Practical Assessment, Research, and Evaluation

Volume 27 Article 4

2022

An Enhancement to the Theory and Measurement of Purpose

Larry H. Ludlow
Boston College

Theresa O'Keefe
Boston College

Henry Braun
Boston College

Ella Anghel
Boston College

Olivia Szendey
Boston College

See next page for additional authors

Follow this and additional works at: https://scholarworks.umass.edu/pare

Recommended Citation

Ludlow, Larry H.; O'Keefe, Theresa; Braun, Henry; Anghel, Ella; Szendey, Olivia; Matz, Christina; and Howell, Burton (2022) "An Enhancement to the Theory and Measurement of Purpose," *Practical Assessment, Research, and Evaluation*: Vol. 27, Article 4.
DOI: https://doi.org/10.7275/c5jb-rr95
Available at: https://scholarworks.umass.edu/pare/vol27/iss1/4

This Article is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Practical Assessment, Research, and Evaluation by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

An Enhancement to the Theory and Measurement of Purpose

Cover Page Footnote

We thank Dr. Nicole Brocato, Laura Hix, MA and Eric Gaudiello, MA at Wake Forest University's Wellbeing Collaborative for their assistance with the initial pilot sampling and administration, and Dr. Jessica Greene Associate Vice Provost, Assessment & Accreditation at Boston College for validation sampling assistance. Funding support was provided by the Boston College, Lynch School of Education and Human Development Dean's Office; and a Research Across the Disciplines and Schools award from the Boston College Provost Office.

Authors

Larry H. Ludlow, Theresa O'Keefe, Henry Braun, Ella Anghel, Olivia Szendey, Christina Matz, and Burton Howell

This article is available in Practical Assessment, Research, and Evaluation: https://scholarworks.umass.edu/pare/vol27/iss1/4

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Copyright is retained by the first or sole author, who grants right of first publication to *Practical Assessment, Research & Evaluation*. Permission is granted to distribute this article for nonprofit, educational purposes if it is copied in its entirety and the journal is credited. PARE has the right to authorize third party reproduction of this article in print, electronic and database forms.

Volume 27 Number 4, March 2022 ISSN 1531-7714

An Enhancement to the Theory and Measurement of Purpose[1]

Larry H. Ludlow, *Department of Measurement, Evaluation, Statistics, and Assessment, Lynch School of Education and Human Development, Boston College, Chestnut Hill, MA, USA*
Theresa O'Keefe, *School of Theology and Ministry, Boston College, Chestnut Hill, MA, USA*
Henry Braun, *Department of Measurement, Evaluation, Statistics, and Assessment, Lynch School of Education and Human Development, Boston College, Chestnut Hill, MA, USA*
Ella Anghel, *Department of Measurement, Evaluation, Statistics, and Assessment, Lynch School of Education and Human Development, Boston College, Chestnut Hill, MA, USA*
Olivia Szendey, *Department of Measurement, Evaluation, Statistics, and Assessment, Lynch School of Education and Human Development, Boston College, Chestnut Hill, MA, USA*
Christina Matz, *School of Social Work, Boston College, Chestnut Hill, MA, USA*
Burton Howell, *Intersections, Division of Mission and Ministry, Boston College, Chestnut Hill, MA, USA*

Development of purpose is an important goal of post-secondary education. This study advances the measurement of purpose by (a) enriching the construct through incorporating the facet of horizon; (b) providing a framework for Rasch/Guttman Scenario score interpretation; and (c) providing evidence of convergent, divergent, and known groups validity.

Introduction

Many higher education administrators and researchers argue that a college education is not only about future employability, but also about expanding a sense of purpose (see Glazer et al., 2017 for a review). Yet, in an age when metrics are a pervasive aspect of educational management and accountability, it is still difficult for colleges to measure reliably the presence and development of purpose in students; purpose is abstract and complex, requiring both a clear definition and a sophisticated measurement tool that can produce nuanced, interpretable, and actionable scores (Ludlow et al., 2021).

Under the auspices of the Boston College Living A Life of Meaning and Purpose (BC-LAMP) portfolio project, we are designing and testing various scales based on the so-called Rasch/Guttman Scenario (RGS) methodology (Ludlow et al., 2014). Our first effort, BC-LAMP-A, served as a "proof of concept" for measuring purpose using this approach (Ludlow et al., 2020). The present article describes and reports on our next, more substantial second step, BC-LAMP-B. Specifically our goal with BC-LAMP-B is to 1) introduce *horizon* as a facet (in addition to clarity, effort, and frequency) within each of three dimensions of purpose: meaningfulness, goal orientation, and beyond-the-self; 2) provide an interpretive framework

[1] We thank Dr. Nicole Brocato, Laura Hix, MA and Eric Gaudiello, MA at Wake Forest University's Wellbeing Collaborative for their assistance with the initial pilot sampling and administration, and Dr. Jessica Greene Associate Vice Provost, Assessment & Accreditation at Boston College for validation sampling assistance. Funding support was provided by the Boston College, Lynch School of Education and Human Development Dean's Office; and a Research Across the Disciplines and Schools award from the Boston College Provost Office.

for the scale scores; and 3) provide reliability and validity evidence for the measures. In the next section, we discuss the background and rationale for developing an enhanced measure of purpose..

Background and Rationale

Defining Purpose and the Claremont Purpose Scale

Although many definitions of the construct of purpose exist in the literature (see Martela and Steger, 2016 for a review), most involve having a meaningful goal and working towards achieving it (Emmons, 2005; McKnight & Kashdan, 2009; Singer, 1996). Other scholars add different elements such as life's mattering (Heintzelman & King, 2014), positive affect (Battista & Almond, 1973), or a sense of authenticity (Leontiev, 2006). For Damon et al., (2003) purpose is "a stable and generalized intention to accomplish something that is at once meaningful to the self and of consequence to the world beyond the self" (p. 121). Since we are particularly interested in purpose among college students, we adopted the definition developed by Damon and his colleagues for three reasons. First, according to Damon (2008), the definition is a "rough consensus" of how purpose had been defined over the years by developmental researchers. Second, the definition was developed with adolescents in mind, making it reasonably appropriate for undergraduates. Third, the definition of purpose comprises three distinct dimensions, which aligned well with the measurement methodology we intended to use.

Damon et al.'s (2003) definition of purpose led to the development of the *Claremont Purpose Scale* (CPS) by Bronk et al. (2018). The CPS is a 12-item instrument measuring the three dimensions of meaningfulness, goal orientation, and beyond-the-self, using four questions for each dimension. The CPS is relatively new, but it served as an attractive starting point for our work based on its clear conceptual framework and strong psychometric qualities (Bronk et al., 2018).

Limitation to CPS

In spite of its attractive features, the CPS has some limitations for our objectives. For one, the CPS uses traditional, short-stemmed, narrowly focused Likert-scored statements. As a result, score interpretation is clinically opaque in the sense that it is difficult to understand what an individual "looks like" for a given score (i.e., what does a score of 18 tell us about a person and how does that person differ from another with a 26?). As reported by Bronk et al. (2018):

> High scores mean participants meet all or most of the criteria for purpose, midrange scores mean they meet some of the criteria, and low scores mean participants meet few, if any criteria for purpose. (p. 110)

Such general statements limit understanding what a person may be thinking or experiencing at a given score level, what the difference between two scores means in substantive terms, and what deliberate steps might be taken to raise a person to a higher level (score). In short, Likert-based items are typically incapable of capturing the full complexity of a construct (Kyllonen & Bertling, 2014).

A second limitation lies in the definition of the construct. Damon (2008) recognized a limitation inherent in the three-dimensional definition of purpose when he further distinguished between "ignoble" and "noble" purpose, where noble purpose is "something worth doing and doing it in an honorable manner" (Damon, 2008). Others, too, have suggested extensions to Damon's (2008) definition of purpose (see Martela and Steger, 2016). The CPS, however, focuses on the three dimensions without further enhancement.

The third limitation lies in the response options. In the CPS, the three dimensions of purpose (meaningfulness, goal orientation, and beyond-the-self) are perfectly confounded with their response options (clarity, effort, and frequency, respectively). For example, the response options for questions of meaningfulness are framed in terms of levels of clarity. This confounding leads to an ambiguity that makes the appropriate interpretation of a CPS score problematic (e.g., is it a meaningfulness score or a clarity score?).

BC-LAMP-B directly confronts each of these limitations. The first is addressed by employing Rasch/Guttman Scenario (RGS) methodology, which we explain in detail in the Methods section below. The RGS methodology constitutes a systematic measurement approach to generate scenario-style items that are authentic descriptions of lived-experiences with respect to meaningfulness, goal orientation, and beyond-the-self. The resulting scores

are thereby linked to a rich representation of what a person "looks like" at a given score level, thus serving as a starting point for how a person's dimensions of purpose might be interpreted and, perhaps, enhanced.

The second limitation is addressed through the inclusion of "horizon" as a facet in the definition of each of the three dimensions of purpose. In the following subsection, we argue that the addition of horizon offers a stronger philosophical and developmental grounding for measuring and understanding meaningfulness, goal orientation, and beyond-the-self.

To address the third difficulty, the RGS approach retains the CPS's three dimensions of purpose but it simultaneously incorporates clarity, effort, frequency, and horizon as complementary facets in the development of each scenario for each dimension. Furthermore, each facet is characterized by different levels (or degrees) of intensity and these levels are systematically combined across the four facets to generate the scenarios comprising the items for a given dimension.

Enhancing the Construct

To enhance the construct of purpose we looked to two domains: philosophy and constructive-developmental psychology. Purpose, with its dimensions of meaningfulness, goal orientation, and beyond-the-self, inherently implies a conception of life's ultimate ends. Even if an ultimate end is never articulated, a choice for one good over another reflects a deeper conception of "the good life" (Taylor, 1989). This suggests that even if a person describes a purpose that is immediately self-interested, it implies an ultimate vision or goal. The philosophical term for this ultimate vision, acknowledged or not, is horizon. Gadamer (1999) defines horizon as follows:

> The horizon is the range of vision that includes everything that can be seen from a particular vantage point….A person who has no horizon does not see far enough and hence over values what is nearest to him. On the other hand, 'to have a horizon' means not being limited to what is nearby but being able to see beyond it. A person who has a horizon knows the relative significance of everything within this horizon, whether it is near or far, great or small. (p. 302).

It is worth noting that Gadamer's definition does not make a claim for a particular horizon (e.g., religious, philosophical), only that horizon expresses some sense of what is ultimate. We borrow the concept of horizon to add a fourth facet in the definition of each of the dimensions of purpose. In so doing, we create the opportunity to recognize and potentially measure "how far" the individual articulates their purpose: immediate to themselves, aware of social expectations, or claimed as ultimate.

The opportunity to incorporate horizon is particularly appropriate for measuring purpose in undergraduates, whose sense of purpose is likely changing and expanding during their college experience (King, 2009; Parks, 2011). The construct of purpose, enhanced by the philosophical concept of horizon, fits well with constructive-developmental psychology – a developmental theory that recognizes the expanding capacity of the human person to interpret the world and make meaning of it and of themselves (Baxter Magolda, 2008; Kegan 1982; Kegan, 1994; Parks, 2011). Constructive-developmental theory builds on the cognitive development theory of Piaget (Piaget, 1952). Kegan augments Piaget's focus on cognitive capacities by also attending to interpersonal and intrapersonal capacities, all of which pertain to a sense of purpose. Furthermore, Kegan extends the Piagetian model by identifying the capacity of the maturing person to dis-embed from the immediate, concrete, and self-referential proclivity of later childhood, so as to recognize and make sense of a wider world of persons, relationships, concepts, consequence, and time future and past (Kegan, 1982; Kegan 1994). The capacity to dis-embed and gain a wider horizon is directly connected to the individual's ability to make better sense of themselves in the world. This is particularly salient for college students, who are actively introduced to expanding spheres of ideas, people, and contexts. These college experiences can prompt students' ability to achieve a wider understanding of self and purpose and claim these as their own or, in other words, develop the capacity for self-authorship (Baxter Magolda, 2008; Baxter Magolda, 2009; King, 2009; Parks, 2011; Perez et al., 2012).

For both Kegan and Parks, through the development of meaning-making capacity, an individual is increasingly able to see themselves not

only as distinct from others but also multiply connected to the world around them. An expanding capacity to make meaning of one's life in the world translates into an expanding capacity to conceive of and act on one's purpose. Growth in meaning-making allows a person to recognize a wider circle of relationships, a more complex sense of agency and consequence, as well as a more capacious perspective on history and the future. Concomitantly, growth of meaning-making in a college student would be indicated by a move from purpose that is assumed, immediate, concrete, and self-referential, to purpose that is intentionally named by the student, socially constructed, and possibly, to purpose that is long-range, and defined in terms of ultimate ends (Baxter Magolda, 2008; Kegan, 1994; Parks, 2011).

With these considerations in mind, horizon nicely complements clarity, effort, and frequency in the framing of lived-experience scenarios for the three dimensions. Adding it allows us to measure, How distant is the goal? How expansive is the meaning? How far beyond the self is the consideration? Note that we do not treat horizon as a dimension to be measured in its own right; rather, at an operational level, we employ it as an additional facet to more broadly define each dimension. Furthermore, including horizon as a facet of each dimension helps us assess, from a constructive-developmental perspective, how those measured make sense of their purpose in light of their particular contexts (e.g. family, friends, school, work). Including horizon enhances the overarching construct of purpose, resulting in specific, interpretable, and actionable scores. In the next section, we introduce the Rasch/Guttman Scenario methodology (RGS) and then describe how it uses the four facets of clarity, effort, frequency, and horizon to measure purpose through the three dimensions of meaningfulness, goal orientation, and beyond-the-self.

Measurement Methodology

Theoretical Foundation of Instrument Design

As argued above, purpose is a complex construct, requiring a powerful tool to measure it well. We utilize Rasch/Guttman Scenario methodology (Ludlow et al., 2014) throughout the BC-LAMP Portfolio to accomplish this objective. This methodology offers a sophisticated approach that operationalizes a complex construct by first breaking it down into *dimensions*. Each dimension is then captured through a series of *scenarios* where each scenario is systematically constructed from multiple *facets*. The facets, aspects of life commonly used to describe the dimension, are represented as ordered, categorical variables distinguished by high, medium, and low levels of intensity. Facets can then be combined in different ways within a scenario.

This process yields an orderly progression of plausible lived experience scenarios defining each dimension, which in turn contributes to a more refined measurement of the overall construct. By responding to all the scenarios across the dimensions, the respondent generates scores that reflect the multi-dimensional complexity of the construct. In sum, this methodology combines aspects of Guttman facet theory and Rasch measurement principles to create a novel measurement tool with multiple conceptual and practical benefits.

Guttman Facet Theory and Sentence Mapping. Guttman facet theory (FT) enables the use of quantitative measures to convey qualitative information about those measured (Guttman, 1959). Rather than using traditional questions or statements as items in a scale, we use FT to generate rich qualitative descriptions of lived experiences expressed through scenarios. Hackett (2014) explains that an important contribution of facet theory to measurement is that it offers a mechanism for decomposing a construct "into parts that are significant to its subjects and then these components can be pulled together as a meaningful whole" (p. 4). Theory and empirical research provide guidance on decomposing a construct into dimensions and facets that resonate with respondents.

The process of creating the scenarios relies on another useful feature of facet theory, *sentence maps* (Borg & Shye, 1995). These maps diagram the syntactic structure of a scenario that then is filled in with different elements (words, phrases), define the range of facet levels available for a scenario, and illustrate the relationships among the selected facets (Guttman & Greenbaum, 1998). When facets are combined in different ways, using the sentence maps, they create the scenarios intended to capture the lived experience complexity of a dimension.

Not only is sentence mapping valuable for creating scenarios, but it also provides the basis for score

interpretation. By virtue of sentence mapping, a person's quantitative scale score can be interpreted in qualitative terms directly linked to the rich descriptions found in the scenarios. These descriptions are what Canter (2019) affirms is a central strength of facet theory: the opportunity to combine qualitative and quantitative methodologies. In the results section, we demonstrate how these scenarios, when linked to a person's score on a scale, support a rich qualitative interpretation of that score.

While Guttman facet theory offers a methodology for facet specification and scenario construction, Rasch measurement theory provides a foundation for how to systematically combine the facets in order to generate scenarios depicting a hypothesized progression from lower through higher levels of the focal dimension.

Rasch Measurement Principles and Psychometric Model. The construction of the scenarios defining a scale follows the measurement principles of Georg Rasch (1960/1980), summarized in Ludlow et al., (2014):

(1) A scale's items should measure a single dimension and capture a range of lower to higher levels spanning the dimension;

(2) The items written for the varying levels should define a clear, substantively meaningful, hierarchical progression with respect to the dimension; and

(3) The *a priori* underlying theory of what the dimension is measuring should be reflected in the empirical results.

These principles drive the operationalization of the construct through the subsequent specification of the dimensions and the development of the items. This means that items are deliberately written to follow principles 1 and 2. The third principle is a cornerstone to the evaluation of the content validity of the hypothesized construct's structure. That is, given that scenario content is based on specific combinations of the facet levels, this process generates an *a priori* expected structure, or ordering, of the scenarios in terms of their levels of intensity along the hypothesized continuum for each of the construct's dimensions. The subsequent analysis of the scenario scale response data then reveals the extent to which the empirical ordering of the scenarios does indeed reflect the hypothesized progressive structure for each dimension.

The Rasch rating scale model (*RSM*) is a common psychometric model used to analyze the extent to which data conform to the above measurement principles (Andrich, 1978; Wright & Masters, 1982). Specifically, do the empirical results provide construct validity evidence congruent with the hypothesized scale structures? Furthermore, the RSM is appropriate when the polytomous ordered response options are intended to be interpreted similarly for all items, as is the case in BC-LAMP-B. For a particular dimension, the model may be expressed as:

$$\pi_{nix} = \frac{e^{\sum_{j=0}^{x_{ni}} [\beta_n - (\delta_i + \tau_j)]}}{\sum_{k=0}^{m} e^{\sum_{j=0}^{k} [\beta_n - (\delta_i + \tau_j)]}} \qquad (1)$$

For our purposes, the important parts of the model are: π_{nix} is the probability of person *n* on scenario *i* responding to category *x;* β_n is person *n*'s estimated level, or location, on a particular dimension; δ_i is the difficulty estimate location associated with scenario *i*; and τ_j is the difficulty estimate corresponding to moving from one response category in a scenario to the next higher one. In the current context, scenario difficulty refers to how challenging it is for a student to score high on the scenario, with higher level response options corresponding to higher levels of intensity. All estimates are in the logit metric (Ludlow & Haley, 1995).

Construct maps, often called variable maps, are an essential component of Rasch model analyses. These maps are graphical representations that simultaneously display the scenario difficulty estimates (δ_i) and the respondent level of purpose estimates (β_n), both as locations along a quantitative continuum calibrated in logits. This continuum provides the basis for both a content and construct validity test for the scale and for a substantive diagnostic interpretation of a person's score.

Since scenario sets are designed to represent a sequence of increasingly higher levels for a particular dimension, the first analytic task is to compare the empirical difficulty order of the scenarios with their

hypothesized order (presented in detail below). Ideally, a well-designed scale will appear on the construct map as a ladder-like progression of scenario locations that are congruent with their hypothesized order. Next, we describe how we employed the Rasch/Guttman Scenario methodology in this study.

Development of the Boston College Living A Life of Meaning and Purpose-B Instrument. BC-LAMP-B comprises three scales, each of which measures one of the Damon et al. (2014) dimensions of *meaningfulness*, *goal orientation,* or *beyond-the-self*. Each scale consists of a set of scenarios describing fictional people exhibiting different degrees of purpose with respect to the corresponding dimension. The original CPS facets of clarity, effort, and frequency, as well as the added facet of horizon, are employed in the construction of the scenarios for each of the three dimensions.

The wording describing the facet levels for effort, clarity, and frequency is the same across the three dimensions. The horizon facet, however, has slightly different meanings and wordings for each dimension. For *meaningfulness*: a person with a "low" horizon focuses on what is meaningful only to themselves; a person with a "medium" horizon considers what is meaningful to others around them; and a person with a "high" horizon reflects on life's ultimate meaning. For *goal orientation*: a person with a "low" horizon only has immediate aims; a person with "medium" horizon has intermediate goals; and the one with "high" horizon has long-term goals. For *beyond-the-self*, the horizon facet measures the expansiveness of whom the individual aims to impact: a person with a "low" horizon focuses on themselves; a person with "medium" horizon attends to others around them; and the one with "high" horizon considers all people.

The sentence maps (see Figure 1) guided the development of the scenarios through combinations of facet levels to create realistic depictions of lived experiences. The initial scenarios were then modified to reduce redundancy, avoid obvious contradictions and "double-barreled" interpretations, and better illustrate their different intensity levels. Finally, connector words were added and the placement of facet phrases within scenarios was varied.

Participants were asked to compare the lived experience represented in a scenario with an assessment of their self-perceived circumstances. They were given the following instructions:

> This next set of unusual items present different descriptions of people. Think about how each description of person "X" compares to you, in general, at this particular point in your life and then respond with: (a) I have more direction in my life than "X"; (b) I feel like I have more direction in my life than "X"; (c) My direction in life is comparable to "X"; (d) I feel like "X" has more direction in life than me; (e) "X" has more direction in life than me.

We adopted the following scoring scheme: (a)=5 points, (b)=4, (c)=3, (d)=2, and (e)=1. The first scenario is a training item to minimize initial confusion attributable to the novel form of the items. We avoided the words "meaning and purpose" in order to minimize cueing or social desirability bias. We chose "direction" as a neutral term on the basis of which respondents compare themselves to the individual in the scenario. "I feel..." is used to express less certainty than "I have more...", thus representing a weaker response option. The scenarios within a dimension/scale were presented in random order. Each scenario response was scored and the sum of the scenario responses yielded a dimension, or scale, score for the individual. Higher scores represent a higher degree, or intensity, of purpose for that dimension.

Revision Process

The first round of piloting and analysis, utilizing a convenience sample of undergraduates from a large southeast research university (N = 2282), revealed several problems. The difficulty estimates of the scenarios were not as uniformly spaced along the construct maps as intended, some scenarios were not as difficult to score high on as expected, and some pairs of scenarios were not clearly distinguished from one another in their scale locations. In addition, upon further reflection, some of the scenarios had confusing and contradictory wording.

Extensive discussions around the conceptual foundations of the purpose construct and the operational definitions of the three dimensions led to substantial adjustments in the facet level and scenario descriptions. This process involved three cycles of revisions with re-administrations to new samples using Amazon's Mturk survey platform (see Difallah et al., 2015).

Figure 1. The Boston College Living A Life of Meaning and Purpose-B Sentence Maps

(a) Meaningfulness

	Facet (clarity) {extremely clear}		
{Name} has a	{somewhat clear}	sense of purpose in life.	
	{not at all clear}		
	Facet (effort) {a tremendous amount of effort}		
He/She places	{some effort}		into discovering a satisfying purpose in life.
	{almost no effort}		
	Facet (frequency) {almost all the time}		**Facet (horizons)** {what life is all about.}
He/She	{sometimes}	reflects upon	{is important to those around him/ her.}
	{almost never}		{make his/ her life worthwhile.}

(b) Goal Orientation

	Facet (clarity) {extremely clear}		
{Name} has a	{somewhat clear}	vision of how to make his/ her goals a reality.	
	{not at all clear}		
	Facet (effort) {a tremendous amount of effort}		
He/She places	{some effort}		towards making his/ her long term aims a reality.
	{almost no effort}		
	Facet (frequency) {almost all the time}		
He/She	{sometimes}		spends time in the day engaged in activities that bring
	{almost never}		
		Facet (horizons) {long term}	
him/her closer to his/ her		{immediate}	goals.
		{next}	

(c) Beyond-the-self

Facet (clarity)
{Name} has a {extremely clear} / {somewhat clear} / {not at all clear} desire to leave the world a better place than he/she found it.

Facet (effort)
He/she places {a tremendous amount of effort} / {some effort} / {almost no effort} into making a meaningful contribution to the broader world.

Facet (frequency)
{Name} {almost all the time} / {sometimes} / {almost never} hopes that the work that he/ she does will positively influence

Facet (horizons)
{all other people.}
{those around him/her.}
{him/herself.}

Once the scales exhibited the intended scenario ordering, we administered the instrument to a final sample of students from a different university, along with some of the validation scales used by Bronk et al. (2018; see below). Table 1 displays for each of the three dimensions the scenarios presented in their original hypothesized intensity order from the highest down to the lowest level. All results reported below come from the final sample.

Validity Instruments

Bronk et al. (2018) report Cronbach alpha's (α) of .91, .92, and .87 for meaningfulness, goal orientation, and beyond-the-self, respectively. In the present study, the three CPS subscale reliabilities are .88, .88, and .89, respectively. Bronk and colleagues also showed that the CPS is highly correlated (r =.81) with the Purpose in Life scale (*PIL,* a 7-point, 20-item scale, with response options that differ according to the item (Crumbaugh & Maholick, 1964) and, as predicted, the CPS is positively correlated with life satisfaction and negatively correlated with depression (e.g., Bigler et al., 2001; Gillham et al., 2011). To demonstrate convergent and divergent validity in the current study, we examined the correlations of the BC-LAMP-B measures with the CPS, as well as the instruments selected by Bronk et al. (2018). They report a split-half reliability for the PIL of .81. The PIL α in our study is .89.

The Life Satisfaction scale (*LS*; Diener et al., 1985) is a 5-item scale measuring a person's satisfaction with their life. Bronk et al. (2018) report an α of .87; the LS α in our study is .85. Life satisfaction is expected to be associated with high levels of purpose in life, and Bronk et al. (2018) report a correlation of .65 between the LS and CPS.

The Patient Health Questionnaire (PHQ-D; Kroenke et al., 2001) is a 9-item scale on which participants rate the frequency in which they experienced several depressive symptoms in the previous two weeks. They report an α of .89 among primary care patients; our α is .87. Depression is expected to be inversely related to purpose in life, and Bronk et al. (2018) report a -.34 correlation between the CPS and the PHQ-D.

Table 1. The Boston College Living A Life of Meaning and Purpose--B Scenarios

Meaningfulness	Goal-Orientation	Beyond-the-Self
Olivia is absolutely clear about her sense of purpose in life. She reflects on what gives life meaning all the time and puts a tremendous amount of effort into understanding what makes life worthwhile.	Bill knows how to make his long-term goals a reality and constantly exerts tremendous effort towards accomplishing them.	Jane is absolutely clear that she wants to make the world a better place for all people and she exerts a tremendous amount of effort in doing so. She's constantly engaged in activities that involve a meaningful contribution.
Sally is clear about what gives life meaning and direction. She regularly works hard on activities related to her purpose.	Jill is sure she knows how to achieve her future goals. She is actively engaged in efforts to make her future goals a reality and wishes she could be even more engaged.	Frank regularly exerts great effort in making a meaningful contribution because he knows it is important to leave the world a better place for everyone.
Joey is pretty clear about his purpose in life and exerts some effort on a regular basis to reflect on what's important to others around him.	Kammy has a pretty clear idea about how to make her next goals a reality. She occasionally places some effort towards achieving them.	Jer has a well-developed objective to leave the world a better place. Some of his activities are related to work he believes will positively influence the lives of others like him.
Kelsey is somewhat clear about her purpose in life. She occasionally puts some effort into thinking about what makes her life worthwhile.	Tyler thinks he understands how his immediate goals could become a reality. He sometimes puts some effort into making this happen.	Jim is sort of clear that he wants to make the world a better place. Although he sometimes hopes to make a meaningful contribution for people in his community, only a few of his activities are involved with positively influencing others.
Although it is not clear to Dan what his purpose in life is, he occasionally puts some effort into thinking about what would make his life more meaningful for those around him.	Jim is unclear on how to make his immediate goals a reality. Although he is rarely engaged in activities that move him closer to these goals, he does put some effort into working towards them.	Jess hopes that she will make a meaningful contribution to her friends and family. Although she is pretty clear that she desires to leave the world a better place, she places little effort in doing so.
Jack sometimes thinks about what makes his life satisfying, but he puts no real effort into understanding what gives his life meaning. So, he is not clear at all in his sense of purpose.	Ross doesn't know how to achieve his goals, and does not spend time or effort working toward them.	Angela seldom hopes that the work she does will positively influence her friends and family. She has a vague desire to leave the world a better place and is minimally engaged in activities that will make a meaningful contribution to the broader world.
Sarah has no sense of purpose in life and she does not spend any effort or time thinking about what makes her life worthwhile.		Mary does not work on anything intended to have a positive influence on anyone but herself because she is not sure it is important to leave the world a better place.

Procedure and Participants

The final sample validation data come from a subset of the 5,000 undergraduate students at a medium-sized, private, northeast research university who were randomly invited to participate. Participants were entered into a raffle to win one of twenty $25 Amazon gift cards. Of those invited, 813 students responded but we retained only the 722 who completed at least 50% of the survey. Of the 722, we removed an additional 61 (8.4%) with extremely inconsistent response patterns (discussed below). Of the final 661, 62.2% were female (n = 411), 59.9% of the respondents identified as white (n = 396), and 4.2% of the total were international students. The sample comprises an approximately uniform distribution across class years: freshmen (n = 193), sophomores (n = 154), juniors (n = 158), and seniors (n = 149). The pilot and validation studies were conducted with the approval of the university IRB board (IRB number 19.293.01E).

All surveys were administered over the Qualtrics platform. About half of the sample took all three BC-LAMP-B scales and the other group took one BC-LAMP-B scale, the CPS, either the PIL or the LS, and the PHQ-D. In the validation samples, 409 students completed the meaningfulness scale, 428 completed goal orientation, and 430 completed beyond-the-self. For our validation measures, 331 students completed the CPS, 166 completed the PIL, 158 completed the LS, and 157 completed the PHQ-D. The median time to complete a survey was 5 minutes and 3 seconds.

Results

Construct Maps and Score Interpretation

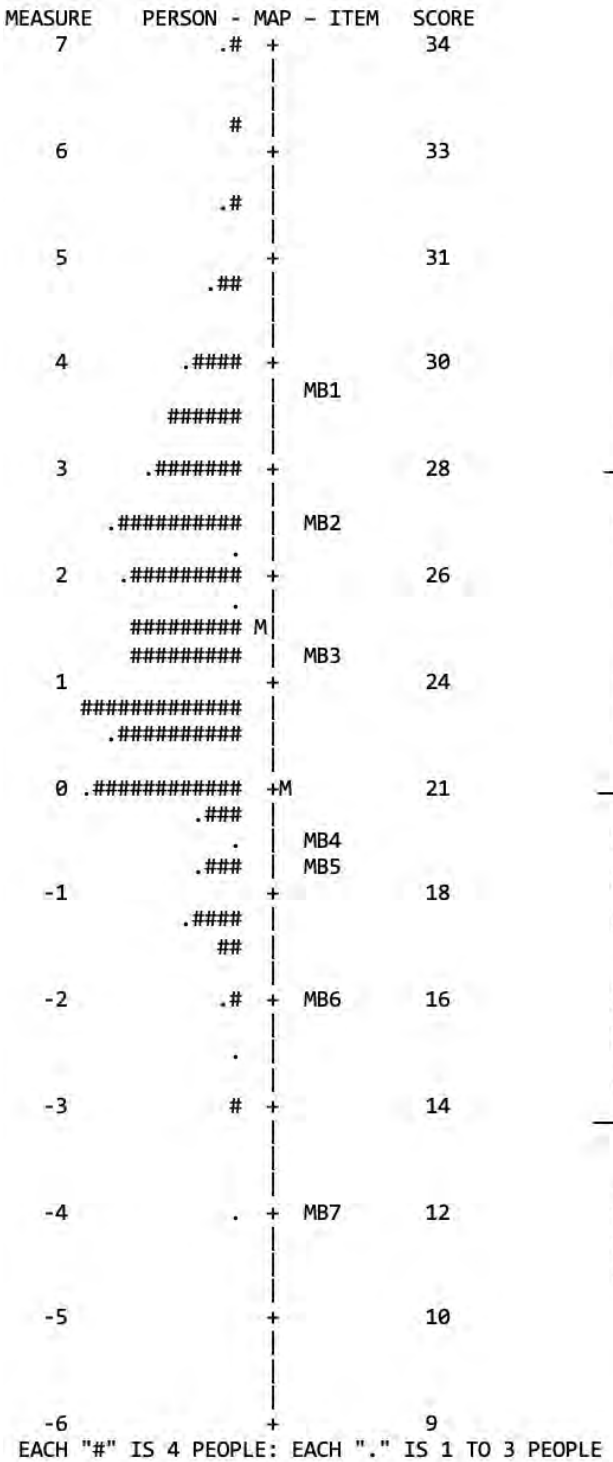
Figure 2 contains the construct maps for the meaningfulness, goal orientation, and beyond-the-self scales. On both sides of the vertical lines, "M" represents the mean location estimates. The right sides show the scenarios in order of their difficulty estimates from easiest to say "I have more direction in my life than 'X'" (bottom) to hardest (top). Each scenario has an alphanumeric code, with the numeral corresponding to its intended order in the hierarchy, e.g., "MB1" refers to meaningfulness, version B, 1st (hardest) scenario. On the left sides are the student locations indicated by "#" symbols representing 1 to 4 students at a given location. The "MEASURE" column contains the students' location estimates in logits and the corresponding raw "SCORE" for the students is reported next to them. Horizontal lines correspond to an average response of "2", "3" and "4". Overall, the scenarios and students display an excellent range from lower to higher levels on each dimension.

The locations of the scenarios defining each dimension are consistent with their intended order, as presented in Table 1. Based on the systematic development of the scenarios: (a) MB1, GB1, and BB1, were written to be the most challenging for a student to say "I have more direction in my life than 'X'" and their locations at the top of each construct map confirm that the students did, in fact, find them most difficult; (b) MB4, GB4, and BB4, written to be mid-level in their representation of the dimensions, are located midway along their respective continua; and (c) MB7, GB6, and BB7, constructed to be the least challenging to say "I have more direction in my life than 'X'", define the lower levels of the dimensions.

The text boxes in Figure 2 provide a substantively grounded and nuanced description for interpreting individual scores. For example, in Figure 2a a student with a meaningfulness score of 12 or below has little to no sense of purpose in life and spends no time or effort in figuring out what makes his/her life worthwhile. In contrast, a person with a score of 30 or more is clear on their sense of purpose, has spent considerable effort discovering a satisfying purpose, and regularly reflects upon what life is all about. With respect to the newly introduced facet of horizon, someone with a score of 12 or below identifies with a degree of meaningfulness that is limited to the self, if named at all; someone with a slightly higher score at least recognizes what is meaningful to those people around them; and those with a score of 30 or more are able to imagine a perspective on "what life is all about." This progression reflects the constructive-developmental shift from self-interest (score of 12 or below), to a social awareness (score of 13-29), to those (score of 30 or more) able to identify a perspective of ultimate meaning (Baxter Magolda, 2009; Kegan, 1994; Parks, 2011).

Figure 2. Construct Maps for the BC-LAMP-B Scales.

(a) Meaningfulness



Absolutely clear about purpose in life, puts tremendous effort into discovering a satisfying purpose in life, constantly reflects on what life is all about.

Average score of 4

Clear about what gives life purpose, works hard on activities related to that purpose, regularly reflects on what life is all about.

Pretty clear about purpose, sometimes exerts some effort to reflect on what's important to others.

Average score of 3

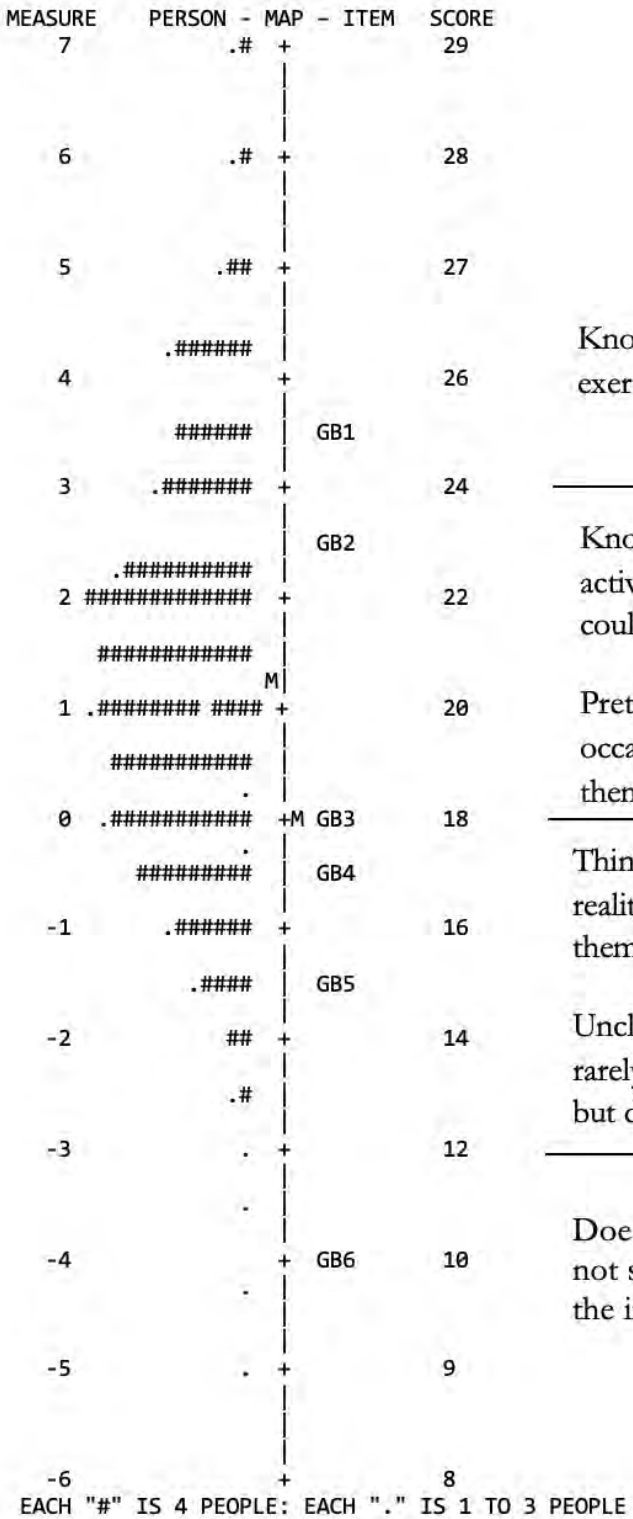
Somewhat clear about purpose in life, occasionally puts effort into thinking about what makes life important to those close at hand.

Not clear about purpose, puts little effort into understanding life's meaning, sometimes thinks about what is important to themselves and others close at hand.

Average score of 2

No sense of purpose in life, does not put in any effort in developing one, and does not consider what makes life worthwhile beyond what is personally relevant.

(b) Goal orientation



Knows how to make long term goals a reality, constantly exerts tremendous effort towards accomplishing them.

Average score of 4

Knows how to make their long-term goals a reality, is actively engaged to achieve them, and wishes they could be even more engaged.

Pretty clear how to make immediate goals a reality, occasionally places some effort towards achieving them.

Average score of 3

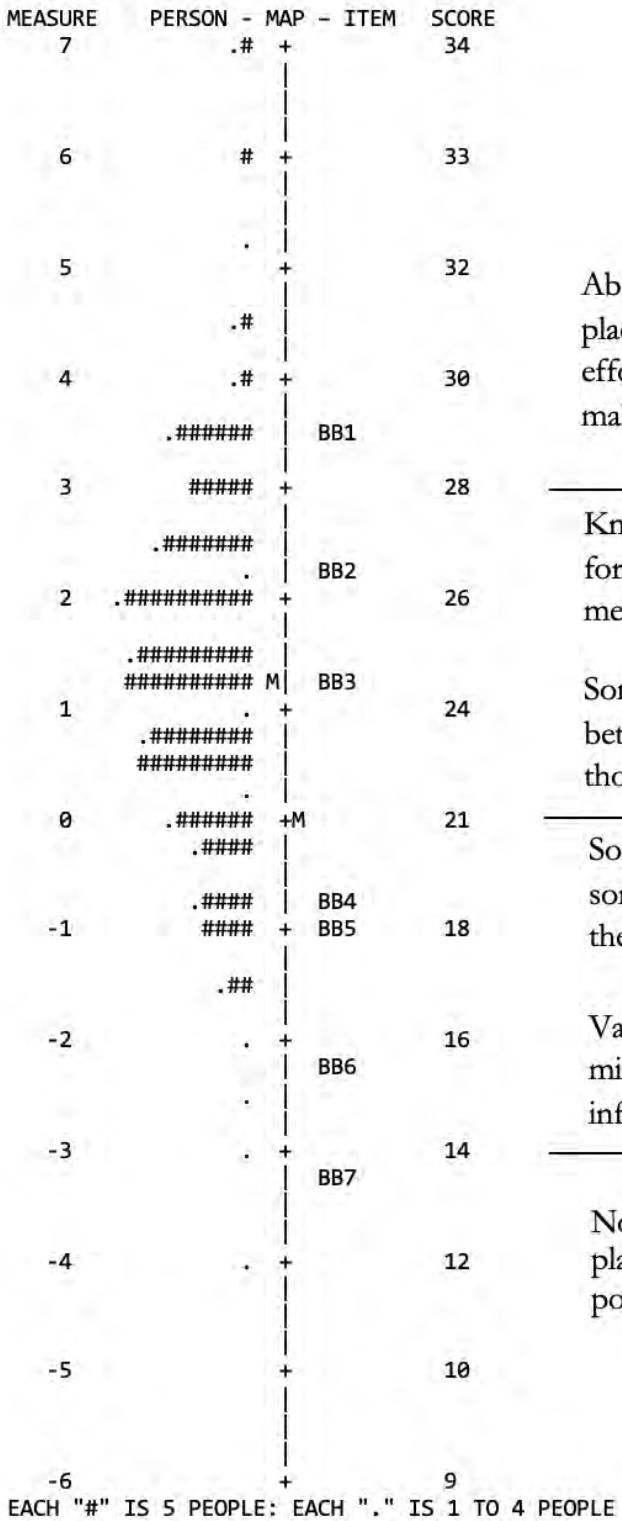
Thinks they understand how to make immediate goals a reality, and sometimes puts some effort into making them happen.

Unclear on how to make immediate goals a reality, rarely engages in activities that move closer to the goals, but does put some effort towards next steps.

Average score of 2

Does not know how to achieve their goals and does not spend time or effort working on anything beyond the immediate.

(c) Beyond-the-self



Absolutely clear in desire to make the world a better place for all people, exerts tremendous amount of effort in doing so, constantly engaged in activities that make a meaningful contribution.

Average score of 4

Knows it is important to leave the world a better place for all people, exerts great effort in making a meaningful contribution.

Somewhat clear about wanting to leave the world a better place, some effort exerted to positively influence those around them.

Average score of 3

Some interest in making the world a better place, sometimes makes some effort to positively influence their own community

Vague desire to leave the world a better place, places minimal effort in doing so, seldom hopes to positively influence friends and family.

Average score of 2

Not sure it is important to leave the world a better place, does not work on anything intended to have a positive influence on anyone but themselves.

In Figure 2b a student with a score of 10 or less on goal orientation does not have a clear vision of how to achieve their immediate goals and spends little time or effort on activities that bring about their goals. A person with a score of 24 or more, however, has an extremely clear vision of how to make their goals a reality and puts in tremendous effort engaging in activities to achieve goals that are long term. As they pertain to horizon, the scores reflect a progression from an inability to consider beyond the immediate (score of 10 or less), to some awareness of consequence beyond the immediate (score of 11 – 23), to finally (score of 24 or more) an awareness of consequence that is quite distant (Kegan, 1994).

Finally, in Figure 2c, a student low on beyond-the-self with a score of 14 or less is not clear why they would want to leave the world a better place, does not place any effort into making a meaningful contribution, and rarely hopes their work has a positive influence on anyone beyond themselves. A person with a score of 28 or more, by contrast, has a clear desire to leave the world a better place, places tremendous effort in making a meaningful contribution, and frequently hopes their work will have a positive influence on all people. As it pertains to horizon, this progression reflects the self-interest normal to later childhood and early adolescence (score of 14 or below), to the person's ability to see and account for their impact on the surrounding community in their purpose (score of 15 to 27), and finally (score of 28 or more) the individual's ability to recognize their impact on the wider world (Kegan, 1994).

Fit Statistics

In addition to the construct maps, goodness-of-fit statistics evaluate how well the observed responses from the participants fit the responses predicted under the Rasch model. Residuals are the differences between the observed and predicted responses. Mean squared residual statistics for "infit" (a variance weighted statistic) and "outfit" (an unweighted statistic) above 1.4 indicate potential model misfit (Linacre, 2019b). The goodness-of-fit analysis for the three scales from the initial pilot, three revisions, and final administration data revealed that meaningfulness MB7, goal orientation GB6, and beyond-the-self BB7 all demonstrated a relatively high frequency of unexpected responses. These three scenarios are the easiest to respond with "I have more direction in my life than 'X'." The problem, however, was that some high-scoring students, expected to respond with "I have more direction in my life than 'X'" on these three easy scenarios, responded unexpectedly low to them.

A detailed analysis of the observed and expected responses on each scale revealed two disturbing patterns: (a) there were some students with multiple responses nearly opposite what was expected given their overall score, and (b) there were some students who simply marked the middle response category for all scenarios. Pattern (a) is illogical—a student cannot legitimately respond with "I have more direction in my life than X" to a high-level scenario while at the same time responding with "X has more direction in life than me" to a low-level scenario. Pattern (b) is the typical response set known to occur when participants simply pick one option and then move quickly through the items without paying any attention to them.

These types of construct irrelevant variance are not unusual in survey practice (Kyllonen & Bertling, 2013). One solution for ensuring the greatest generalizability of a new instrument is to remove, or trim, influential aberrant observations from the analysis (Belsley et al., 1980; Winer, 1971). Extremely high misfitting student data records (infit and outfit statistics greater than 2.5) were identified and compared statistically on a number of background characteristics with all other students. No statistically significant differences between these two groups were found with regard to gender, year in college, race (white/non-white), length of time to complete the survey, whether or not they completed it, or whether they took all three BC LAMP-B scales or just one in conjunction with a validation scale.

Those misfitting data records were removed from each scale's data set: meaningfulness (n = 25, 5.4% of the sample), goal orientation (n = 25, 5.2%), and beyond-the-self (n = 24, 4.9%). Interestingly, across these three groups of deleted records, there were five students who provided misfitting response patterns on two scales and four who misfit on all three scales.

All analyses were rerun (including the validity instrument reliabilities and inter-scale correlations reported below) and those results are what we present in this paper. The final scenario estimates and their goodness-of-fit statistics are reported in Table 2. Not only did the degree of misfit decrease but the "student separation" statistic (separation > 2.0 indicates increasingly greater, and finer differentiation, in the

spread of the student location estimates; Wright & Masters, 1983) for meaningfulness increased from 2.14 to 2.57, for goal orientation the improvement was from 2.0 to 2.25, and for beyond-the-self, it was from 2.03 to 2.33).

Internal Consistency Reliability

The Cronbach αs for the original full sample were .83, .81, and .81 for meaningfulness, goal orientation, and beyond-the-self, respectively. For the trimmed sample, the corresponding alphas are .86, .84, and .84. Since 1-α is an estimate of the error variance within one's scale scores, the removal of the extremely misfitting student records reduced the measurement error for each scale by 21%, 19%, and 19%, respectively.

Scale Correlations

The inter-scale Pearson correlations for the sample are reported in Table 3. Meaningfulness and goal orientation correlated at $r = .75$, meaningfulness and beyond-the-self at $r = .70$, and goal orientation and beyond-the-self at $r = .64$ (all at $p < .001$).[2] These moderate correlations, along with the corresponding scatterplots (Appendix B), demonstrate that a number of students see themselves at substantially different levels on the three dimensions. Case #737, for example, scored high on meaningfulness, moderately on goal orientation, and very low on beyond-the-self. This student (a 3rd-year undergraduate Hispanic male) may spend time and effort thinking expansively on what life is all about and has set goals a few years out (like a college degree or first job), but that vision apparently does not include consideration of how the student's life impacts anyone but the student. Another example is Case #819 who had a medium score on meaningfulness but very low scores on goal orientation and beyond-the-self. This student (a 2nd-year undergraduate White male) may have a clear sense of what is important to those around him but appears not to have spent any time or effort thinking beyond immediate goals or impacting anyone beyond himself.

As these two examples well illustrate, using the three scales in combination provides a more nuanced picture of an individual across the three dimensions than is offered by a single summary score. This is a reasonable strategy when trying to understand a student's profile on the three dimensions of purpose because even with the highest correlation of .75 (meaningfulness and goal orientation), only 56.25% of the variance in those pairs of scores is shared.

Convergent and Divergent Evidence

Table 3 also presents the three BC-LAMP-B scale correlations with the validity measures: the CPS, PIL, LS, and PHQ-D. The correlations between all three scales and the CPS, PIL, and LS validation scales are positive and statistically significant. Although the correlations between all three scales and the PHQ-D (depression) were negative, as expected, they were only significant for goal orientation. Overall, the correlations of the BC-LAMP-B with the validity scales are consistent with those reported by Bronk et al. (2018) for the CPS; the differences may be attributed in part to measurement error introduced by the novel item format, sampling error, and demographic differences between the samples.

Known Groups Evidence

Construct validity is typically provided when an instrument can differentiate between those known to have a particular trait and those who do not have it. For example, in the present case, it is reasonable to expect that over the four years of college students will gain in maturity and that this development should be reflected by higher scores related to purpose (Baxter Magolda, 2008; Kegan, 1994; Parks, 2011). Indeed, the cross-sectional analysis of the four-year undergraduate responses reveals that the average responses for meaningfulness and beyond-the-self increased each year (Appendix C). Although goal orientation showed a dip from year-1 to year-2, there was a sharp increase in the year-3 and year-4 averages. We believe that the slight dip in goal orientation makes sense in that most students enter college with some sense of direction (e.g., choice of major), but that frequently is overturned or interrupted by the experience of college itself. In fact, the increase in meaningfulness and engagement with people beyond themselves may contribute to both disruption and reorientation.

[2] The respective correlations corrected for attenuation are .88, .82, and .76.

Table 2. Fit statistics for the meaningfulness, goal orientation, and beyond-the-self scenarios.

(a) Meaningfulness

Item	Total count	Measure	Model *SE*	Infit MNSQ	Outfit MNSQ
MB1	436	5.28	0.09	1.08	1.03
MB2	437	3.56	0.09	0.99	0.94
MB3	437	1.98	0.10	0.96	0.93
MB4	438	-0.75	0.09	0.80	0.81
MB5	436	-1.18	0.09	0.81	0.86
MB6	436	-2.89	0.10	1.03	1.59
MB7	437	-6.00	0.15	1.41	2.43
Mean	436.7	0.00	0.10	1.01	1.23
SD	0.70	3.60	0.02	0.19	0.55

Note: *SE* = Standard error, MNSQ = Mean square, *SD* = Standard deviation

(b) Goal orientation

Item	Total count	Measure	Model *SE*	Infit MNSQ	Outfit MNSQ
GB1	458	4.74	0.09	1.11	1.06
GB2	457	3.49	0.09	1.06	1.03
GB3	458	-0.15	0.09	0.90	1.10
GB4	456	-0.74	0.09	0.85	0.84
GB5	457	-2.01	0.09	0.80	1.10
GB6	456	-5.34	0.13	1.19	2.25
Mean	457.0	0.00	0.09	0.98	1.23
SD	0.80	3.36	0.01	0.15	0.46

(c) Beyond-the-self

Item	Total count	Measure	Model *SE*	Infit MNSQ	Outfit MNSQ
BB1	454	4.55	0.09	1.03	1.10
BB2	455	2.94	0.09	0.92	0.94
BB3	458	1.47	0.09	0.84	0.83
BB4	455	-0.87	0.09	0.94	1.01
BB5	456	-1.40	0.08	0.76	1.37
BB6	455	-2.80	0.09	1.19	1.10
BB7	457	-4.16	0.10	1.18	1.42
Mean	455.7	0.00	0.09	0.98	1.11
SD	1.30	2.94	0.01	0.15	0.20

Table 3. Correlations table between the BC-LAMP-B scales and the validity measures

	MB	GB	BB	CPS	PIL	LS
MB						
GB	.77 (< .001)					
BB	.68 (< .001)	.63 (< .001)				
CPS	.76 (< .001)	.73 (< .001)	.70 (< .001)			
PIL	.76 (< .001)	.50 (< .001)	.52 (< .001)	.67 (< .001)		
LS	.36 (.008)	.49 (< .001)	.32 (.023)	.48 (< .001)	-	
PHQ-D	-.16 (.251)	-.41 (.002)	-.19 (.192)	-.25 (.001)	-	-.47 (< .001)

Note: *p*-values are presented in parentheses. M = Meaningfulness, G = Goal Orientation, B = Beyond-the-Self, CPS = Claremont Purpose Scale, PIL = Purpose in Life Test, LS = Life Satisfaction Scale, PHQ-D = Patient Health Questionnaire – Depression. The PIL was never administered with the LS and PHQ-D.

Furthermore, there was a statistically significant positive linear trend across time for meaningfulness ($F(1, 431) = 7.29$, $p = .007$) and the year-1 and year-4 averages differed significantly (Dunnett's *t*-test with year-1 as the target, $p = .027$). For goal orientation, the test of linear trend was marginally nonsignificant ($F(1, 431) = 3.11$, $p = .078$) as was the Dunnett *t*-test between year-1 and year-4 ($p = .096$). For beyond-the-self the linear trend was significant ($F(1, 431) = 12.15$, $p = .001$) as was the Dunnett *t*-test of the year-1 and year-4 average responses ($p = .001$).

In keeping with our objective to provide substantive score interpretations, the increase in mean scores from year-1 to year-4 for meaningfulness may be understood as a shift from "pretty clear about purpose in life while exerting some effort on a regular basis to reflect on what's important to others around them" to "clear about what gives life meaning and direction while regularly working hard on activities related to her purpose." For goal orientation, the increase is from "pretty clear about how to make her next goals a reality while occasionally placing some effort towards achieving them" to "sure she knows how to achieve her future goals, is actively engaged in efforts to make future goals a reality, and wishes she could be even more engaged." Finally, for beyond-the-self, the change reflects movement from "has a well-developed objective to leave the world a better place and some activities are related to work he believes will positively influence the lives of others like him" to "regularly exerts great effort in making a meaningful contribution because it is important to leave the world a better place for everyone."

Discussion

A complex construct such as purpose in life is difficult to define and measure. This fact may explain the many frameworks proposed for studying purpose. These frameworks lead to competing definitions and assessment instruments with varying degrees of psychometric quality and utility. For example, the definition of purpose proposed by Damon et al., (2003) and Damon (2008) led to the *Claremont Purpose Scale* (Bronk et al., 2018). While the CPS demonstrated high internal consistency reliability and is helpful in conceptualizing purpose, we found it too limiting for our objectives. In this article we attended to these limitations by (a) demonstrating the utility of Rasch/Guttman Scenario measurement methodology to construct sets of scenarios representing plausible progressions of lived experiences, (b) providing rich and actionable interpretations of scores along those progressions, and (c) enhancing the construct's definition by drawing on the constructive-developmental theory of Kegan (1994) to introduce horizon as a facet for evaluating whether an individual's purpose is near or far, great or small.

Employing the RGS methodology, we showed how an intricate item development process operating on a set of discrete facets, each of which may be systematically altered in their level of intensity, can create meaningful progressions of lived experiences along multiple dimensions that, in turn, can be combined to represent a complex construct - in our case, purpose. Although the interpretation of a scenario may differ somewhat for each person because of the unpredictable facet interactions perceived by an individual, the authentic mosaics of experience defining a construct can be shaped and modeled into plausible depictions of human development.

Furthermore, the utility of constructing a separate scale for each of the three dimensions, rather than a single summary score, was demonstrated through analyses that revealed numerous students with substantially different scores across the three dimensions; that is, differences in their perceptions of their lived experiences. The BC-LAMP-B uncovers this divergence, which then invites a more nuanced interpretation of the student. Consequently, proactive efforts to facilitate enriched purpose would entail strategies appropriate to the revealed profile. Such implications are discussed in the following subsection.

Finally, our results support the convergent and discriminant validity of the BC-LAMP-B scales with respect to other self-report measures of purpose (i.e., the CPS and PIL scales) and measures of life satisfaction and depression. Known group validity based on undergraduate class analyses for each of the dimensions was also supported.

Implications

By virtue of being able to link score profiles to rich qualitative descriptions of self-perceptions, the BC-LAMP-B provides a powerful approach to interpreting both individual and aggregate score differences. This point is especially salient. Once an individual's location on a dimension's continuum has been determined, it is possible to consider what might facilitate further progress. For example, interventions supporting further development along the goal orientation dimension might include career counseling, academic advisement, and mentorship programs. Interventions in support of beyond-the-self might include service learning projects, co-curricular student organizations, independent or counselor-guided structured reflection activities. With regard to meaningfulness, coursework in the liberal arts, retreats, or facilitator-led psychoeducational or therapeutic groups focused on different objectives might prove efficacious. Finally, the BC-LAMP-B offers data at the student level, as well as at the aggregate level, making it possible to determine what kinds of interventions are needed in different years, whether and how they are effective.

Limitations and Future Work

The instrument development work documented here is fertile ground for future research efforts. Although we initially tested our pilot instrument using a variety of samples from different settings, our final sample comprised undergraduate students all attending a single university, so it was quite demographically homogeneous. Further studies should explore whether the instrument's psychometric properties hold across different types of institutions (e.g., community colleges; public colleges and universities; private, secular colleges and universities; faith-based colleges, etc.) and student populations (e.g., students of different racial, economic, social, and religious backgrounds, etc.).

Second, in this scale, we were working within some of the conceptual parameters of CPS, which investigated purpose without explicit consideration of context or student diversity. In future work, it would be important to explore the extent of measurement invariance in the meaning of the lived experience scenarios across students of different cultural backgrounds and languages (Ashlee et al., 2018). It would also be helpful to discover if and how context shapes purpose.

Third, although our long-term plan includes developing a tool to capture within-person change over time (naturally occurring or in response to intervention), in the current study we only measured status and have not yet empirically tested sensitivity to change. To evaluate whether interventions have contributed to changes in students' sense of purpose will require comparative studies specifically designed to answer that question. These are important tasks for future work.

In building and continuing to modify the BC-LAMP portfolio, our goal is to create tools that (a) measure the presence, and degree, of meaning and purpose in the lives of persons 18-30, (b) are sufficiently easy to administer, (c) facilitate rich

authentic score interpretations, and (d) are useful for mapping developmental change over time. With the incorporation of the concept of horizon in the measurement of purpose and refinements in the RGS methodological development of lived experience scenario-based scales, our results demonstrate that the BC-LAMP-B scales have the potential to help colleges measure the presence and development of purpose in students.

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*, 561–573.

Ashlee, A. A., Ashlee, K. C., Okello, W., & Karikari, S. N. (2017-2018). Critical race reflections on self-authorship and the Learning Partnerships Model. *Journal of Student Affairs,* Vol XXVII, 63-74.

Battista, J., & Almond, R. (1973). The development of meaning in life. *Psychiatry*, *36*(4), 409-427.

Baxter Magolda, M. (2008). Three elements of self-authorship. *Journal of College Student Development, 47*(4), 269-284.

Baxter Magolda, M. (2009). The activity of meaning making: A holistic perspective on college student development. *Journal of College Student Development, 50*(6), 621-639.

Baxter Magolda, M., King, P., Taylor, K., & Wakefield, K. (2012). Decreasing authority dependence during the first year of college. *Journal of College Student Development, 53*(3), 418-435.

Belsley, D. A., Kuh, E. & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York, NY: Wiley & Sons.

Bigler, M., Neimeyer, G. J., & Brown, E. (2001). The divided self revisited: Effects of self-concept clarity and self-concept differentiation on psychological adjustment. *Journal of Social and Clinical Psychology*, *20*(3), 396-415.

Borg, I., & Shye, S. (1995). Facet theory. Form and content, Vol. 5. *Advanced quantitative techniques in the social sciences series. Thousand Oaks, CA/London/New Delhi: SAGE*.

Bronk, K. C., Riches, B. R., & Mangan, S. A. (2018). Claremont Purpose Scale: A measure that assesses the three dimensions of purpose among adolescents. *Research in Human Development*, *15*, 1-17.

Canter, D. (2019). Qualitative structural theory: A basis for decision-making. *International Studies of Management & Organization*, *49*, 265-282.

Crumbaugh, J. C., & Maholick, L. T. (1964). An experimental study in existentialism: The psychometric approach to Frankl's concept of noogenic neurosis. *Journal of Clinical Psychology*, *20*, 200-207.

Damon, W. (2008). *The Path to Purpose: How Young People Find Their Calling in Life*. New York: Free Press.

Damon, W., Menon, J., & Bronk, K. C. (2003). The development of purpose during adolescence. *Applied Developmental Science*, *7*, 119-128.

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*, 71-75.

Difallah, D. E., Catasta, M., Demartini, G., Ipeirotis, P. G., & Cudré-Mauroux, P. (2015, May). The dynamics of micro-task crowdsourcing: The case of amazon mturk. In *Proceedings of the 24th international conference on world wide web* (pp. 238-247). International World Wide Web Conferences Steering Committee.

Emmons, R. A. (2005). Striving for the sacred: Personal goals, life meaning, and religion. *Journal of Social Issues*, *61*, 731-745.

Gadamer, H. G. (2004). *Truth and Method*. New York: Continuum Publishing.

Gillham, J., Adams-Deutsch, Z., Werner, J., Reivich, K., Coulter-Heindl, V., Linkins, M., & Contero, A. (2011). Character strengths predict subjective well-being during adolescence. *The Journal of Positive Psychology*, *6*(1), 31-44.

Glanzer, P. L., Hill, J. P. & Johnson, B.R. (2017). *The Quest for Purpose: The Collegiate Search for a Meaningful Life*. Albany: State University of New York Press.

Guttman, L. (1959). Introduction to facet design and analysis. Paper presented at the proceedings of the Fifteenth International Congress of Psychology, Amsterdam, Netherlands.

Guttman, R., & Greenbaum, C. W. (1998). Facet theory: Its development and current status. *European Psychologist*, *3*, 13-36.

Hackett, P. M. (2014). *Facet Theory and the Mapping Sentence: Evolving Philosophy, Use and Application*. Basingstoke: Palgrave Macmillan.

Heintzelman, S. J., & King, L. A. (2014). Life is pretty meaningful. *American Psychologist*, *69*(6), 561-574.

Kegan, R. (1982). *The Evolving Self: Problem and Process in Human Development.* Cambridge, MA: Harvard University Press.

Kegan, R. (1994). *In Over Our Heads: The Mental Demands of Modern Life.* Cambridge, MA: Harvard University Press.

King, P.M. (2009). Principles of development and developmental change underlying theories of cognitive and moral development. *Journal of College Student Development, 50*(6), 597-620.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606-613.

Kyllonen, P. C., & Bertling, J. P. (2004). Innovative questionnaire assessment methods to increase cross-country comparability. In (L. Rutkowski, M. von Davier, D. Rutkowski Eds.) Handbook of International Large-scale Assessment. CRC Press.

Leontiev, D. A. (2006). Meaningful living and the worlds of art. In A. Della Fave (Ed.), *Dimensions of well-being research and intervention* (pp. 529–537). Milano: Franco Angeli.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, *3*(1), 85-106.

Linacre, J. M. (2019a). WINSTEPS, V 4.4.5. [Computer program]. Beaverton, OR: Winsteps.com.

Linacre, J. M. (2019b). Fit diagnosis: infit outfit mean-square standardized. Retrieved from https://www.winsteps.com/winman/misfitdiagnosis.htm

Ludlow, L. H., Baez-Cruz, M., Chang, W.-C., & Reynolds, K. (2020). Rasch/Guttman Scenario Scales: A Methodological Framework. *Journal of Applied Measurement*, *21*(4), 361-378.

Ludlow, L. H., & Haley, S. M. (1995). Rasch model logits: Interpretation, use, and transformation. *Educational and Psychological Measurement, 55,* 967-975.

Ludlow, L. H., Matz-Costa, C., Johnson, C., Brown, M., Besen, E., & James, J. B. (2014). Measuring engagement in later life activities: Rasch-based scenario scales for work, caregiving, informal helping, and volunteering. *Measurement and Evaluation in Counseling and Development*, *47*, 127-149.

Ludlow, L. H., Reynolds, K., Baez-Cruz, M., & Chang, W.-C. (2021). Enhancing the interpretation of scores through Rasch-based scenario-style items. In (Harbaugh AG: Ed). *Basic Elements of Survey Research in Education: Addressing the Problems Your Advisor Never Told You About.* https://www.infoagepub.com/products/Basic-Elements-of-Survey-Research-in-Education.

Martela, F., & Steger, M. F. (2016). The three meanings of meaning in life: Distinguishing coherence, purpose, and significance. *The Journal of Positive Psychology*, *11*, 531-545.

McKnight, P. E., & Kashdan, T. B. (2009). Purpose in life as a system that creates and sustains health and well-being: An integrative, testable theory. *Review of General Psychology*, *13*, 242-251.

O'Keefe, T. A. (2018). *Navigating Toward Adulthood.* New York: Paulist Press.

Parks, S. D. (2011). *Big Questions, Worthy Dreams: Mentoring Emerging Adults in Their Search for Meaning, Purpose, and Faith*. San Francisco: Jossey-Bass.

Piaget, J. (1952). *The Origins of Intelligence in Children.* New York: International Universities Press. Originally published in 1936.

Rasch, G. (1980). Probabilistic models for some intelligence and attainment tests. Chicago, IL: University of Chicago Press. (Original work published 1960).

Ryff, C. D., & Singer, B. (1996). Psychological well-being: Meaning, measurement, and implications for psychotherapy research. *Psychotherapy and Psychosomatics*, *65*, 14-23.

Taylor, C. (1989). *Sources of the Self: The Making of the Modern Identity*. Cambridge MA: Harvard University Press.

Winer, B. J. (1971). *Statistical principles in experimental design*. New York, NY: McGraw-Hill.

Wright, B. D., & Masters, G. N. (1982). *Rating Scale Analysis*. Chicago, IL: MESA Press.

Citation:

Ludlow, L.H., O'Keefe, T., Braun, H., Anghel, E., Szendey, O., Matz, C., & Howell, B. (2022). An Enhancement to the Theory and Measurement of Purpose. *Practical Assessment, Research & Evaluation*, 27(4). Available online: https://scholarworks.umass.edu/pare/vol27/iss1/4/

Corresponding Author:

Larry H. Ludlow
Lynch School of Education
Boston College
Chestnut Hill, MA, USA

Email: larry.ludlow [at] bc.edu

Appendix A. Examples of Horizons Language

	Meaningfulness	Goal Orientation	Beyond the Self
Low	…serves my immediate purposes	…it's what I am working on now	…it serves my needs
Medium	…what is important to those around me	…it seems to be the next thing to work on	...it serves the needs of people I like
High	…what I think life is all about	…a long-term goal worth working towards	…it serves the needs of all people

Appendix B. Scatterplots of BC-LAMP-B scores

Figure B-1. Scatterplot for Meaningfulness (MB) and Goal Orientation (GB)

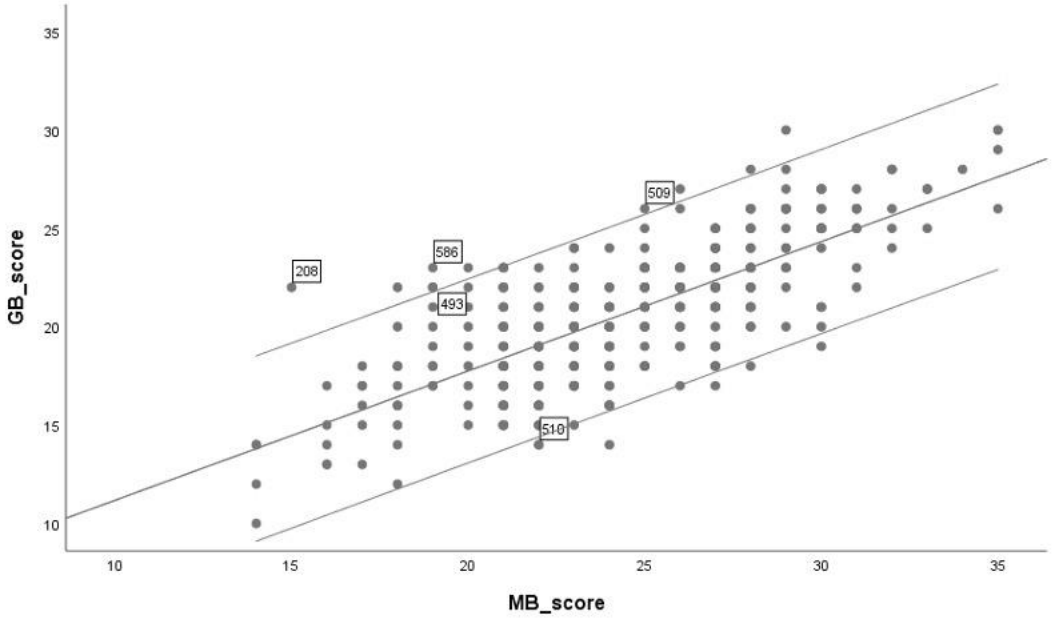


Figure B-2. Scatterplot for Meaningfulness (MB) and Beyond-the-Self (BB)

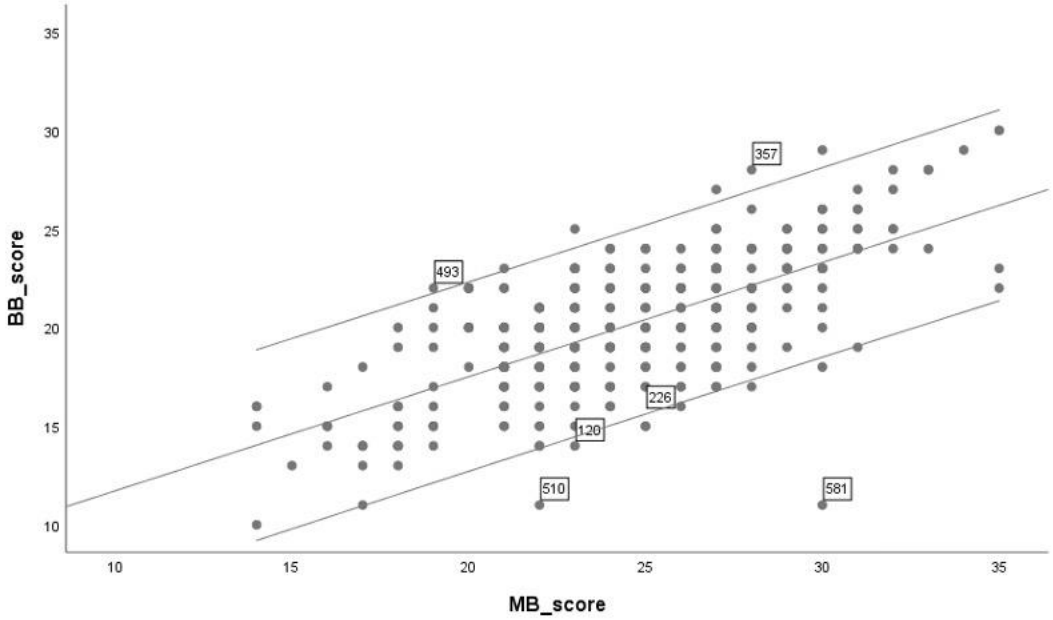
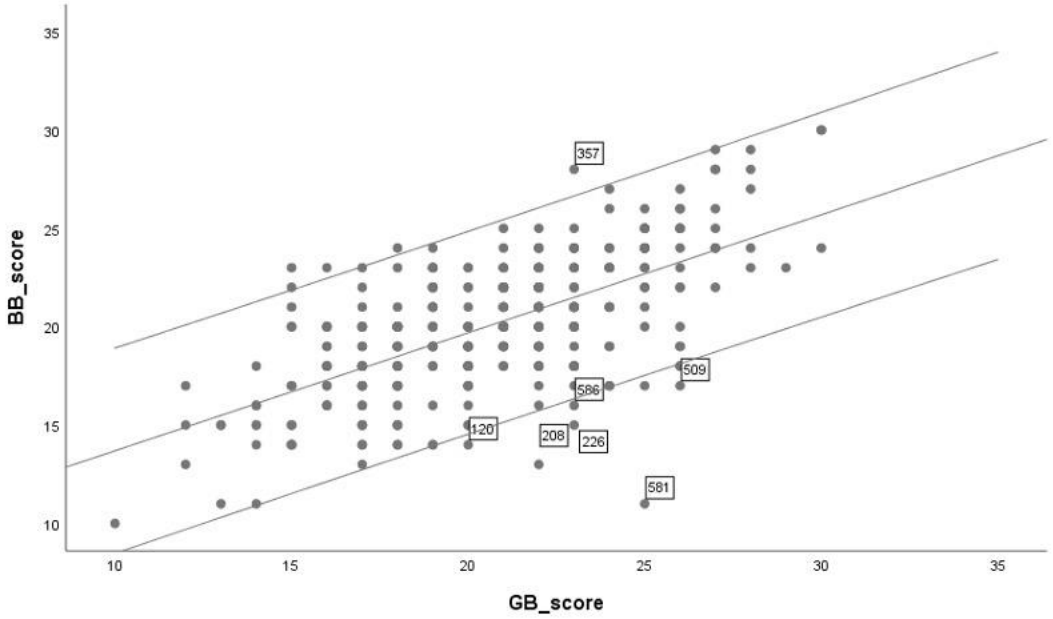


Figure B-3. Goal Orientation (GB) and Beyond-the-Self (BB)



Appendix C. BC-LAMP-B scale scores by undergraduate year

Figure C-1.