500 completions are available.

We note the method's limits honestly. Simulation work shows that normative recovery is better when there are many traits, when items within a block differ in how strongly and in which direction they are keyed, and when the trait correlations are not extreme, and that it can be poor in less favorable designs (Bürkner, Schulte & Holling, 2019). Our design, with 18 traits but only positively keyed statements and four appearances each, may limit how well normative scores can be recovered. Whether Thurstonian scoring delivers usable between-person scores for the Summit Map is an empirical question that our first-phase analysis will answer. Until then, we report gifts only within-person.

5.4 Why the gift ranking is within-person, and what that means

The gift order in a report answers the question "which of *my* gifts do I lean on most?" It does not answer "do I have more of this gift than other people?" Two people can both have Builder as their #1 gift while differing greatly in how strongly Builder-like they are compared with the general population.

For the person taking the assessment, this is usually the right question. Personal development decisions are within-person decisions: which tools to reach for, which shadow to watch. For anyone who wants to compare people, it is the wrong question, and that is one reason the Summit Map must never be used for selection (Section 9.2). Every report states this in plain language.

5.5 Foundation: Likert items with a 30-day frame

Each of the nine squares is measured by three statements about the last 30 days, rated on five labeled points from "Not at all" to "Completely." Two statements in each square are positively keyed and one is reverse-keyed. Adjacent items never come from the same row or the same column of the grid.

Illustrative examples only. These items were written for this paper. They are not items from the Summit Map.

In the last 30 days...

"I ended most days knowing what I had moved forward." (positively keyed)

"Other people's requests decided how most of my evenings went." (reverse-keyed)

Three choices deserve explanation.

- **A 30-day frame.** Foundation is a state. Asking about a defined recent period, rather than about "the kind of person I am," aligns the measure with the construct and makes it sensitive to change. It also makes items harder to answer with a flattering self-image.
- **Concrete, behavioral wording.** Items describe observable behavior and specific experiences where possible, which reduces ambiguity and the scope for interpretation.
- **Labeled response options.** Every point on the scale is labeled. Translators keep the labels evenly spaced in meaning (Section 7.4).

Scoring. Reverse-keyed items are flipped ($6 - x$), the three items are averaged, and the mean is rescaled to 0–100. Each square is banded: **Built** (70–100), **Building** (40–69) or **Bare** (0–39). Level scores (Structure, Stability, Direction) and column scores (Think, Act, Be) are the means of their three squares.

The bands are currently **criterion-referenced**. They describe where answers fall on the response scale (for example, a Built square averages "Mostly" or higher), not where a person stands relative to others. Once a sufficiently large and diverse sample exists, we plan to report norm-referenced

percentiles alongside the bands. The 70 and 40 cut points are provisional and will be reviewed against data.

A known limitation. Three items per square is short. Short scales are efficient and reduce fatigue, but their reliability is limited, and a single reverse-keyed item per square carries much of the weight of the consistency check. The pilot item bank was larger, and the final three items per square will be reviewed once real-world data are available. If reliability is inadequate for any square, we will lengthen it.

5.6 Reverse keying and response validity

Reverse-keyed items help detect inattentive and acquiescent responding, but they introduce their own problems. They can be misread, especially when negated, and they can behave differently across cultures (Weijters, Baumgartner & Schillewaert, 2013; Wong, Rindfleisch & Burroughs, 2003). We use one reverse-keyed item per square, phrased without double negatives.

The engine computes three response-validity flags, following standard approaches to detecting careless responding (Meade & Craig, 2012; Curran, 2016):

Flag	Rule
Straight-lining	Identical answers to all 27 foundation statements
Inconsistent	In three or more squares, strong agreement with both the positive items and the reverse-keyed item
Too fast	Completion in under three minutes

A flag does not invalidate a report or accuse the person of anything. It softens the report's language, explains that the results may be less precise than usual, and invites a retake. Flagged responses are excluded from norms and from research analyses.

6. The validity program and honest claims

6.1 Earned claims: what we say, and when

"Science-backed" is a claim about evidence, and evidence accumulates. We follow a public rule: **we never claim more than our data show**. Each claim is earned only when its evidence exists.

Stage	Evidence required	Public claim
Launch	Constructs grounded in published research; transparent scoring	"Built by a PhD psychologist on established research."
~500 completions	Adequate reliability; internal structure supported	"Research-based and reliable."
Convergent and predictive results	Sensible relations with established measures; prediction of 90-day outcomes	"Validated," stated with the specific evidence
Full technical manual	Complete validity evidence, norms and invariance results (target: within 12 months of launch)	Claims specified in the manual

Our approach follows the modern, unified view of validity, which treats validity as an accumulation of evidence for a specific interpretation and use of scores, not as a property a test simply has (Cronbach & Meehl, 1955; Messick, 1995; AERA, APA & NCME, 2014).

6.2 Reliability

We will estimate internal consistency for each square and level using McDonald's omega (McDonald, 1999), with a working target of .70 for squares. For the gifts, we will estimate the reliability of Thurstonian trait scores where the model permits. We will estimate test–retest reliability over two to three weeks, a period over which neither gifts nor foundation should change much, and stability over 90 days, over which gifts should hold steady and foundation should move. We will also estimate the reliability of the Alignment score, which we expect to be lower than that of the squares themselves.

6.3 Structure: the crossed factor model

The grid makes a strong structural claim: each foundation item reflects both a **level** (Structure, Stability or Direction) and a **column** (Think, Act or Be). We will test this with a *crossed* confirmatory factor model, in which each item loads on one level factor and one column factor. The logic resembles a multitrait–multimethod analysis (Campbell & Fiske, 1959). If both sets of factors are needed to explain the data, both dimensions of the grid are real. If one set collapses, the grid will need to change, and we will say so. We will use conventional fit criteria (Hu & Bentler, 1999) as guides, not as pass–fail thresholds.

6.4 Convergent and discriminant validity

Participants who opt in will complete a short set of established, freely available measures after their report: the Satisfaction With Life Scale (Diener, Emmons, Larsen & Griffin, 1985), the presence subscale of the Meaning in Life Questionnaire (Steger et al., 2006), the Brief Self-Control Scale (Tangney et al., 2004) and a public-domain Big Five measure from the International Personality Item Pool (Goldberg et al., 2006). We expect, for example, that Purpose & Values will correlate with presence of meaning, and that Follow-Through will correlate with self-control and conscientiousness, but that neither will be redundant with them.

6.5 Incremental validity beyond the Big Five

The most important skeptical question is: *isn't this just the Big Five in a costume?* The Big Five traits predict many important life outcomes (Roberts, Kuncel, Shiner, Caspi & Goldberg, 2007), and any new measure should show that it adds something. We will test whether foundation scores predict life satisfaction, meaning and 90-day goal progress *beyond* the Big Five. Our reason to expect added value is theoretical: the foundation measures current, changeable conditions in a person's life, which trait inventories are not designed to capture. If it adds nothing, that will be an important finding, and we will publish it.

6.6 Prediction over 90 days

The strongest test of the model is whether it predicts what happens next. Our pre-registered 90-day study (Nosek, Ebersole, DeHaven & Mellor, 2018) will test four hypotheses:

1. Lower foundation scores predict less progress on the person's chosen 90-day focus.
2. Building the Next Camp square predicts broader gains in the foundation and in well-being.
3. Gains at the Structure level precede gains at the Stability level, which precede gains at the Direction level (the build-order hypothesis).
4. Closing alignment gaps predicts gains in well-being.

Hypothesis 3 is the model's boldest developmental claim. If the data do not support it, we will revise the model's language accordingly.

6.7 The blind-report test

Almost everyone accepts vague, flattering personality descriptions as accurate. Forer (1949) gave students an identical personality sketch, assembled largely from horoscope-style statements, and most rated it as a good description of themselves. Meehl (1956) named this the "Barnum effect," and a substantial literature has confirmed it (Dickson & Kelly, 1985). It means that people's ratings of a report's accuracy say little about whether the report is personal.

We will run a direct test. Participants will see their own report and a report generated from another person's answers, without knowing which is which, and rate each for accuracy. If the Summit Map is personal, people should rate their own report as more accurate significantly more often than

chance. A related test will check whether people endorse the shadow predicted for them more than other shadows. We will report the results whatever they are.

6.8 Fairness and invariance

We will test whether the instrument measures the same constructs in the same way across gender, age groups, life stages and languages, using standard measurement invariance procedures (Section 7.5).

6.9 What is not yet established

To be specific, at the time of writing **none of the following has yet been demonstrated** for the Summit Map:

- The reliability of the squares, the gifts or the Alignment score.
- The crossed 3×3 factor structure.
- The relationships with established measures, or incremental validity beyond the Big Five.
- Prediction of 90-day outcomes.
- The Structure → Stability → Direction build order.
- The Shadow principle as a predictive claim.
- Resistance to the Barnum effect.
- Measurement invariance across groups or languages, and the equivalence of the 11 language versions.
- Normative scores for the gifts, or norm-referenced bands for the squares.
- That using the report causes improvement in people's lives.

What *is* established is the research on which each construct is based, and the transparency of the scoring. That is what our launch claim says, and no more.

7. Across cultures and walks of life

7.1 Independent and interdependent selves

Markus and Kitayama (1991) proposed that cultures differ in how they construe the self. An **independent** self-construal, common in North American and Western European settings, emphasizes the self as autonomous and defined by internal attributes. An **interdependent** self-construal, more common in many East Asian, Latin American, African and South Asian settings, emphasizes the self as embedded in relationships and defined by roles and obligations. These differences shape cognition, emotion and motivation. Later work found that relational self-construal also varies considerably *within* Western samples (Cross, Bacon & Morris, 2000).

This distinction matters for a model whose third column is called **Be**.

7.2 How **Be** and Purpose express themselves relationally

In an independent frame, "who I am" and "what I stand for" may be expressed as personal convictions and individual goals. In an interdependent frame, they are often expressed through roles and relationships: being a good child, parent, neighbor or community member, keeping harmony, honoring obligations. Research suggests that the sources and correlates of meaning and well-being differ across cultures in this way (Steger, Kawabata, Shimai & Otake, 2008; Uchida & Kitayama, 2009). The ideal emotional state also differs: some cultures value calm, low-arousal positive states more, and others value excitement and high-arousal states (Tsai, Knutson & Fung, 2006). Even self-concept clarity was found to be lower on average in a Japanese sample than in a Canadian one, which the authors interpreted as a cultural difference in how the self is construed, not a deficit (Campbell et al., 1996).

The Summit Map is designed to accommodate this. The **Be** column includes relationships explicitly through Boundaries & Bonds. The Purpose & Values square asks about living by one's values, not about pursuing individual ambition. Gifts such as Rope Team, Guardian, Cornerstone, Bridge Builder and Coach are relational by nature. And in adaptation, translators are instructed to preserve meaning rather than literal wording, so that an item about values can be answered naturally by someone whose values center on family and community.

We do not assume this is sufficient. Whether the **Be** column functions equivalently across cultures is a question for invariance testing, not for assumption.

7.3 Cautions about cultural dimension models and WEIRD samples

National culture models, such as Hofstede's dimensions (Hofstede, 2001), have been influential, but they have also been strongly criticized for treating nations as uniform and for inferring individual psychology from national averages (McSweeney, 2002). A meta-analysis of individualism and collectivism found that cultural differences were often smaller and less

consistent than assumed, and depended heavily on how they were measured (Oyserman, Coon & Kemmelmeier, 2002). A large multinational study found that cultural models of selfhood vary along many dimensions and do not reduce to a simple East–West split (Vignoles et al., 2016). We therefore do not tailor reports by nationality or apply cultural "profiles." We adapt language and test equivalence instead.

The broader problem is sampling. Henrich, Heine and Norenzayan (2010) showed that most published psychological research relies on samples from Western, Educated, Industrialized, Rich and Democratic (WEIRD) societies, which are often unusual in global terms. Much of the research beneath the Summit Map shares this limitation. Effects established in American undergraduates, such as the benefits of reappraisal over suppression, may differ elsewhere. We name this limitation plainly, and we will recruit validation samples beyond WEIRD populations.

There are also measurement-specific hazards. Response styles such as acquiescence and extreme responding vary across cultures (Johnson, Kulesa, Cho & Shavitt, 2005). The **reference-group effect** means that people judge themselves against those around them, which can mask or even reverse real cross-cultural differences in Likert ratings (Heine, Lehman, Peng & Greenholtz, 2002). Reverse-keyed items can behave differently across languages (Wong et al., 2003). For these reasons, we do not compare raw scores across countries, and we will treat cross-cultural comparisons as valid only after invariance is established.

7.4 Adaptation into 11 languages

The Summit Map is available in 11 languages: English, Arabic, Chinese, French, German, Hindi, Italian, Japanese, Korean, Portuguese and Spanish. The adaptation process follows the International Test Commission's Guidelines for Translating and Adapting Tests (International Test Commission, 2017). The core principle of those guidelines is that adaptation must preserve the *meaning and psychological properties* of items, not merely their words.

Our process has these steps:

1. **A glossary first.** For each language, a glossary fixes the translation of every gift name, square name, level and key term before any item is translated, so terms are used consistently.
2. **Meaning, intensity and difficulty.** Each item is adapted to preserve its meaning, its intensity and how hard it is to endorse. A "Mostly" answer must mean the same thing in every language, and the five response labels are kept evenly spaced.
3. **Balanced pairs.** In each forced-choice pair, the two statements are kept equally attractive in the target language, because a translation that makes one statement sound more appealing would bias the result.
4. **Natural, plain register.** Items use warm, everyday language at roughly an eighth-grade level, with the form of address that is culturally normal for personal-development material, and gender-neutral wording wherever the language allows.

5. **Culturally fitting names.** Gift names are chosen to be clear and memorable in each language rather than translated literally, and to avoid collisions with well-known type names in that language.
6. **Local safety resources.** Crisis lines are retained and supplemented with well-known local equivalents where they exist.
7. **Structural checks and reviewer notes.** Automated checks confirm that every language version has exactly the same structure as the source. Translators record every adaptation choice that a native-speaking reviewer should check.

We want to be clear about the current state of this work. The 11 versions were produced to a written adaptation standard, but **independent native-speaker review, back-translation of key items, and cognitive interviews in each language are still to be completed.** The ITC guidelines also require empirical evidence of equivalence. Until that evidence exists, non-English versions should be regarded as provisional.

7.5 Measurement invariance testing

To compare scores meaningfully across groups, an instrument must measure the same thing in the same way in each group. Measurement invariance testing checks this in steps (Vandenberg & Lance, 2000; Putnick & Bornstein, 2016):

- **Configural invariance:** the same items relate to the same factors in every group.
- **Metric invariance:** the items relate to the factors with equal strength, so relationships between scores can be compared.
- **Scalar invariance:** the items have the same starting points, so average scores can be compared.

We will test invariance across language versions, gender, age groups and life stages as samples allow, using change-in-fit criteria suited to large samples (Chen, 2007). Where full invariance fails, we will investigate the items responsible and report which comparisons are and are not justified.

7.6 Across life stages, circumstances and hard seasons

The Summit Map is meant to be useful to a nineteen-year-old and a seventy-year-old, to a business owner and to someone between jobs. The person tells us their life stage in the first section, and the report frames the climb for that stage without changing the scoring:

Life stage	How the climb is framed
Students	Learning what you care about apart from grades and other people's plans for you
Early career	Trying paths and noticing which work and which people fit your values
Mid-career	Checking whether the life you have built still points at what matters most
Business owners	Making sure the business serves your values and the life you want
Caregivers	Finding small, steady ways to honor your own values while caring for others
People in transition	A chance to choose again what matters most and where to head next
Retirees	How to spend time, wisdom and energy on what matters

These framings draw on developmental research. Emerging adulthood is a distinct period of identity exploration (Arnett, 2000). As people perceive their time horizons as shorter, they tend to prioritize emotionally meaningful goals and relationships (Carstensen, Isaacowitz & Charles, 1999). Personality traits continue to change across the life course, especially in young adulthood (Roberts, Walton & Viechtbauer, 2006). Whether the instrument functions equivalently across age groups is part of our invariance testing.

Socioeconomic circumstances. The Summit Map measures how well-built a life *feels from the inside*, not income, status or achievement. That is a deliberate choice, but it has implications. A person working two jobs may have little room for routines, not for lack of skill but for lack of time. There is evidence, still debated, that financial strain consumes cognitive resources and can impair decision-making (Mani, Mullainathan, Shafir & Zhao, 2013). For this reason, Next Camp actions are sized to the time the person says they have, with a light option measured in minutes rather than hours, and the report never implies that a thin foundation is a personal failing.

Hard seasons, and why a high purpose score in hardship can be genuine. Suppose a person whose life is visibly difficult, through illness, grief, job loss or caregiving, scores high on the Summit level. Is the score wrong? Not necessarily. Frankl's central claim, and the research described in Section 3.8, suggest that people can hold real meaning and purpose in hardship. Resilience is the most common response to adversity (Bonanno, 2004), and people in poorer nations report *more* meaning on average, not less (Oishi & Diener, 2014). We therefore do not treat a high purpose score in hardship as an error.

But self-report can also be rosy. People defend their self-image, and reports of positive change after adversity do not always match measured change. Frazier and colleagues (2009) found that self-reported post-traumatic growth correlated poorly with actual change measured before and after a traumatic event. A responsible instrument must allow for both possibilities.

How the grid checks rosy self-report. The Summit Map does not read the Summit level alone. Four safeguards work together:

1. **Concrete, time-bound items.** Specific questions about the last 30 days are harder to inflate than general statements about the kind of person one is.
2. **Reverse-keyed items in every square** and the inconsistency flag.
3. **The grid checks itself.** If the Direction level is high while the Structure level is thin, the report says that the person's sense of purpose is strong but the trail beneath it is still thin, and that purpose becomes a life when ordinary days carry it. The Alignment score makes a similar check within each level: a high Purpose & Values score with a low Aimed Effort score is named as a gap, not celebrated as a summit.
4. **Forced-choice gifts**, which cannot all be rated highly.

This approach respects both possibilities. It does not tell a person in hardship that their purpose is an illusion, and it does not tell a person whose days do not yet reflect their values that they have arrived.

When every square is Bare, the report changes tone. It acknowledges that the person may be in a hard season, suggests one very small step, encourages them to talk with someone they trust, and provides crisis resources. The Summit Map is not a screening tool for depression or risk, and it does not claim to identify people in crisis (Section 9).

8. How people use it

8.1 Individuals

Most people take the Summit Map alone. The free result gives them their Summit Code, their top gifts, their Map with bands, and the name of their Next Camp. The full report adds the Shadow, Alignment by level, the quoted evidence, the full Think · Act · Be plan, a seven-day starter plan and a 30-day measure of success. The report is designed around one practical outcome: a person should finish reading it knowing what to do this week.

8.2 The 90-day retake

Gifts should hold steady. The foundation should move. The Summit Map is designed to be taken again after 90 days, and the second report shows what changed. There are three reasons for this design. Progress monitoring improves goal attainment (Harkin et al., 2016). Ninety days is long enough for meaningful change in routines and conditions, while acknowledging that habit formation often takes longer (Lally et al., 2010). And retake data are the core of our predictive validity program: they let us test whether the model's predictions about change come true.

8.3 Coaching

Coaches can use the Summit Map as a structured starting point for conversation. A meta-analysis of workplace coaching found positive effects on performance, well-being, coping and goal-directed self-regulation (Theeboom, Beersma & van Vianen, 2014). The Summit Map is designed to give a coaching relationship a shared language and a clear first focus, and the 90-day retake gives coach and client a way to see change. Coaches should treat scores as the client's self-description, to be explored together, not as a verdict.

8.4 Couples and teams (future)

We plan versions for couples and teams, where the Summit Code and the grid may help people understand how their gifts complement each other and where their shared foundation is thin. These uses are **not yet available and not yet validated**. They raise specific issues: within-person gift rankings cannot be used to compare people's levels of a gift, and in group settings there is a risk that self-report is shaped by what others will see. Any team version will be designed for development conversations only, and never for evaluation.

9. Ethics and limitations

9.1 Not a diagnosis

The Summit Map is a personal-growth tool. It is not a clinical instrument and it does not diagnose, screen for or rule out any mental health condition. A low foundation score is not a symptom, and a high one is not a clean bill of health. People who are struggling should speak with a qualified professional.

9.2 Never for hiring or selection

The Summit Map must not be used to hire, promote, select, place or evaluate people. There are several independent reasons:

- **Within-person scores.** Gift rankings describe what a person leans on, not how much of a gift they have relative to others.
- **State measurement.** Foundation scores describe the last 30 days and are expected to change with circumstances, including circumstances outside a person's control.
- **Self-report in high stakes.** Self-report measures can be faked when outcomes depend on them, and the use of personality measures in selection is contested even for well-validated instruments (Morgeson et al., 2007).
- **No adverse-impact evidence.** The professional standards for testing require evidence that a measure is valid and fair for each specific use (AERA, APA & NCME, 2014). No such evidence exists for selection uses of the Summit Map, and we do not intend to seek it.

9.3 Data privacy

The Summit Map involves personal reflections, and we treat them with care. Our commitments are:

- Scoring happens on our server. The scoring key is never sent to the browser.
- Questions about the respondent's life context are used only to tailor the report and are never scored.
- Participation in research measures is optional and separately consented.
- Shareable result cards show the Summit Code but no scores by default.
- Research analyses use de-identified data, and published findings report only aggregate results.

The privacy policy published on our site governs how data are collected, stored and deleted, and these commitments are principles it must honor. Where data protection law, such as the EU General Data Protection Regulation, gives people additional rights, we honor those rights.

9.4 Crisis resources

When a profile suggests a very hard season, the report includes crisis resources: in the United States, call or text 988, the Suicide & Crisis Lifeline, and elsewhere a local emergency number or a well-known local crisis line where one is available in that language. We include these resources out of care, not because the instrument can detect crisis. It cannot, and no one should rely on it to do so.

9.5 The limits of self-report

Every score in the Summit Map comes from what people say about themselves. Self-report is the most practical way to learn about people's inner lives, but it has well-documented limits. People's self-assessments are often only modestly related to objective measures of their performance (Dunning, Heath & Suls, 2004). Responses are influenced by the wish to look good to others and to oneself (Paulhus, 1984). When all measures come from the same source, their correlations can be inflated by shared method variance (Podsakoff, MacKenzie, Lee & Podsakoff, 2003), which applies directly to our convergent validity studies. And people judge themselves against their own reference group (Heine et al., 2002).

We mitigate these limits where we can: concrete, time-bound items; reverse keying; validity flags; forced-choice gifts; a grid that checks its own Summit; and outcome measures collected 90 days later rather than at the same moment. But we cannot eliminate them. A Summit Map report is a structured mirror. It reflects what a person reports about their life, organized by research, and it is most useful when held up against what the person and the people who know them can see.

9.6 Other limitations

- The research on which the model rests comes mainly from WEIRD samples (Section 7.3).
- Several of the model's most distinctive claims, including the build order, the Shadow principle and the role of alignment, are hypotheses still to be tested.
- Each square has only three items, which limits reliability.
- Gift scores are currently ipsative, and Thurstonian scoring may not fully solve that problem in our design.
- Band cut points are provisional and not normed.
- The report's recommendations are drawn from evidence-based techniques, but the report itself has not been tested as an intervention.
- The Summit Map is a commercial product, created by its author. That is a conflict of interest. Our safeguards are pre-registration, a published technical manual, independent expert review and a commitment to publish results whatever they show.

10. Conclusion: Now Go and Do

The Summit Map starts from a simple observation. Most people who stall are not short of talent or desire. Their gifts are running on ground that cannot carry them, or their thoughts, actions and identity are pulling in different directions.

The model answers with a map. Across the grid, Think · Act · Be ask whether a person's mind, behavior and sense of self tell the same story. Up the grid, Structure, Stability and Direction trace the climb from a Base Camp of priorities, routines and protected relationships, through footing that holds under pressure, to a Summit: a purpose a person can see, pursue and live. Eighteen gifts describe how a person naturally builds, and their shadows warn how those gifts misfire on thin ground. The Next Camp turns all of it into one concrete next step.

Every element rests on published research, and every distinctive claim is framed as a hypothesis to be tested. We have described what we know, what we do not yet know, and how we will find out. We will publish what we learn, including what proves us wrong.

The report ends with the same words this paper does, because insight is only the start of the work.

Now Go and Do.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Aeon, B., Faber, A., & Panaccio, A. (2021). Does time management work? A meta-analysis. *PLoS ONE*, *16*(1), e0245066.
- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*(5), 469–480.
- Ashforth, B. E., Kreiner, G. E., & Fugate, M. (2000). All in a day's work: Boundaries and micro role transitions. *Academy of Management Review*, *25*(3), 472–491.
- Baumeister, R. F., Vohs, K. D., Aaker, J. L., & Garbinsky, E. N. (2013). Some key differences between a happy life and a meaningful life. *The Journal of Positive Psychology*, *8*(6), 505–516.
- Bem, D. J. (1972). Self-perception theory. *Advances in Experimental Social Psychology*, *6*, 1–62.
- Biswas-Diener, R., Kashdan, T. B., & Minhas, G. (2011). A dynamic approach to psychological strength development and intervention. *The Journal of Positive Psychology*, *6*(2), 106–118.
- Bolier, L., Haverman, M., Westerhof, G. J., Riper, H., Smit, F., & Bohlmeijer, E. (2013). Positive psychology interventions: A meta-analysis of randomized controlled studies. *BMC Public Health*, *13*, 119.
- Bonanno, G. A. (2004). Loss, trauma, and human resilience: Have we underestimated the human capacity to thrive after extremely aversive events? *American Psychologist*, *59*(1), 20–28.
- Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement*, *71*(3), 460–502.
- Brown, A., & Maydeu-Olivares, A. (2018). Ordinal factor analysis of graded-preference questionnaire data. *Structural Equation Modeling: A Multidisciplinary Journal*.
- Bryan, C. J., Walton, G. M., Rogers, T., & Dweck, C. S. (2011). Motivating voter turnout by invoking the self. *Proceedings of the National Academy of Sciences*, *108*(31), 12653–12656.
- Buckingham, M., & Clifton, D. O. (2001). *Now, discover your strengths*. Free Press.
- Bürkner, P.-C. (2019). thurstonianIRT: Thurstonian IRT models in R. *Journal of Open Source Software*.
- Bürkner, P.-C., Schulte, N., & Holling, H. (2019). On the statistical and practical limitations of Thurstonian IRT models. *Educational and Psychological Measurement*, *79*(5), 827–854.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105.
- Campbell, J. D., Trapnell, P. D., Heine, S. J., Katz, I. M., Lavallee, L. F., & Lehman, D. R. (1996). Self-concept clarity: Measurement, personality correlates, and cultural boundaries. *Journal of Personality and Social Psychology*, *70*(1), 141–156.
- Cao, M., & Drasgow, F. (2019). Does forcing reduce faking? A meta-analytic review of forced-choice personality measures in high-stakes situations. *Journal of Applied Psychology*, *104*(11), 1347–1368.
- Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, *54*(3), 165–181.

- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504.
- Clear, J. (2018). *Atomic habits: An easy and proven way to build good habits and break bad ones*. Avery.
- Clifton, D. O., & Harter, J. K. (2003). Investing in strengths. In K. S. Cameron, J. E. Dutton, & R. E. Quinn (Eds.), *Positive organizational scholarship: Foundations of a new discipline* (pp. 111–121). Berrett-Koehler.
- Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*, 113(3), 492–511.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302.
- Cross, S. E., Bacon, P. L., & Morris, M. L. (2000). The relational-interdependent self-construal and relationships. *Journal of Personality and Social Psychology*, 78(4), 791–808.
- Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, 66, 4–19.
- Damon, W., Menon, J., & Bronk, K. C. (2003). The development of purpose during adolescence. *Applied Developmental Science*, 7(3), 119–128.
- Dickson, D. H., & Kelly, I. W. (1985). The "Barnum effect" in personality assessment: A review of the literature. *Psychological Reports*, 57(2), 367–382.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49(1), 71–75.
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087–1101.
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69–106.
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. Dryden Press.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.
- Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, 80(6), 1011–1027.
- Forer, B. R. (1949). The fallacy of personal validation: A classroom demonstration of gullibility. *Journal of Abnormal and Social Psychology*, 44(1), 118–123.
- Frankl, V. E. (2006). *Man's search for meaning*. Beacon Press. (Original work published 1946)
- Frazier, P., Tennen, H., Gavian, M., Park, C., Tomich, P., & Tashiro, T. (2009). Does self-reported posttraumatic growth reflect genuine positive change? *Psychological Science*, 20(7), 912–919.
- Gerber, A. S., Huber, G. A., Biggers, D. R., & Hendry, D. J. (2016). A field experiment shows that subtle linguistic cues might not affect voter behavior. *Proceedings of the National Academy of Sciences*.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1), 84–96.
- Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54(7), 493–503.

- Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology*, 38, 69–119.
- Grant, A. M., & Schwartz, B. (2011). Too much of a good thing: The challenge and opportunity of the inverted U. *Perspectives on Psychological Science*, 6(1), 61–76.
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, 2(3), 271–299.
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, 85(2), 348–362.
- Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., ... Zwienerberg, M. (2016). A multisite preregistered paradigmatic test of the ego-depletion effect. *Perspectives on Psychological Science*, 11(4), 546–573.
- Harkin, B., Webb, T. L., Chang, B. P. I., Prestwich, A., Conner, M., Kellar, I., Benn, Y., & Sheeran, P. (2016). Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence. *Psychological Bulletin*, 142(2), 198–229.
- Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference-group effect. *Journal of Personality and Social Psychology*, 82(6), 903–918.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83.
- Hicks, L. E. (1970). Some properties of ipsative, normative, and forced-choice normative measures. *Psychological Bulletin*, 74(3), 167–184.
- Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94(3), 319–340.
- Hilgard, E. R. (1980). The trilogy of mind: Cognition, affection, and conation. *Journal of the History of the Behavioral Sciences*, 16(2), 107–117.
- Hill, P. L., & Turiano, N. A. (2014). Purpose in life as a predictor of mortality across adulthood. *Psychological Science*, 25(7), 1482–1486.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Sage.
- Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. *PLoS Medicine*, 7(7), e1000316.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- International Test Commission. (2017). *The ITC guidelines for translating and adapting tests* (2nd ed.). International Test Commission.
- Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology*.
- Kaiser, R. B., & Overfield, D. V. (2011). Strengths, strengths overused, and lopsided leadership. *Consulting Psychology Journal: Practice and Research*.
- Kant, I. (1987). *Critique of judgment* (W. S. Pluhar, Trans.). Hackett. (Original work published 1790)
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284.

- Lally, P., van Jaarsveld, C. H. M., Potts, H. W. W., & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology*, 40(6), 998–1009.
- Lazarus, R. S. (2003). Does the positive psychology movement have legs? *Psychological Inquiry*, 14(2), 93–109.
- Lewin, K. (1936). *Principles of topological psychology*. McGraw-Hill.
- Lincoln, A. (1953). Annual message to Congress, December 1, 1862. In R. P. Basler (Ed.), *The collected works of Abraham Lincoln* (Vol. 5). Rutgers University Press. (Original work delivered 1862)
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9), 705–717.
- Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science*, 341(6149), 976–980.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253.
- Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, 41(9), 954–969.
- McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality*, 63(3), 365–396.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates.
- McSweeney, B. (2002). Hofstede's model of national cultural differences and their consequences: A triumph of faith—a failure of analysis. *Human Relations*, 55(1), 89–118.
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455.
- Meehl, P. E. (1956). Wanted—a good cookbook. *American Psychologist*, 11(6), 263–272.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749.
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, 102(2), 246–268.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, 60(3), 683–729.
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115(11), 2600–2606.
- Oettingen, G. (2012). Future thought and behaviour change. *European Review of Social Psychology*, 23(1), 1–63.
- Oishi, S., & Diener, E. (2014). Residents of poor nations have a greater sense of meaning in life than residents of wealthy nations. *Psychological Science*, 25(2), 422–430.
- Oyserman, D. (2009). Identity-based motivation: Implications for action-readiness, procedural-readiness, and consumer behavior. *Journal of Consumer Psychology*, 19(3), 250–260.
- Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, 128(1), 3–72.
- Park, C. L. (2010). Making sense of the meaning literature: An integrative review of meaning making and its effects on adjustment to stressful life events. *Psychological Bulletin*, 136(2), 257–301.

- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598–609.
- Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. American Psychological Association; Oxford University Press.
- Plato. (1992). *Republic* (G. M. A. Grube, Trans.; C. D. C. Reeve, Rev.). Hackett.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903.
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90.
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2(4), 313–345.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, 132(1), 1–25.
- Rogers, C. R. (1959). A theory of therapy, personality, and interpersonal relationships, as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of a science: Vol. 3. Formulations of the person and the social context* (pp. 184–256). McGraw-Hill.
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, 25, 1–65.
- Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, 55(1), 5–14.
- Sheeran, P. (2002). Intention–behavior relations: A conceptual and empirical review. *European Review of Social Psychology*, 12(1), 1–36.
- Sheeran, P., & Webb, T. L. (2016). The intention–behavior gap. *Social and Personality Psychology Compass*, 10(9), 503–518.
- Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, 76(3), 482–497.
- Snyder, C. R. (2002). Hope theory: Rainbows in the mind. *Psychological Inquiry*, 13(4), 249–275.
- Starcke, K., & Brand, M. (2012). Decision making under stress: A selective review. *Neuroscience and Biobehavioral Reviews*, 36(4), 1228–1248.
- Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The Meaning in Life Questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, 53(1), 80–93.
- Steger, M. F., Kawabata, Y., Shimai, S., & Otake, K. (2008). The meaningful life in Japan and the United States: Levels and correlates of meaning in life. *Journal of Research in Personality*, 42(3), 660–678.
- Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state–trait theory and research in personality and individual differences. *European Journal of Personality*, 13(5), 389–408.
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72(2), 271–324.

- Theeboom, T., Beersma, B., & van Vianen, A. E. M. (2014). Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *The Journal of Positive Psychology*, 9(1), 1–18.
- Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology*, 90(2), 288–307.
- Uchida, Y., & Kitayama, S. (2009). Happiness and unhappiness in east and west: Themes and variations. *Emotion*, 9(4), 441–456.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4–70.
- Vignoles, V. L., Owe, E., Becker, M., Smith, P. B., Easterbrook, M. J., Brown, R., ... Bond, M. H. (2016). Beyond the 'east–west' dichotomy: Global variation in cultural models of selfhood. *Journal of Experimental Psychology: General*, 145(8), 966–1000.
- Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, 132(2), 249–268.
- Weijters, B., Baumgartner, H., & Schillewaert, N. (2013). Reversed item bias: An integrative model. *Psychological Methods*, 18(3), 320–334.
- White, C. A., Utzl, B., & Holder, M. D. (2019). Meta-analyses of positive psychology interventions: The effects are much smaller than previously reported. *PLoS ONE*, 14(5), e0216588.
- Wong, N., Rindfleisch, A., & Burroughs, J. E. (2003). Do reverse-worded items confound measures in cross-cultural consumer research? The case of the material values scale. *Journal of Consumer Research*, 30(1), 72–91.
- Wong, P. T. P. (2011). Positive psychology 2.0: Towards a balanced interactive model of the good life. *Canadian Psychology*, 52(2), 69–81.
- Wood, W., & Neal, D. T. (2007). A new look at habits and the habit–goal interface. *Psychological Review*, 114(4), 843–863.
- Wood, W., & Rünger, D. (2016). Psychology of habit. *Annual Review of Psychology*, 67, 289–314.