Measuring Meaning-Making Among College Students: The Living a Life of Meaning and Purpose—C Portfolio

Authors:

Larry H. Ludlow, Ph.D.
Boston College

Henry I. Braun, Ph.D.
Boston College

Theresa A. O'Keefe, Ph.D.
Boston College

Christina Matz, M.S.W., Ph.D.
Boston College

Lauren Warner, Ph.D. Candidate
Boston College

ABSTRACT

This article describes the development and initial validation of a scenario-based instrument to assess meaning-making in college students, grounded in constructive developmental theory (O'Keefe et al., 2025). Using Rasch/Guttman Scenario (RGS) scale methodology (Ludlow et al., 2020a), we constructed two parallel instruments employing school and family contexts. Each instrument comprised an ordered sequence of scenarios designed to increase in complexity across four facets of the construct: cognitive ability, relational awareness, conflict resolution, and sense of responsibility. Results from a diverse undergraduate student sample supported the hypothesized scenario hierarchy and proper functioning of the response categories. The instruments also provide substantive interpretations of respondent score locations that can inform longitudinal tracking of meaning-making growth in emerging adults. These findings offer preliminary support for using the RGS framework to model meaning-making as a structured, developmental construct, offering a theory-aligned alternative to qualitative interviews.

Correspondence E-mail: ludlow@bc.edu

Keywords: meaning-making, constructive developmental theory, Rasch/Guttman scenario scales

The college years represent a critical period of developmental transitions as students move from adolescence into early adulthood. While postsecondary education is often evaluated in relation to intellectual growth and career readiness, there is a growing recognition that institutions of higher learning also bear a responsibility to support the formation of students as whole persons capable of leading lives of meaning and purpose (Dalton & Crosby, 2010; Glanzer et al., 2017). For educators and administrators who consider formation a worthwhile educational outcome, the need for psychometrically sound tools to measure such development becomes both practical and urgent.

Yet the measurement of formative outcomes remains inherently challenging. Constructs like meaning-making are conceptually rich, developmentally dynamic, and not easily captured by traditional survey instruments. In response to this measurement problem, we developed the Living a Life of Meaning and Purpose–C (LAMP-C), a self-report instrument designed to assess meaning-making capacity among traditional-age college students. LAMP-C comprises two context-specific instruments—school and family—each composed of developmentally ordered scenarios.

The design of LAMP-C is grounded in two complementary frameworks. The first is constructive developmental theory (CDT), which provides the conceptual foundation for the meaning-making construct. The second is the Rasch/Guttman Scenario (RGS) scale methodology, which guides scenario development, response scaling, and item calibration (Ludlow et al., 2014, 2020b).

This paper outlines our theoretical and methodological rationale for the development of LAMP-C. We begin by summarizing the relevant scholarship on meaning-making and offer a reframing of CDT for the purpose of construct operationalization. We then describe the application of RGS methodology to instrument construction and conclude with a discussion of the psychometric characteristics of the school and family scales, next steps, and implications for future assessment research.

Constructive development theory (CDT)

The literature on meaning-making, cognitive development, self-authorship, and leadership within educational contexts, particularly among college students and emerging adults, presents a rich tapestry of interconnected ideas that deepen our understanding of individual development. Building upon the cognitive development work of Jean Piaget (1952) theorists such as Kegan (1982, 1994), Parks (2000), Baxter Magolda (2001, 2009), King (2009), McAdams (2013) and Helsing & Howell (2014) have explored how individuals construct and evolve their understanding of themselves and their place in the world. In the aggregate, this work is called constructive developmental theory (CDT).

Briefly, CDT posits that human development is a process of progressively more complex meaning-making with individuals potentially moving from childhood to mature adulthood toward increasingly complex capacities for understanding themselves and the world. Enhancement of meaning-making impacts and is impacted by cognitive capacity, identity, and relationships. Longitudinal research that follows traditional-age college students into mature adulthood has been crucial for understanding what contributes to young adults achieving self-authorship, a key developmental milestone. “Self-authorship” is understood as the capacity to internalize and ‘own’ one’s values, beliefs, and identities, while acknowledging connections with and responsibility to others (Kegan, 1994; Baxter Magolda, 2001).

Studies of cognitive and moral development within higher education settings suggest that the cognitive and moral reasoning capacities of students evolve in response to both educational and social experiences. In particular, narrative identity formation illustrates how individuals continually revise their life stories as they encounter new experiences and challenges, and highlights how critical incidents and supportive relationships can facilitate or hinder the journey toward self-authorship (King, 2009; McAdams, 2013; Pizzolato, 2003). The extension of CDT into the realm of leadership illustrates how the principles of CDT can inform leadership training and assessment (Helsing & Howell, 2014). Finally, aligning educational experiences with the developmental capacities of emerging adult learners exemplifies how theoretical insights can be translated into practical strategies for enhancing a broad range of educational outcomes (Steward & Wolodko, 2016). In sum, the literature suggests that meaning-making, self-authorship, and cognitive development are deeply interconnected processes that are central to personal and educational growth; CDT provides a robust framework for understanding these processes.

In nearly each case, the research cited above involved some form of qualitative interviews, such as the Subject-Object Interview developed by Kegan and colleagues (Leahy et al., 2011). Qualitative interviews, while potentially highly informative, are problematic for large-scale administration because they are time consuming, both in administration and the interpretation of data. Accordingly, they are usually employed in small samples. A few self-report questionnaires and scales have been created to assess self-authorship as it pertains to career readiness (Creamer et al., 2010; Fallar et al., 2019). However, these measures have limitations, particularly in their ability to capture the dynamic and developmental nature of meaning-making across contexts. Our goal was to create a practical, quantitative assessment of meaning-making, drawing on CDT to inform instrument development.

Facets of CDT

As explained in greater detail below, the Rasch/Guttman Scenario psychometric approach operationalizes complex constructs, like meaning-making, by breaking the construct down into its constituent characteristics, or facets. These facets are then combined in a specific manner to construct an ordered sequence of scenarios that describe responses to situations occurring in common contexts in which college students must exercise meaning-making. Within a context, the scenarios are incrementally modified to reflect gradations of the facets, representing different responses to the situations that reflect different capacities of meaning-making. We offer warrants for both the facets and their gradations by drawing from our analysis of the theories of Kegan, Parks, and Baxter Magolda. Below, each facet is described and associated with at least one theorist's framework.

With regard to facet gradations, all of the theorists cited above identified developmental steps in which individuals exhibit qualitatively different capacities for interpreting and responding to their worlds. Kegan (1994) identifies five possible "orders of consciousness" across the lifespan (p. 314-315). Parks (2000) identifies four "forms" spanning from adolescence to mature adulthood (p. 90). Baxter Magolda (2001) names four "phases," starting in college and extending into mature adulthood (p. 40). With each shift of order, form or phase, a person does not forget or reject what was previously known,

but subsumes it and organizes it in a broader framework of meaning. In our focus on traditional–age college students, we propose five gradations, which we call "positions" (O'Keefe et al., 2025).

Across the developmental span between late adolescence and early adulthood, traditional college students are actively shifting how they make sense of their lives and relationships. The theories of Kegan, Parks, and Baxter Magolda each describe this transition in slightly different terms, but the general pattern is consistent: students move from externally guided, socially defined ways of knowing—what Kegan calls the Third Order and Baxter Magolda describes as Following External Formulas—toward internally defined, self-authored frameworks such as Kegan's Fourth Order or Baxter Magolda's Self-Authorship.

The LAMP-C model uses five positions to capture this developmental progression. Positions 2 and 3 represent earlier, socially driven forms of meaning-making, while Positions 4 and 5 capture the emergence and consolidation of self-authorship. Position 1 anchors the continuum, reflecting more concrete, categorical ways of thinking often observed at entry to college. This structure allows us to see not just where students begin and where they might ultimately arrive, but also the shifts that occur in between.

There are two reasons why this level of detail matters. Theoretically, five positions reflect the nuanced changes described by each framework, aligning with the complexity of meaning-making growth during the college years. Practically, the five-position structure enables greater measurement precision. Our preliminary Rasch calibrations demonstrated that five distinct positions allow clearer discrimination among student responses and support the construction of vertically ordered scenarios that increase in developmental complexity. Table 1 shows how these five positions correspond to the developmental stages identified by Kegan, Parks, and Baxter Magolda.

We elaborate on the five positions by identifying four facets, or characteristics, that reflect the cognitive, interpersonal, and intrapersonal characteristics of meaning-making capacity (O'Keefe et al., 2025). They are: cognitive ability, relational awareness, conflict resolution, and sense of responsibility. The first facet of CDT is cognitive ability. Appendix A presents the range in cognitive ability across the five positions.

Table 1. Correspondence Between Five LAMP Positions and Three Constructive Developmental Frameworks

LAMP Position	Kegan (1994)	Parks (2000)	Baxter Magolda (2001)
1	2nd Order	—	—
2	Emerging 3rd Order	Adolescent/Conventional	Following external formulas
3	3rd Order	Adolescent/Conventional	Following external formulas
4	Emerging 4th Order	Emerging adult	Crossroads
5	4th Order	Tested adult	Self-authorship

Note. The five LAMP meaning-making positions are aligned with comparable stages from three developmental theories: Robert Kegan's orders of consciousness, Sharon Daloz Parks' model of adult meaning-making, and Marcia Baxter Magolda's theory of epistemological development. Dashes (—) indicate no directly corresponding classification in the respective framework.

With development comes the ability to better recognize and make sense of the world as students move from concrete-operational, to early-, then later-formal-operational cognitive capacity. For example, whereas those in position 1 are not guided by values as such, those in positions 2 and 3 are able to recognize and talk about the values they hold and try to live by, but may be challenged to prioritize among competing values. Finally, those in positions 4 and 5 are not only comfortable in discussing values, but also are able to prioritize among competing values. Similarly, increased cognitive complexity is also reflected in their ability to recognize the relationship between cause and effect, and among behaviors, ideas and worldviews.

The second facet is relational awareness (Appendix B). With appropriate challenges and support, it becomes increasingly apparent to the developing person that they have been and are in relationships with others. It is not just that they are more able to engage in relationships with mutuality but that they are beginning to recognize the relationships of which they are already a part. The person in Position 1 is unselfconsciously self-interested. They may recognize that others have a point of view, just as they do, but they do not recognize they have relationships, with implicit expectations, with others. The move towards Position 2 brings a capacity to recognize relationships, and to begin to take others into consideration. In time and with encouragement (Position 3), tacit belonging is replaced by conscious alignment, whereby one takes on and even can become a proponent for the values and expectations of particular relationships. Position 4 marks a shift in one's sense of dependence on, and obligation to, one's relationships, with the awakening recognition that one's perspective is distinct from that of important others, and one may risk sharing that perspective. By Position 5 one is more confident in claiming their own voice and, in general, is not distressed by diverse perspectives within their relationships.

Conflictual situations are spaces in which intrapersonal meaning-making is made apparent, for how one chooses to resolve conflict reflects a sense of self in relation to others. In conflict resolution (facet 3, Appendix C), the person in Position 1 is able to think logically about their perceptions, but unable to simultaneously consider the perceptions of others; they tend to focus on their own interests and preferences. Therefore, in a conflict, they usually resolve in favor of their own needs and desires, but not for any principled reason. Those in Position 2 have become tacitly aware of the relationships and expectations around them, and so tend to conform to what "everyone" expects. In time, it may develop into a greater sense of identity, grounded in particular relationships and social alignments (Position 3), and so begin to resolve conflicts based on explicit values or relational expectations. As one develops the capacity to distinguish between oneself and one's relationships (Position 4), one becomes more likely to risk resolving a conflict in a manner different from what "everyone expects," even while mindful of those expectations. Finally, the person in Position 5 is more integrated, bravely resolving conflicts in a manner that reflects their principles and values, even as the relationship or context varies.

Facet 4, sense of responsibility, reflects both the interpersonal and the intrapersonal dimension where a person's sense of responsibility can be a strong indicator of meaning-making capacity (Appendix D). In Position 1, since relationships and obligations to others are not yet acknowledged, there is no felt sense of responsibility to others. On the other hand, the person in Position 2 has become tacitly aware of their relationships and social connections, resulting in a move to feel responsible to those relationships. As one moves to Position 3, there is a stronger alignment with particular

relationships, and the values they represent, thus enabling an ability to be more mindful of, and responsible to, those relationships and values. As the person comes to see themselves as having relationships, but not being defined by them, they can begin (Position 4) to determine how they will be in them; that is, they increasingly determine their own rules of engagement for their relationships. As they become more consistently able to name who they are in their relationships (Position 5), they can better see how their attitudes and actions shape the relationships for themselves and others. As such, they become more responsible for those relationships; they recognize they are part of determining the rules of engagement for everyone.

Across these four facets (cognitive ability, relational awareness, conflict resolution, and sense of responsibility), students move from concrete, self-focused interpretations of experience toward increasingly complex capacities to integrate diverse perspectives and act with principled autonomy. Cognitive ability reflects how individuals understand ideas and systems, progressing from concrete, then categorical thinking to the ability to manage competing priorities. Relational awareness captures a shift from unexamined self-interest to mutual, reciprocal relationships. Conflict resolution traces how individuals engage differences, moving from avoidance or compliance to actively balancing self-expression and relational integrity. Finally, sense of responsibility highlights how individuals come to first recognize then potentially shape the "rules of engagement" in their communities. These gradations informed scenario writing by providing a developmental blueprint: each scenario systematically combines facet levels to portray a single, coherent position along the meaning-making continuum.

Scenario construction

For each context, the development of the ordered scenario sequence representing the five levels of meaning-making capacity adhered to the Rasch/Guttman Scenario (RGS) scale construction framework as detailed in Ludlow et al. (2014, 2020a, 2021). This framework consists of the following iterative and theory-driven stages: (a) specifying the target construct; (b) identifying relevant facets and drafting initial facet-level descriptions; (c) articulating developmental gradations within each facet; (d) determining the structural template for scenario construction; (e) formulating and integrating mapping sentences to build theory-anchored scenarios; (f) selecting response formats and drafting respondent instructions; and (g) empirically evaluating the degree of alignment between theoretical expectations and model-based estimates. In the present application, meaning-making capacity serves as the target construct, operationalized through four integrated facets: cognitive complexity, relational awareness, conflict resolution, and sense of responsibility. Each scenario within a given context reflects a fixed position on the developmental continuum beginning with Position 1 and ascending through Position 5. The resulting scenario sets are thus designed to function as a vertically ordered scale with increasing complexity across the meaning-making continuum, as defined by constructive developmental theory and later supported through Rasch model calibration.

The scenarios were deliberately constructed to be brief yet semantically rich portrayals of context-specific situations that were intended to be both ecologically valid and experientially recognizable to college students. Each scenario is aligned with one of five hypothesized meaning-making positions and is written in language representative of that developmental level. This design feature serves a dual function: first, to anchor the

narrative in familiar domains of student experience; and second, to differentiate the scenarios linguistically and conceptually across levels of developmental complexity. At the lower end of the scale (Position 1), the character's language is rooted in concrete, immediate preferences and personal advantage. By contrast, by Position 5 the language reflects the capacity to articulate abstract principles, consider competing values, and make intentional commitments. This progression in linguistic and conceptual sophistication across the scenario sequence is essential to supporting the scale's developmental structure and interpretive utility.

The current implementation includes two parallel Rasch/Guttman Scenario (RGS) instruments—one for the school context and one for the family context—each designed to assess meaning-making capacity within a contextually grounded, potentially conflictual situation. These two contexts were purposefully selected due to their salience in the developmental lives of college students.

The instructions were the same for both instruments.

> Below you will find a description of two contexts (school and family) and a situation within each of those contexts. Following the contextual description, you will find 5 scenarios in which a character is making sense of the situation in that context. Within each context, a single character is named for all scenarios. Obviously, one character cannot see a situation from all perspectives simultaneously. The use of the same character is to emphasize that all elements of the context and situation remain the same for all scenarios; what changes is how the character makes sense of the situation. Your task is to read each scenario and determine how similarly or differently from the character you would approach the situation if you were in it.

For both instruments we utilized a "comparative response option" approach (Ludlow et al., 2014). Here the respondent is asked to compare themselves to the character in each scenario. More specifically, the instructions direct the respondent "to imagine yourself in each situation and consider the degree of similarity between your approach to that of the character's approach." They are then offered five progressively higher-level meaning-making capacity options:

1. In the past I might have approached this kind of situation similarly to X, but I don't now.
2. Sometimes I approach this kind of situation similarly to X, but not regularly anymore.
3. Now, I regularly approach this kind of situation similarly to X.
4. Sometimes I approach this kind of situation similarly to X, but not as regularly as I would like to.
5. Someday I might approach this kind of situation similarly to X, but I don't now.

These response options reflect plausible changes over time, and require the respondent to situate themselves vis-a-vis each scenario in a manner that reflects their current meaning-making capacity. Tables 2 and 3 present the situation and meaning-making scenarios for the school and family contexts, respectively.

Table 2. *Context and Scenarios for School*

Scenario	Description
Context	John is a student in a college and a program that he likes. He is enrolled in a course required for his program of study and he gets on fine with the professor. There has never been any indication from this professor that there are consequences for holding diverse perspectives from the professor. In a question on an exam, the professor instructs students to offer their perspective on an issue, applying the course material to support their answer. Each scenario below reflects a different way John might approach the situation.
John1	"I don't like questions like this. It feels like a trap. Do they want my opinion or the course material? Which is it? If they know what answer they want, they should just tell us what it is and not try to mess with us."
John2	"I really don't see how my opinion matters on this. There's got to be a right way to answer this question, whether it's naming the right steps or the right theorists. My job is to figure out what the right answer is."
John3	"There are multiple ways to answer this question, and I really think differently about this than the professor does. But there's no way in the world I'm going to offer anything different from what I think the professor wants."
John4	"There are multiple ways to answer this question, and some are better than others. On this question, I really think differently than the professor does. It feels a little risky, but I think if I make a strong, well-supported argument using the course material, I should be alright."
John5	"There are multiple ways to answer this question, and some are better than others. If I have the space and time, I think I will lay out a few options and argue the merits and limits of each. But in the end, it will be really important that I make a clear argument for which one I think is most effective. I like questions like this; they offer a chance to really engage with the material."

Note to the reader: This table presents a developmental sequence of five possible responses by John, illustrating increasingly complex ways of engaging with the exam question and course material. Scenario 1 reflects the least complex meaning-making perspective, whereas Scenario 5 reflects the most complex meaning-making perspective.

Table 3. *Context and Scenarios for Family*

Scenario	Description
Context	Maria comes from an intact family made up of immediate and extended members who have occasion to interact with some regularity and are considered a relatively close family. By and large, everyone gets on well. There is no history of abuse or trauma. When it comes to questions of conflict or differences of opinion among family members or between Maria and family members, these are possible ways Maria might make sense of the situation.
Maria1	"When it comes to disagreements, it's just a matter of what you like and don't like. If I am doing something that my family disagrees with, I just try not to get caught. But sometimes when I get caught, they say I 'should have known better'. I don't get it. What should I have known?"
Maria2	"In my family, we really know what family stands for, which is important to me. For the most part, I'm happy to contribute to making things go smoothly. But on the rare cases we disagree, I just keep that to myself. I don't want to make a scene or get anyone upset. Voicing a different opinion from my family is not that important to me."
Maria3	"Sometimes I just want to shout from the rooftop that I don't agree with what they are all thinking, but I never do, because it would be too upsetting. Other times I try and it's a major fail; it comes out all wrong. I think I'm right and don't understand why they don't agree with me. They are my family; shouldn't they get me? Most of the time, I just keep things to myself."
Maria4	"It may seem easier to keep my mouth shut, but sometimes the issues are too important to let go. As risky as it feels, I'm getting better at expressing myself to my family on things that are important to me. It's important to me that we get along as a family, which means I have to contribute to how we listen and speak to each other."
Maria5	"Even though we're family and we love one another, we have very different perspectives and personalities. Talking to each other about difficult issues takes real care. Not only do I have to recognize their point of view, I have to express mine in a way they can understand."

Note. This table presents a developmental sequence of five possible responses by Maria, illustrating increasingly complex ways of navigating conflict and differences of opinion within her family. Scenario 1 reflects the least complex meaning-making perspective, whereas Scenario 5 reflects the most complex meaning-making perspective.

Preliminary field trials of the initial LAMP-C instrument yielded encouraging evidence of construct alignment (O'Keefe et al, 2022); however, the hypothesized "ladder-like" hierarchy of scenario difficulties did not materialize. In addition, we observed a few student responses that departed from model expectations—most often on the lowest complexity scenarios. Some students at higher locations on the scale responded in ways that indicated disagreement or disengagement with scenarios that had been written to be low-complexity. In one case, a student provided a pattern of responses that represented severe model misfit and raised questions about either respondent interpretation or scenario design. Rather than simply dismiss these as outliers, we treated them as important indicators of how respondents were making sense of the task. These response patterns invited closer scrutiny of scenario clarity, response instructions, and response options.

We collaborated with a *Student Voices Advisory Group*—a demographically and experientially diverse group of undergraduate students—who reviewed not just the scenarios themselves, but also the scoring prompts and overall framing of the instrument. This collaborative effort led to several revisions intended to enhance clarity and developmental coherence: (a) the CDT scenario framework was recast from seven to five positions to more clearly reflect psychometrically distinguishable stages; (b) the conflictual situation within each context was standardized rather than varied; (c) a single recurring character (John or Maria) was introduced to anchor each scenario set; and (d) the response options were restructured to align more closely with a developmental self-assessment format. These changes were not trivial; they represented an attempt to make the task more ecologically valid and more interpretable for students, while preserving the structure needed for Rasch calibration. Following additional cognitive interviews with SVAC members, the revised instrument was administered in full-scale form.

Sample

Data were collected through an anonymous online survey, administered via Qualtrics, and disseminated to students at a private university in the northeastern United States. They were recruited through multiple channels, including email invitations sent via university listservs, direct outreach from faculty and staff, and physical flyers distributed in high-traffic student areas such as dining halls, residence halls, and the athletic center. The research team collaborated with university departments and student organizations to maximize outreach. Recruitment materials emphasized the voluntary nature of participation and highlighted the opportunity to enter a raffle for a $50 Amazon gift card, providing a 1 in 10 chance of winning. Upon completion, they were directed to a separate webpage where they could enter the raffle without their contact information being linked to their responses, ensuring anonymity.

The study was approved by the Boston College Institutional Review Board (Protocol 23.074.01e) (Ludlow, 2023). Inclusion criteria required participants to be at least 18 years old and currently enrolled. There were no exclusions based on gender, race, or ethnicity. Recruitment efforts targeted a diverse student population to ensure broad representation and to assess potential demographic differences in responses. These efforts yielded 114 undergraduate and graduate students—Appendix A presents the demographic breakdown.

While these procedures yielded a somewhat diverse group of respondents, the reliance on a single-institution sample represents an important limitation. This private university has distinctive characteristics in terms of institutional culture, demographics, and developmental supports, which may influence meaning-making processes. Consequently, our findings should be interpreted with caution and not assumed to generalize to students in other institutional settings, such as public universities, community colleges, or institutions with different missions or student populations.

Analysis

The Rasch rating scale model (Andrich, 1978; Wright & Masters, 1982) was used to analyze the extent to which data conformed to model-based expectations; that is, if the empirical results provide construct validity evidence congruent with the theorized scale structures. For a particular context, the model may be expressed as:

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

For the present analysis, the key parameters of the model are: π_{nix} denotes the probability that person *n*, when presented with scenario *i*, responds in category *x*; x_{ni} is the observed score of person *n* on item *i*; *m* is the maximum possible rating scale category and k =0 identifies the lowest score category in the response set (this ensures that the denominator sums the category characteristic curves from 0 through m so that the probability for category *x* is properly normalized); β_n represents person *n*'s estimated standing, or location, on the latent meaning-making continuum within a given context; δ_i is the scenario difficulty parameter, indicating the location of scenario *i* along the same latent continuum; and τ_j is the threshold parameter associated with the transition from response category *j* to the next higher category within a scenario. In this framework, scenario difficulty reflects the extent to which a scenario tends to elicit higher-level responses, with higher level response categories corresponding to greater meaning-making capacity. The analysis was conducted using WINSTEPS (Linacre, 2024).

Results

For each context, the systematic Rasch/Guttman scenario construction process generates an *a priori* expected ordering of the scenarios. Construct (or variable) maps, are an essential component of Rasch instrument development and analysis procedures. They are graphical representations that simultaneously display the scenario difficulty estimates and the respondent level of meaning-making estimates, both as locations along a logit continuum (Ludlow & Haley, 1995). If our reformulation of constructive developmental theory is correct, and the scenarios have been well designed, and the respondents are appropriate for the task, then the empirical ordering of scenarios within a context should

agree with the theoretical expectations for that set. Furthermore, the ordered sequence should define a quantitative continuum along which a student may then be located. For each context, students' locations on the corresponding scale thus provide a basis for a specific, actionable interpretation of their meaning-making capacities.

In addition to the construct maps, goodness-of-fit statistics evaluate how well the observed responses conform to the responses predicted under the Rasch model. Residuals are the differences between the observed and predicted responses. Variance weighted mean squared residuals ("Infit MNSQ") and unweighted mean squared residuals ("Outfit MNSQ") above 1.4 indicate potential model misfit (Linacre, 2019).

Figure 1 and Figure 2 present the construct maps for the school and family contexts, respectively. To the right of the vertical lines are the names of the scenarios in descending order of their difficulty estimates from hardest/most complex (John5 and Maria5) to say "Someday I might approach this kind of situation similarly to X, but I don't now" (top of the continuum) to easiest/least complex (John1 and Maria1) to say "In the past I might have approached this kind of situation similarly to X, but I don't now" (bottom of the continuum). The full text of each scenario is presented roughly adjacent to each label in order to qualitatively define each meaning- making position.

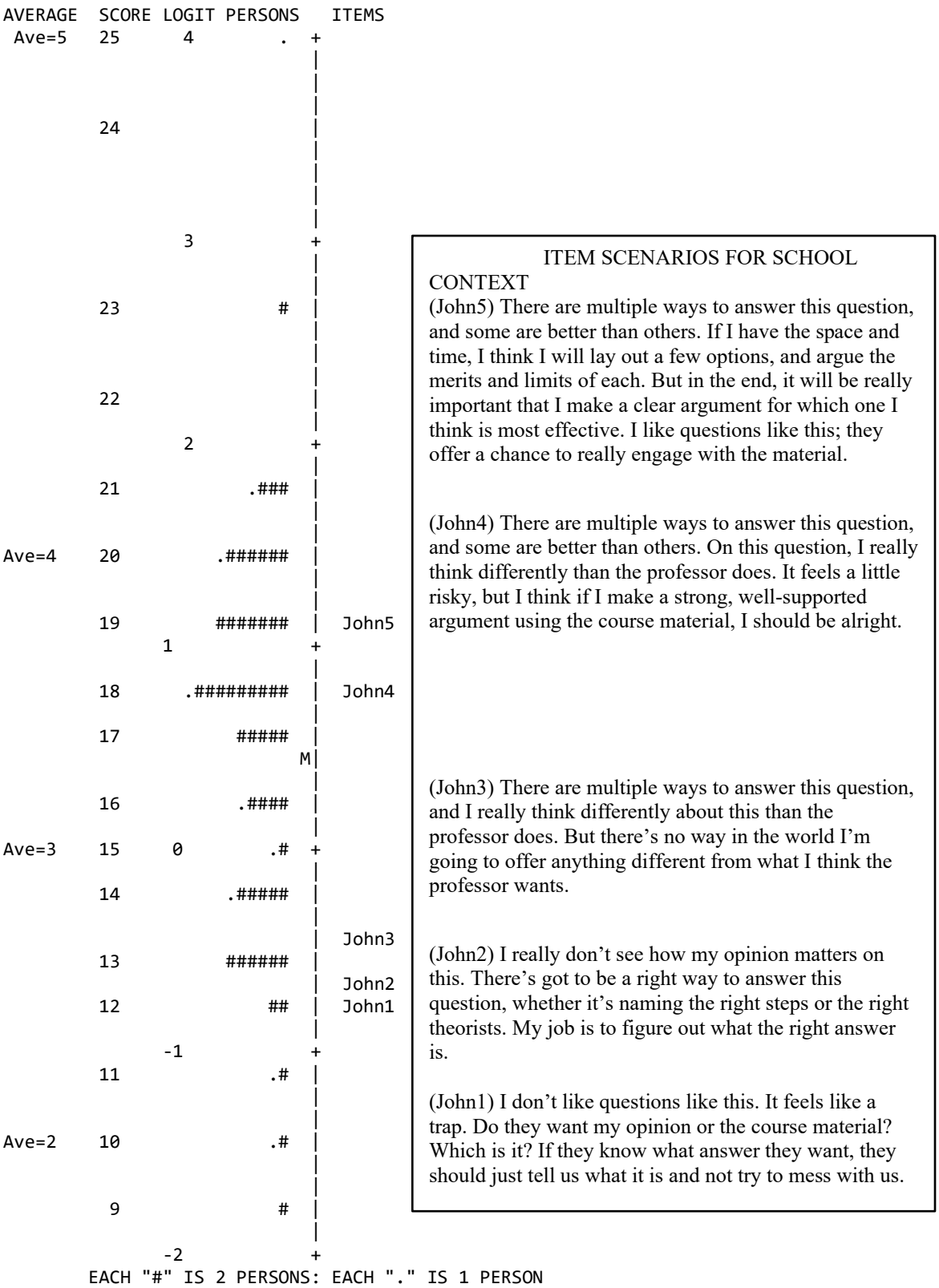
On the left sides of the vertical lines are the students meaning-making location estimates based on their raw score responses to the scenarios. The "#" symbols represent either 1 or 2 students at a given location. The "M"s represent the mean student location estimates. The "AVERAGE" column indicates where the raw score average responses of "2", "3" and "4" are located. The "SCORE" column corresponds to where the students are located in terms of their actual raw score. Furthermore, the Score column allows a reader to answer the scenario prompts, add up their responses, and locate where their responses put them on the respective continua. The "MEASURE" column contains the corresponding student and item locations in logits. Overall, both the scenario estimates and students' estimates display an appropriately broad range from lower to higher levels of complexity in each context.

Below Figure 1 and Figure 2 we provide a narrative to help the reader understand the connection between what a student's score means in terms of the scenarios adjacent to their location. That is, we answer the question: how is a given score to be interpreted?

Based on the systematic development of the scenarios: (a) John5 and Maria5 were written to be the most challenging for a student to say "Someday I might approach this kind of situation similarly to X, but I don't now" and their locations at the top of each construct map confirm that the students did, in fact, find them to be the most challenging. Likewise, John4/Maria4 and John3/Maria3 and John2/Maria2 were written to be mid-level in their degree of complexity, and John1/Maria1 were constructed to be the least challenging to say "In the past I might have approached this kind of situation similarly to X, but I don't now". The locations of these scenarios operationalizing each context, displayed in Figures 1 and 2, are consistent with their intended ordered positions, as presented in Tables 2 and 3.

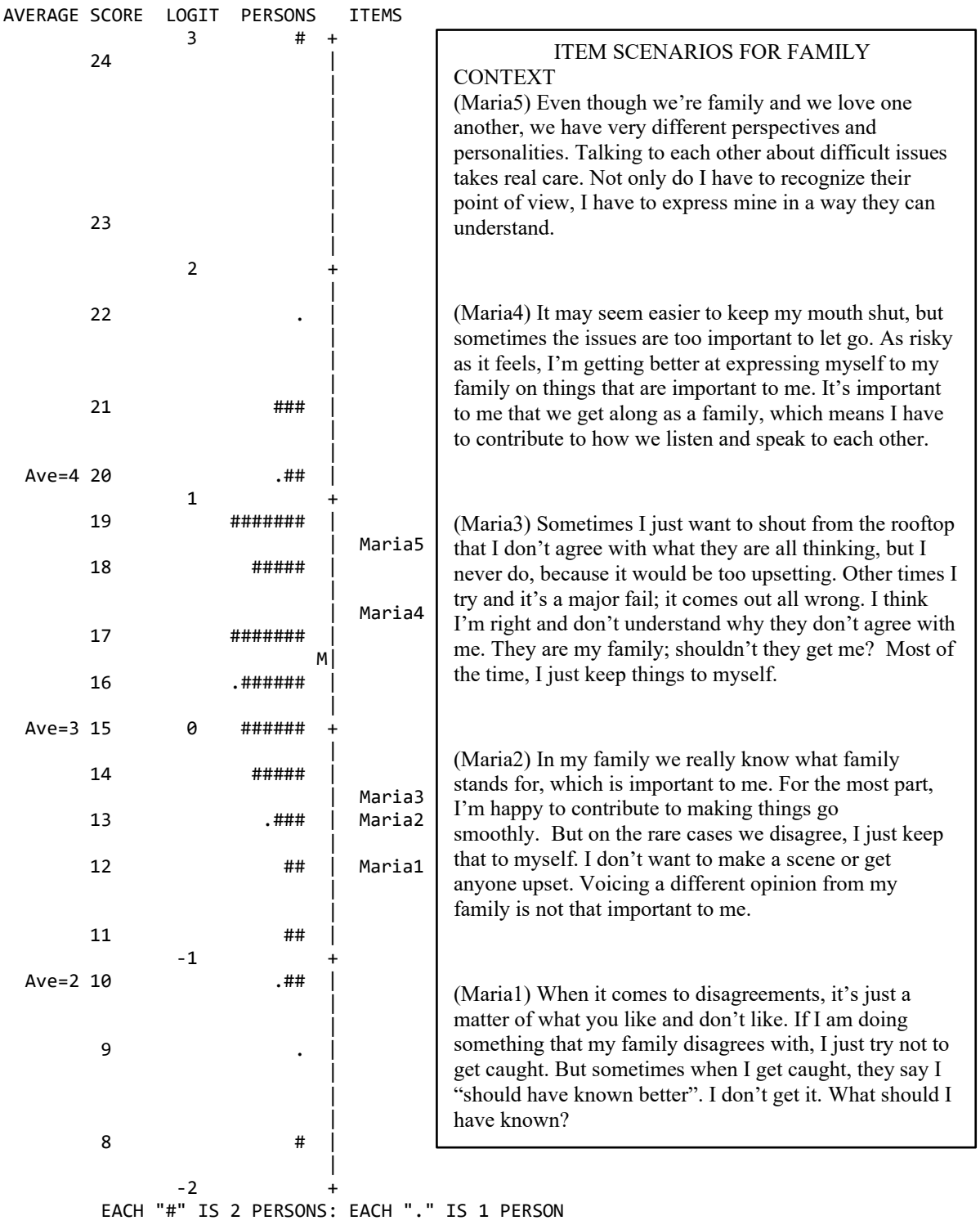
The text boxes in Figures 1 and 2 provide substantively grounded and nuanced descriptions for interpreting individuals' scores. These interpretations derive directly from the construction of the scenarios, based on the integrated facet levels. For example, in Figure 1, students with meaning-making scores of 10 or below don't like questions that contain multiple perspectives, they tend to think concretely, may be unable to take the

Figure 1. Construct map for the School context.



Legend: AVERAGE=the total scale score/5; SCORE=the sum of the 5 item responses; LOGIT=the logistic unit of measurement into which the raw scores are transformed (Ludlow & Haley, 1995).

Figure 2. Construct map for the Family context



Legend: AVERAGE=the total scale score/5; SCORE=the sum of the 5 item responses; LOGIT=the logistic unit of measurement into which the raw scores are transformed (Ludlow & Haley, 1995).

professor's point of view, focus on their own self-interest, and take no responsibility beyond their self-interest. At scores around 12, they still don't like this kind of question but will try to figure it out because they are beginning to recognize what the professor values, they are aware of the expectations for being an engaged member of the class, their self-interest begins to take into account the expectations of the class, and they begin to feel a need to respond appropriately within the class. At scores around 15 they don't necessarily dislike this kind of question as they begin to understand the professor's point of view and purpose and they do have a sense of belonging, but want to avoid conflict as they engage more fully with the task. At scores around 18, they begin to take the work more seriously and they are aware that they may think differently from the professor, but they are willing to take a risk and some responsibility to express their own position. Finally, at scores 20 and above, they recognize what the professor is asking of them, they are confident in their ability to express their point of view, and they are very aware of their relationship to the professor and how they can contribute to that relationship.

In Figure 2, a student with meaning-making scores of 10 or below tries to stay out of family conflicts, they don't understand why the family disagrees with them, they are unable to take the family's point of view, they focus on their own self-interest, and they take no responsibility beyond their self-interest. At scores around 12, they still try to avoid family disagreements even though they are aware the family disagrees with them and their self-interest keeps them quiet although they will engage when things are going smoothly. At scores around 15, they want to express their opinions but don't yet and they are conscious of the different positions of other members but are still not confident in their ability to express themselves appropriately. At scores around 18 and above, they begin to recognize the priorities and points of view of other family members and they are beginning to take some risks and responsibility for expressing and contributing to the family. Finally, at scores 20 and above, they accept that everyone has different experiences and perspectives and points of view about almost everything in family life, they have their own positions that they are confident in expressing, and they take a responsibility in contributing to the well-being of the family.

Appendix B and Appendix C contain the scenario characteristic statistics for the school and family contexts, respectively. The scenarios are displayed in the order in which they were presented. Both sets of scenario difficulty estimates (in their logit values and corresponding "Total scores") are monotonically increasing, as would be expected under the theoretical model. The "Model SE" is the standard error of the scenario estimates. Both sets of Infit and Outfit mean-squares suggest there were no strong patterns of unexpected responses. The item-total correlations for both sets of scenarios are all moderately positive (ranging from .45 to .66), as intended.

Appendix D and Appendix E present the category characteristic statistics for the school and family contexts, respectively. "Category Label" refers to the code for the response options. "Count" refers to how many times that response was provided. Both the "Observed Average" (in logits) for those who responded in each of the categories and the Andrich thresholds (the *transition* between adjacent ordered categories where a respondent becomes equally likely to choose category k instead of category k–1 (Andrich, 1978) monotonically increase, as intended.

The Infit and Outfit mean squares for the category responses show a slight degree of misfit for category one in both the school and family contexts but are otherwise

unremarkable. An inspection of the observed and expected responses for the entire sample revealed one student with an observed pattern of "1 5 4 4 4" for the five school scenarios. Their expected response to scenario one (John1) was 4.2 and their z-residual was -4. Two other students had -2 z-residuals on scenario one. For the family scenarios, there were four students with surprising responses to scenario one (Maria1) with z-residuals of -3, -2, -2, and -2.

A type of sensitivity analysis was performed, where the three students with the unusual John1 school responses and the four students with the unusual Maria1 family responses were removed from their respective data sets. The analyses were re-run and the item separation and reliability for the school scale increased from 6.63 and .98 to 7.02 and .98, respectively. The items maintained their same monotonic order, levels of acceptable fit statistics, and the category 1 fits decreased from 1.34 and 1.33 to 1.16 and 1.14 for Infit and Outfit, respectively.

The item separation and reliability for the family scale improved from 4.92 and .96 to 5.34 and .97, respectively. The items maintained their same monotonic order, levels of acceptable fit statistics, and the category 1 fits dropped from 1.52 and 1.50 to 1.4 and 1.38 for Infit and Outfit, respectively. In contrast, the item difficulty standard errors increased in each scale due to the decrease in the sample sizes. Although it is unclear why these different students responded relatively unexpectedly to the first scenarios in the two different contexts, their surprising responses may reflect a "start-up effect" due to their unfamiliarity with this type of unique item.

Finally, Rasch model item separation and item reliability indices (Linacre, 2002; Wright & Masters, 1982; Wright & Stone, 1979) help evaluate the quality of the item set in distinguishing among different locations on the latent trait—in this case, meaning-making capacity across school and family contexts. Item separation indicates how well the items in the instrument are spread out along the measured construct (school = 6.63; family = 4.92). Item reliability reflects the replicability of the item difficulty hierarchy if these same items were administered to another sample of similar respondents (school = .98; family = .96).

High separation indices (e.g., 6.63 or 4.92) suggest there are marked difficulty location distinctions among the items across the latent continua. Reliability values >.90 suggest that the item hierarchies are stable and would likely replicate in other samples. Collectively, these indices provide supportive preliminary evidence that the LAMP-C instrument is both psychometrically robust and theoretically well-aligned for measuring developmental progression across contexts.

Discussion

The purpose of this study was to assess the feasibility and psychometric defensibility of assessing meaning-making using the Rasch/Guttman Scenario (RGS) scale methodology, grounded in a reformulation of constructive developmental theory (CDT). Measuring a developmental construct like meaning-making is difficult—it requires clarity not only about the nature of the construct and how it unfolds over time, but also how it can be represented in ways that respondents in the target population can meaningfully engage with. The RGS scale methodology provided a promising strategy for addressing these formidable challenges.

From the outset, we adopted a facet-based approach to construct modeling, drawing on CDT to identify four core facets of meaning-making: cognitive ability, relational awareness, conflict resolution, and sense of responsibility. These facets were theoretically defensible and we anticipated that they could be operationalized in a reasonable way, though we recognized this would need to be tested empirically. Our aim was to integrate these facets into a series of scenarios—set in the contexts of school and family—that would vary systematically in complexity in order to provide a developmental ladder of meaning-making within each context. We used a sentence-mapping strategy to scaffold the complexity of each scenario corresponding to increasing orders of consciousness as described in Kegan's model.

As reflected in the construct maps and scenario calibration statistics, the instruments yielded measurement structures that were consistent with our developmental hypotheses. In both the school and family contexts, the scenario difficulty estimates increased monotonically (as displayed in Figures 1 and 2, and reported in Appendices B and C), and the student meaning-making estimates spanned a wide range of the latent trait continuum (also seen in Figures 1 and 2). Rasch fit statistics, separation and reliability indices, and item-total correlations supported the internal structure of each scale and the ordered response thresholds confirmed that the rating categories functioned as intended (as reported in Appendices D and E). These are promising results, particularly given the complexity of the construct and the nascent stage of instrument development.

The iterative process of defining the construct, building scenarios, testing, identifying misfit, revising, and retesting—what Braun and Kirsch (2025) refer to as establishing procedural validity—is typical of how Rasch models are used to support the measurement of developmental constructs. As Ludlow et al. (2020a) have emphasized, model misfit is not simply a problem to fix; it is valuable diagnostic information that invites theory refinement and deeper construct understanding. In this case, the feedback we received from students in early phases of our pilot testing confirmed that some of our early assumptions about developmental clarity, linguistic accessibility, and contextual nuance needed to be reconsidered.

Because the Rasch/Guttman Scenario (RGS) scale methodology produces calibrated scores linked to specific, interpretable scenario responses, the results are theoretically very actionable and could be used to inform advising, program design, and longitudinal evaluation. For example, the construct maps in Figures 1 and 2 show how student locations along the continua correspond to qualitatively different ways of making sense of conflict. A student at a score near 10 on either the school or family continuum views conflict through a lens of self-protection and rule-following, whereas a student near 20 appreciates competing perspectives and acts with confidence and responsibility.

These developmental interpretations can guide advising conversations. Advisors can use score locations to match their approach to a student's demonstrated capacity, offering more explicit structure for those at earlier levels and encouraging autonomy and leadership for those at later levels. Similarly, program designers can use aggregated results to align courses and co-curricular experiences with institutional goals. Curricula may be intentionally sequenced to provide opportunities for growth across the four facets of meaning-making: cognitive ability, relational awareness, conflict resolution, and sense of responsibility. Further, institutions could track student growth over time to evaluate the effectiveness of interventions. In this way, LAMP-C creates a direct link between theory

and practice: a developmental framework grounded in constructive developmental theory that also informs the day-to-day work of higher education.

Limitations and Future Research Directions

Although this study provides credible, preliminary evidence of construct alignment and developmental coherence for the LAMP-C instruments, several limitations temper interpretation and indicate directions for future research. The single-site, convenience sample severely constrains external validity. Institutional culture, mission, and student supports likely shape how respondents interpret the scenarios and construe developmental challenges; replication across diverse institutional types—including public universities, community colleges, and Minority-Serving Institutions (MSIs), as well as residential versus commuter settings—is needed before generalizing item hierarchies or score distributions. For instance, community colleges tend to serve many nontraditional students who juggle academic demands alongside work and family responsibilities—circumstances that may shape how they approach relational and conflict-resolution challenges. Similarly, students at MSIs may engage with questions of identity and responsibility in culturally specific ways. These differences make it essential to ensure that both the scenarios and the response formats in LAMP-C are relevant and meaningful for a variety of student populations. Expansion to additional contexts will allow testing of measurement invariance and differential item functioning (DIF) across subgroups defined by institution type, class year, first-generation status, race/ethnicity, and language background.

Response-process and linguistic factors also constitute important validity considerations. Scenario comprehension depends on the clarity and cultural resonance of the language and idioms used. Differences in first-language background, disciplinary jargon exposure, or reading fluency could influence ratings independently of meaning-making capacity. Although cognitive interviews with a *Student Voices Advisory Group* informed wording revisions, this group was small and also from a single institution. More systematic response-process evidence across diverse types of students, including linguistic subgroups—can help to identify sources of construct-irrelevant variance. Future testing will, therefore, combine multi-site replication with qualitative response-process studies to strengthen both external and internal validity.

Because the Rasch/Guttman Scenario scale methodology uses comparative, self-referential response options, ratings may also be susceptible to self-presentation or social-desirability effects. Students might over- or under-identify with a character's stance based on perceived desirability rather than genuine reasoning. Future iterations of LAMP-C should include validity checks such as social-desirability short forms and multitrait–multimethod comparisons to signal potential bias. At the same time, examining predictive validity—for instance, how meaning-making scores relate to advising engagement, leadership participation, or retention—will provide evidence of applied utility and real-world relevance.

A further methodological limitation involves the "start-up effects" observed when respondents first encounter a novel assessment format. Unexpected response patterns on initial items (e.g., 1–5–4–4–4) likely reflect poor format acclimation rather than construct differences, inflating residual misfit and attenuating stability. Prior Rasch research (Ludlow, 1983, 1986; Linacre, 2002; van der Linden, 2009) demonstrates that introducing

warm-up items can mitigate these distortions. Future LAMP-C implementations will include randomized item order and brief warm-up scenarios to reduce this source of measurement error.

Another consideration is context dependency, which is inherent to scenario-based measurement. School and family situations were chosen for relevance to college students, yet the degree to which each context mirrors a respondent's lived experience may vary. This variation can influence scenario response choices and consequently, observed developmental order. Future work will consider incorporating site-specific adaptations through focus groups with students, faculty, and staff to ensure that scenarios authentically reflect the developmental challenges salient within each institutional and cultural context. For example, community-college scenarios may highlight work–school balance, while those for MSIs could emphasize identity negotiation or systemic inequities.

Building on these refinements, the next phase of research will pursue a longitudinal validation agenda to test the sensitivity of LAMP-C to developmental change. Administering pre–post and multi-wave assessments within academic courses or co-curricular programs will help determine whether scores capture expected growth patterns, such as movement from compliance or avoidance toward principled self-authorship, and whether interventions targeted to specific facets (e.g., relational awareness) yield differential gains.

In collaboration with partners such as student affairs and career counselors, we also plan to design and evaluate score-reporting tools that translate person measures into developmental bands, accompanied by plain-language narratives and facet-specific growth suggestions. These reports will be tested for clarity, utility, and actionability through advisor and student feedback, ensuring that LAMP-C not only functions as a valid measure but also as a practical developmental feedback tool. Finally, future development will expand the current two-context framework by introducing a third relational context—friend relationships—within the LAMP-C instrument. This expansion will allow cross-context modeling of meaning-making development and provide a more complete picture of how students navigate meaning across personal, academic, and social domains. Collectively, these steps aim to enhance both the psychometric robustness and the practical utility of the LAMP-C framework, with the hope of advancing our understanding of meaning-making as an important component of student development across diverse higher-education settings.

Conclusion

This study focused on students from a single private university in the Northeast, offering an important starting point for our work. While the study findings provide a solid foundation, our broader goal is to develop a set of psychometrically integrated instruments that can capture how meaning-making evolves both within and across different contexts over time. By doing so, we aim to advance both theory and measurement, while gaining deeper insight into how meaning-making supports students' personal and educational growth throughout college and beyond.

The four core facets of meaning-making—cognitive ability, relational awareness, conflict resolution, and sense of responsibility—are not unique to any one institution or type of college. Rather, they reflect universal developmental processes that all students engage with as they navigate the challenges and opportunities of higher education and

beyond. Expanding this work to a wider range of settings and populations will allow us to create tools that are not only grounded in strong theory, but are also practical and adaptable for those working to foster student growth and transformation.

Appendices

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4), 561–573. https://doi.org/10.1007/BF02293814

Baxter Magolda, M. B. (2001). *Making their own way: Narratives for transforming higher education to promote self-development*. Stylus.

Baxter Magolda, M. B. (2009). The activity of meaning making: A holistic perspective on college student development. *Journal of College Student Development, 50*(6), 621–639. https://doi.org/10.1353/csd.0.0106

Baxter Magolda, M. B. (2010). The interweaving of epistemological, intrapersonal, and interpersonal development in the evolution of self-authorship. In M. B. Baxter Magolda, E. G. Creamer, & P. S. Meszaros (Eds.), *Development and assessment of self-authorship: Exploring the concept across cultures* (pp. 25–43). Stylus.

Boston College (2023). "The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Boston College (protocol code: 24.076.01e; approval date: 9 August 2023)."

Braun, H. & Kirsch, I. (2025). Large-scale assessments in international contexts. In L. Cook & M. Pitoniak (Eds.) *Educational Measurement* (5th ed.). Oxford U. Press.

Creamer, E. G., Baxter Magolda, M. B., & Yue, J. (2010). Preliminary evidence of the reliability and validity of a quantitative measure of self-authorship. *Journal of College Student Development, 51*(5), 550–562.

Dalton, J., & Crosby, P. (2010). How we teach character in college: a retrospective on some recent higher education initiatives that promote moral and civic learning. *Journal of College and Character*, 11, 13-17. https://doi.org/10.2202/1940-1639.1264

Fallar, R., Hanss, B., Sefcik, R., Goodson, L., Kase, N., & Katz, C. (2019). Investigating a quantitative measure of student self-authorship for undergraduate medical education. *Journal of Medical Education and Curricular Development, 6*, 2382120519896789. https://doi.org/10.1177/2382120519896789

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.

Helsing, D., & Howell, A. (2014). Understanding leadership from the inside out: Assessing leadership potential using constructive-developmental theory. *Journal of Management Inquiry, 23*(2), 186–204. https://doi.org/10.1177/1056492613500717

Kegan, R. (1982). *The evolving self: Problem and process in human development*. Harvard University Press.

Kegan, R. (1994). *In over our heads: The mental demands of modern life*. 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. https://doi.org/10.1353/csd.0.0104

Lahey, L., Felix, S., Goodman, R., Kegan, R., & Souvaine, E. (2011). *A guide to the subject-object interview: Its administration and interpretation*. Minds at Work.

Linacre, J. M. (2002). *Computer-adaptive testing: Memo 69*. Rasch Measurement SIG. https://www.rasch.org/memo69.pdf

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

Linacre, J. M. (2019). *Fit diagnosis: Infit outfit mean-square standardized*. https://www.winsteps.com/winman/misfitdiagnosis.htm
Linacre, J. M. (2024). *WINSTEPS Version 5.8.3.0*. Winsteps.com.
Ludlow, L.H. (1983). The analysis of Rasch model residuals [Unpublished doctoral dissertation]. University of Chicago.
Ludlow, L.H. (1986). Graphical analysis of item response theory residuals. *Applied Psychological Measurement*, *10*, 217-229.
Ludlow, L.H. (2023). *Measuring meaning-making* [Unpublished IRB protocol]. Lynch School of Education and Human Development, Boston College.
Ludlow, L.H. & Haley, S.M. (1995). Rasch model logits: Interpretation, use, and transformation. *Educational and Psychological Measurement, 55,* 967-975.
Ludlow, L.H., Baez-Cruz, M., Chang, W-C C., & Reynolds, K. (2020a). Rasch/Guttman Scenario Scales: A Methodological Framework. *Journal of Applied Measurement*, *21*(4), 361-378.
Ludlow, L.H., Anghel, E., Szendey, O., O'Keefe, T., Howell, B., Matz-Costa, C., & Braun, H. (2020b). The Boston College Living a Life of Meaning and Purpose (BC-LAMP) Portfolio: An application of Rasch/Guttman Scenario methodology. *Journal of Applied Measurement*, *21*(2), 134-153.
McAdams, D. P. (2013). The psychological self as actor, agent, and author. *Perspectives on Psychological Science, 8*(3), 272–295. https://doi.org/10.1177/1745691612464657
O'Keefe, T., Ludlow, L., Matz, C., Braun, H., Anghel, E., & Szendey, O. (2022, April 23). *Measuring growth in the development of meaning making: The analysis of Rasch model residuals* [Paper presentation]. American Educational Research Association (AERA) Annual Meeting, San Diego, CA, United States.
O'Keefe, T. A., Warner, L., Matz, C., Ludlow, L. H., & Braun, H. I. (2025). A reframing of meaning-making and its measurement among emerging adults. *Religions, 16*(11), 1431. https://doi.org/10.3390/rel16111431
Parks, S. (2000). *Big questions, worthy dreams: Mentoring young adults in their search for meaning, purpose, and faith*. Jossey-Bass.
Piaget, J., & Cook, M. (1952). *The origins of intelligence in children* (Vol. 8, No. 5, pp. 18-1952). New York: International universities press.
Pizzolato, J. E. (2003). Developing self-authorship: Exploring the experiences of high-risk college students. *Journal of College Student Development, 44*(6), 797–812. https://doi.org/10.1353/csd.2003.0074
Stewart, C., & Wolodko, B. (2016). University Educator Mindsets: How Might Adult Constructive-Developmental Theory Support Design of Adaptive Learning?. *Mind, Brain, and Education*, *10*(4), 247-255.
van der Linden, W. J. (2009). Conceptual issues in response-time modeling. *Journal of Educational Measurement, 46*(3), 247–272. https://doi.org/10.1111/j.1745-3984.2009.00080.x
Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.
Wright, B. D., & Stone, M. H. (1979). *Best test design*. MESA Press.