

Undergraduate Lay Theories of Abilities: Mindset, universality, and brilliance beliefs uniquely predict undergraduate educational outcomes

Lisa B. Limeri,^{†*} Nathan T. Carter,[‡] Franchesca Lyra,[§] Joel Martin,^{||}
Halle Mastronardo,^{||} Jay Patel,^{||} and Erin L. Dolan^{||}

[†]Department of Biological Sciences, Texas Tech University, Lubbock, TX 79409; [‡]Department of Psychology, Michigan State University, East Lansing, MI 48824; [§]Department of STEM Education, University of Texas, Austin, TX 78712; ^{||}Department of Biochemistry & Molecular Biology, University of Georgia, Athens, GA 30602

ABSTRACT

Students' beliefs about their abilities (called "lay theories") affect their motivations, behaviors, and academic success. Lay theories include beliefs about the potential to improve intelligence (mindset), who (i.e., everyone or only some people) has the potential to be excellent in a field (universality), and whether reaching excellence in a field requires raw intellectual talent (brilliance). Research demonstrates that each of these beliefs influences students' educational experiences and academic outcomes. However, it remains unclear whether they represent distinct latent constructs or are susceptible to the "jangle fallacy" (i.e., different names given to the same underlying construct). We conducted a multiphase, mixed-methods study to 1) evaluate whether mindset, universality, and brilliance beliefs represent conceptually and empirically discriminable concepts, and 2) evaluate whether mindset, universality, and brilliance beliefs contribute unique explanatory value for both psychosocial (e.g., sense of belonging) and academic outcomes (e.g., course grades). To address these questions, we developed and collected validity evidence for a new measure of science and math undergraduates' lay theories, called the Undergraduate Lay Theories of Abilities (ULTrA) survey. Factor analyses suggest that mindset, brilliance, and universality are distinct and empirically discriminable constructs. Structural Equation Models indicate that each lay theory contributes unique predictive value to relevant outcomes.

INTRODUCTION

Students' academic outcomes are influenced not only by their knowledge and cognitive abilities, but also by their beliefs about their cognitive abilities, collectively called "lay theories." Researchers have identified three types of lay theories: 1) mindset, 2) universality, and 3) brilliance. Mindset refers to beliefs about the extent to which intelligence is improvable or innate (Dweck *et al.*, 1995; Dweck, 1999). Universality refers to beliefs about who (i.e., everyone, or only some people) has the potential for excellence (Rattan *et al.*, 2012a,b). Brilliance refers to beliefs about whether success in a field requires innate brilliance that cannot be taught (Leslie *et al.*, 2015; Meyer *et al.*, 2015). Research on each of these lay theories reveals that they influence a wide variety of outcomes, such as what kinds of goals students set, how students interpret their successes and failures, and how strongly students feel they belong in their institution and discipline, as well as students' academic performance and retention in college (Yeager and Walton, 2011; Yeager and Dweck, 2012; Meyer *et al.*, 2015; Smiley *et al.*, 2016; Rattan *et al.*, 2018).

Ross Nehm, *Monitoring Editor*

Submitted Dec 20, 2022; Revised Jul 26, 2023;
Accepted Aug 2, 2023

CBE Life Sci Educ December 1, 2023 22:ar40
DOI:10.1187/cbe.22-12-0250

*Address correspondence to: Lisa B. Limeri
(llimeri@ttu.edu).

© 2023 L. B. Limeri *et al.* CBE—Life Sciences Education © 2023 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 4.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/4.0>).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.

These three lay theories were developed by different researchers and, to date, they have been largely studied in isolation from each other. They are conceptually similar in that they all focus on students' beliefs about the nature of their abilities and the role those abilities have in student success. Thus, there is a risk that these constructs represent different names given to the same underlying belief, i.e., the “jangle” fallacy (Kelley, 1927; Gonzalez *et al.*, 2021). It is also possible that, even if the constructs are empirically distinguishable, they may not predict unique variance in students' outcomes. For example, it is possible that mindset fully predicts variance in outcomes that universality and brilliance predict, in which case further research should focus on just a single belief. Alternatively, if mindset, universality, and brilliance all contribute uniquely to a student's academic growth and success, then further research should incorporate all three beliefs. Thus, clarity about the conceptual landscape and predictive power of these three lay theories is critical for future research on these psychological factors that influence undergraduates' educational trajectories.

We conducted a multiphase, mixed-methods study to clarify the conceptual space of mindset, universality, and brilliance beliefs, and examine how they uniquely relate to outcomes. In the process of addressing these questions, we developed and collected validity evidence for a new measure of these lay theories for Science and Math undergraduates, the Undergraduate Lay Theories of Abilities (ULTrA) Survey, which we also report here.

Mindset Beliefs

Mindset beliefs (also called “implicit theories of intelligence”) describe students' beliefs about the extent to which intelligence can be improved (growth mindset) or is innate and unchangeable (fixed mindset; Dweck, 1999). Decades of research demonstrates how mindsets act as meaning systems that influence how students interpret and respond to academic cues, especially negative feedback and failure (Hong *et al.*, 1999). Students who believe that their intelligence is fixed tend to avoid challenging situations where they might receive negative feedback, which is threatening because it is seen as a permanent and personal indictment of their fixed abilities (i.e., goal orientation; Burnette *et al.*, 2013; Smiley *et al.*, 2016). Receiving negative feedback (such as failing an exam) is discouraging, and they are more likely to withdraw and disengage (Burnette *et al.*, 2013; Smiley *et al.*, 2016). Students may engage in self-protective behaviors, like self-handicapping, where they exert less effort so that they can blame failure on lack of effort rather than lack of ability (Jones and Berglas, 1978; Rickert *et al.*, 2014). Ultimately, students with a fixed mindset tend to earn lower grades and are more likely to drop out or switch to a different field (Dweck, 1999; Burnette *et al.*, 2013; Limeri *et al.*, 2020a). In contrast, students who believe that their intelligence can improve tend to focus on mastering material. They see negative feedback as a normal part of the learning process and useful for guiding them in improving (i.e., goal orientation; Burnette *et al.*, 2013; Smiley *et al.*, 2016). They are more likely to respond adaptively to negative feedback and improve their performance over time (Jones and Berglas, 1978; Rickert *et al.*, 2014). Ultimately, students with a growth mindset tend to earn higher grades and are more likely to persist (Dweck, 1999; Burnette *et al.*, 2013; Limeri *et al.*, 2020a).

Interventions persuading students to adopt growth mindset beliefs have improved students' academic performance and retention (Blackwell *et al.*, 2007; Yeager *et al.*, 2016, 2019). However, the success of these interventions has been variable (Sisk *et al.*, 2018), suggesting that further research in this domain is necessary to reliably improve student outcomes using mindset interventions.

Universality Beliefs

Universality beliefs describe students' beliefs about the distribution of the potential for high levels of ability (Rattan *et al.*, 2012b; Rattan *et al.*, 2018). The belief that everyone has the potential to reach the highest level of ability is called the universal belief. Individuals with universal beliefs acknowledge that different people may realize their potential to differing degrees depending on their life circumstances. In contrast, the belief that only some people have the potential to reach the highest levels of ability is called a nonuniversal belief. Those with non-universal beliefs believe that only some people are capable of attaining the highest levels of ability, no matter what resources they have access to or how much effort they put in.

Rattan and colleagues (2012b) conceptualized and defined universality beliefs as a lay theory that is distinct from mindset beliefs. They have argued that it is possible, for example, to believe intelligence can be improved (growth mindset), but that different people hold different potentials for how much they can improve (nonuniversal belief). In a series of studies measuring both universality and mindset beliefs, Rattan and colleagues (2012b, 2018) have consistently found small to no correlations (r values $<|0.3|$), supporting the hypothesis that mindset and universality are separate constructs.

Universality beliefs influence undergraduates' educational experiences and outcomes as well. In one study, students who perceived that instructors held universal beliefs felt a greater sense of belonging, and increasing perception of faculty's universal belief caused an increased sense of belonging (Rattan *et al.*, 2018). In addition, an increased sense of belonging positively predicted improved academic performance for underrepresented groups (Rattan *et al.*, 2018). These results suggest that further research on universality beliefs could have implications for creating inclusive educational environments and achieving equity in academic outcomes.

Brilliance Beliefs

Brilliance beliefs describe the belief that a “raw,” innate intellectual talent (i.e., “brilliance”) is required for success (Leslie *et al.*, 2015; Meyer *et al.*, 2015)¹. The brilliance belief was conceptualized as distinct from universality and mindset in that it focuses on what is required for success in a field, rather than about an individual's potential for growth (mindset) or the distribution of that potential (universality). However, no prior work has directly compared universality and brilliance beliefs. There is some evidence that brilliance beliefs are distinct from mindset. One study found that changing mindset messages did not affect the relationship between brilliance and outcomes, suggesting that they are orthogonal constructs (Bian *et al.*, 2018b). Another

¹Brilliance was initially conceptualized to refer to innate talent, but more recent theorizing has been more agnostic about whether talent is innate or improvable. See Porter & Cimpian 2023 for more discussion on this.

study measured both growth mindset and brilliance (also called Field-Specific Ability Beliefs), and found a moderate correlation ($r = -0.53, p < 0.001$) that was not so high as to suggest the two constructs are synonymous (Porter and Cimpian, 2023). However, it remains unclear whether the belief that brilliance is required for success is conceptually distinct and empirically discriminable from universality beliefs.

Research suggests that brilliance beliefs cause representation gaps across fields. Underrepresentation of women and racially/ethnically marginalized groups is more severe in fields that are generally perceived to require brilliance (Leslie *et al.*, 2015; Meyer *et al.*, 2015). This effect is driven by culturally pervasive stereotypes of white men being associated with brilliance and women and people of color being stereotyped as having lower intellectual ability (Leslie *et al.*, 2015; Meyer *et al.*, 2015; Bian *et al.*, 2018b; Storage *et al.*, 2020). This interaction results in discrimination (Bian *et al.*, 2018a), and hostile or cold climates (Vial *et al.*, 2022) that foster greater imposter feelings among women and minoritized students (Muradoglu *et al.*, 2021). Thus, further research on brilliance beliefs may be an important step forward in addressing widespread inequities and underrepresentation across fields.

The Present Study

This work is guided by two research questions. First, we sought to evaluate whether these three lay theories are susceptible to the jangle fallacy.

Research Question 1: Do mindset, universality, and brilliance beliefs represent conceptually and empirically distinct constructs?

Based on prior work, we expected that mindset and universality would be distinct (Rattan *et al.*, 2012b), as would mindset and brilliance (Bian *et al.*, 2018b). However, we were unsure whether brilliance beliefs would be sufficiently distinct from universality beliefs to be empirically discriminable.

Our second goal was to evaluate the extent to which any distinctions we identify are practically meaningful. Conceptual and empirical distinctions among these lay theories would be meaningless if one completely subsumes all explanatory power of another.

Research Question 2: Do mindset, universality, and brilliance beliefs show unique explanatory value for relevant psychosocial and academic outcomes?

To address this research question, we selected four psychosocial factors (sense of belonging, evaluative concern, goal orientation, and self-handicapping) and two academic outcomes (intent to persist in Science and grades) that prior research indicates relate to one or more of the focal lay theories.

Addressing our research questions required high-quality (i.e., supported by strong evidence of validity) measurement tools. Thus, we also developed a new measure of undergraduates' lay theories of abilities, the ULTrA survey, which we report here.

METHODS AND RESULTS

Due to the iterative and developmental nature of this work, we present methods and results together for each of the four phases of this work. We addressed our research questions and developed the ULTrA Survey over four phases: 1) evaluation of response processes; 2) evaluation of content validity;

3) evaluation of internal structure of the ULTrA measure, including the extent to which mindset, universality, and brilliance are distinguishable; and (4) evaluation of the relationships between responses on the ULTrA and relevant outcomes. Each phase roughly corresponds to one type of evidence of validity from the Standards for Educational and Psychological Testing framework (AERA *et al.*, 2014).

All methods were reviewed and approved by the Institutional Review Board at the lead researcher's home institutions (PROJECT00000858 & IRB2021-514) and at institutions where recruitment took place, as applicable.

Data were analyzed using R, version 4.1.2 (R Core Team, 2021) including functions in packages psych (Revelle, 2022), lavaan (Rosseel, 2012), and mirt (Chalmers, 2012). The datasets generated and analyzed during the current study are available in the Open Science Framework repository: https://osf.io/beayz/?view_only=fac206f383c547069c05ce793be5c0e5.

Measurement of Lay Theories

Mindset beliefs are typically measured using the Implicit Theories of Intelligence Scale, which consists of eight items: four growth and four fixed (Dweck, 1999; Cook *et al.*, 2017; Sisk *et al.*, 2018). The wording of the items are context-generic, and evidence suggests that this creates issues in the context of undergraduate education (Limeri *et al.*, 2020b; Sun *et al.*, 2021). Specifically, undergraduates interpret "intelligence," a key referent in every item, unreliably. This introduces undesirable measurement error and poses a threat to validity (Crooks *et al.*, 1996).

Brilliance beliefs and universality beliefs are both relatively newer constructs. Thus, there has not been time or theoretical impetus to invest in developing high-quality measurement tools for undergraduate educational contexts. Universality beliefs have been measured in a variety of ways: a single, bipolar item; a set of items with a positively skewed rating scale; and forced choice between pairs of universal and nonuniversal statements (Rattan *et al.*, 2012b). The researchers who conceptualized brilliance beliefs wrote a set of four items (e.g., "Being a top scholar of [discipline] requires a special aptitude that just can't be taught.") with a 7-point agreement response scale (Leslie *et al.*, 2015).

Out of necessity, these scales were written by the researchers who conceptualized the ideas to establish proof of concept. These seminal studies have provided evidence that these beliefs influence relevant outcomes (e.g., representation across fields, attitude towards educational policies, sense of belonging). Now there is a need for careful, rigorous development of high-quality measures of these concepts to enable advancement in this field and to examine the conceptual space occupied by mindset, universality, and brilliance.

To guide the measurement development process, we relied on the framework established in Standards for Educational and Psychological Testing (AERA *et al.*, 2014). Standards describes four types of evidence of validity, emphasizing that all types are equally important. This approach follows Crooks and Kane's metaphor of validity as chain which is only as strong as its weakest link (Crooks *et al.*, 1996). Throughout the four phases of this project, we collected data relating to each of the four types of validity evidence in the Standards framework: 1) evidence based on response processes, 2) content,

3) internal structure, and 4) relations to other variables. While there are multiple sources of evidence of validity, validity is a unitary concept that refers to the strength of evidence supporting proposed interpretations of survey responses (Kane, 1992; Messick, 1995).

Phase 1: Evidence Based on Response Processes

In the first phase, we collected evidence that undergraduates interpret the items as intended (i.e., evidence based on response processes). We collected this evidence through two rounds of interviews. First, we conducted semistructured interviews to identify language that could be used in the items that undergraduates would interpret consistently. We drafted a set of 50 items initially intended to measure mindset based on these results. Then, we conducted cognitive interviews to check that students interpreted items as intended. We revised items based on these results.

Semistructured Interviews. We began by conducting semistructured interviews with undergraduates to identify language describing intellectual abilities that undergraduates would interpret more consistently than “intelligence” (Limeri *et al.*, 2020b). We interviewed 45 students enrolled in introductory Science and Math courses at 14 institutions. Our sample was diverse in terms of gender (26 women, 18 men, one nonbinary individual), race/ethnicity (14 white, 14 African American or Black, 10 Hispanic or Latinx, five South or Southeast Asian, three East Asian, two Native American or Alaskan Native, one Middle Eastern or North African, and one Native Hawaiian or Pacific Islander), courses (31 in Chemistry, 28 in Biology, 26 in Math, 12 in Physics) and institution type (five very high research activity, two high research activity, one doctoral, one masters granting, two baccalaureate, and three community colleges; four Historically Black Colleges and Universities, two Hispanic-Serving Institutions, and eight primarily white institutions). See Supplemental Material section 1 for additional details about data collection and analysis methods.

We identified three terms that reflected how undergraduates thought about intelligence and that were interpreted consistently: 1) analyzing information, 2) applying knowledge, and 3) ability to learn. We drafted an initial pool of 50 survey items based on these terms and on existing literature defining the lay theories. We also reviewed the items in existing measures and in some instances adapted wording or ideas (e.g., our initial brilliance items were very close to the items written by Leslie, Cimpian, Meyer, and Freeland, 2015).

Cognitive Interviews. We then conducted cognitive interviews to gather further evidence of response process validity and refine our items. Specifically, we asked participants to read survey items and explain aloud how they interpreted the item and their reasoning for their response (Desimone and Le Floch, 2004). We interviewed 29 undergraduates from 11 institutions who were diverse in terms of gender (19 women, 9 men, 1 nonbinary individual), race/ethnicity (12 white, five African American or Black, five East Asian, four South or Southeast Asian, four Hispanic or Latinx, and three Middle Eastern or North African), courses (19 in Chemistry, 16 in Biology, 14 in Math, and six in Physics), and institution type (five very high research activity, one master’s granting, three baccalaureate, and two community

colleges; one Historically Black College, two Hispanic-Serving Institutions, and eight primarily white institutions).

Based on our results, we revised the items further to ensure consistent interpretation. For instance, participants described “intellectual ability” consistently as a broad ability that encompassed applying knowledge and analyzing information. Thus, we used “intellectual ability” as a key term in the items. We made numerous other small word changes based on students’ feedback (e.g., “natural talent” rather than “genius,” “learning effectively” rather than “learning efficiently”).

Phase 2: Evidence Based on Content

In the second phase, we collected evidence that the content of the items faithfully represent the lay theories they are intended to measure (i.e., evidence based on content). We collected this evidence by asking experts in lay theories to review our items and provide feedback about their alignment with theory.

Expert feedback. We asked 11 experts in lay theories and measurement development to review and provide feedback on our draft item set. Experts conducted a Q-sort activity (Nahm *et al.*, 2002) in which they read all 50 draft items, assigned each item to a lay theory, and rated how well the item fit the theory. We also asked experts to comment on whether there were any aspects of lay theories that were not captured by the items. We revised all items for which reviewers’ comments indicated an issue or for which there was less than 75% agreement (i.e., whether four of 11 experts categorized the item counter to our expectations). We consulted the experts to ensure our revisions addressed the issues and that each item aligned with one and only one lay theory. At the end of this process, we had a draft set of 50 items: 21 mindset (11 fixed, 10 growth), 23 universalism (11 universal, 12 nonuniversal), and six brilliance.

Phase 3: Addressing Research Question 1 and Evaluating Internal Structure of the ULTrA

Next we sought to determine whether mindset, universalism, and brilliance beliefs were empirically distinguishable and evaluate the internal structure of the ULTrA. Specifically, we evaluated competing hypotheses that mindset, universalism, and brilliance represent distinct or overlapping constructs by testing alternative confirmatory factor models on responses to the draft item set from a large national sample of undergraduate students in the United States. We then conducted further analyses to investigate the quality and functioning of the items. We ensured that items functioned equivalently across student groups by conducting differential item functioning and measurement invariance analyses. We then also estimated item response theory models to examine the functioning of each item. Based on all these considerations, we selected a subset of items with the best properties that captured the full conceptual range of each construct to create a recommended reduced item set for the ULTrA survey.

Data Collection. We recruited participants by asking instructors of introductory Biology, Chemistry, Physics, and Math courses at 20 institutions to send the survey to their students. We also asked participants to share the study with any peers enrolled in introductory Science or Math courses at any US institution. We selected institutions with diverse characteristics and with diverse student populations so that our sample would

TABLE 1. Demographic information for survey participants in phase 3. Demographic information for survey participants is presented alongside each category's representation in the population of undergraduates in the United States (from NCES data) and chi-square tests comparing the representation in the study sample to national representation. Participants came from 68 institutions with an average of 18 respondents per institution ($SD = 31$). Note: totals do not sum to 100% for racial/ethnic identity and course because participants were able to select more than one racial/ethnic identity and indicate all courses in which they were enrolled and in both cases were counted in all options they selected. In total, 152 (13%) respondents selected more than one racial/ethnic identity.

Descriptor	Number of respondents (% of respondents)	Representation in the population of undergraduates in the US	χ^2 test of sample deviation from national representation
Gender	1183 participants reported their gender identity		
Man	399 (34%)	43.5%	$\chi^2 = 39.4, df = 1, p < 0.001$
Woman	763 (64%)	56.5%	
Nonbinary or transgender	21 (1.8%): 20 nonbinary, 1 transgender	Not available	
Racial/Ethnic identity	1171 participants reported their racial/ethnic identity		
White	602 (51%)	52%	$\chi^2 = 0.168, df = 1, p = 0.682$
Hispanic or Latin(x)	242 (21%)	20%	$\chi^2 = 0.342, df = 1, p = 0.559$
Black or African American	118 (10%)	13%	$\chi^2 = 8.74, df = 1, p = 0.003$
Middle Eastern or North African	30 (2.6%)	Not Available	
East Asian	129 (11%)	7%	$\chi^2 = 504, df = 1, p < 0.001$
South Asian	85 (7.3%)		
Southeast Asian	64 (5.5%)		
Native American or Alaskan Native	35 (3.0%)	0.72%	$\chi^2 = 91.8, df = 1, p < 0.001$
Native Hawaiian or other Pacific Islander	26 (2.2%)	0.27%	$\chi^2 = 121, df = 1, p < 0.001$
Generation in college	1170 participants reported their generation status		
First generation	344 (29%)	Not available	
Continuing generation	826 (71%)		
Language	1190 participants reported their language status		
English is first language	997 (84%)	Not available	
English is not first language	193 (16%)		
Institution ownership	Reported by all participants		
Public	957 (80%)	74%	$\chi^2 = 23.2, df = 1, p < 0.001$
Private	237 (20%)	26%	
Research activity (Carnegie Classification)	Reported by all participants		
Very High Research Activity	398 (33%)	18%	$\chi^2 = 190, df = 1, p < 0.001$
High Research Activity	50 (4%)	9%	$\chi^2 = 33.4, df = 1, p < 0.001$
Doctoral/Research	117 (10%)	6%	$\chi^2 = 29.9, df = 1, p < 0.001$
Master's granting	231 (19%)	21%	$\chi^2 = 2.02, df = 1, p = 0.156$
Baccalaureate	149 (12%)	11%	$\chi^2 = 2.78, df = 1, p = 0.096$
2-yr institution	249 (21%)	29%	$\chi^2 = 38.3, df = 1, p < 0.001$

Note: When available from the National Center for Education Statistics, the national representation for undergraduates in the United States is presented in the middle column for comparison (National Center for Education Statistics, 2019). For most demographic categories, we were able to obtain a close approximation. Notable exceptions are that men are underrepresented in our sample compared with the national population by 10% and students enrolled at 2-yr institutions are underrepresented by 8%.

represent the population of undergraduates in the United States (Table 1). Participants were compensated with a \$10 gift card. We received 1522 complete survey responses and screened out duplicate responses and responses that failed the attention checks, resulting in a final sample of 1194 participants. See more details about data collection and screening in Supplemental Material section 2.

Confirmatory Factor Analysis Methods. We tested a series of competing models using confirmatory factor analysis (CFA)

to evaluate which model best explains the data. Our first model represented the hypothesis that there is complete conceptual overlap among the lay theories: a 1-factor model where all 50 items loaded onto a single factor. Our second model tested the hypothesis that there is conceptual overlap between brilliance and universality, where mindset is distinct. This hypothesis was based on the conceptual similarity between brilliance and universality, which both center on beliefs about whether some individuals have more potential than others. This 2-factor model has all mindset items loading

on one factor and all universality and brilliance items loading onto the second factor. Our third model represented the hypothesis that all three lay theories are conceptually distinct: a 3-factor model where all items for 1) mindset, 2) universality, and 3) brilliance load onto their respective latent factor. We then tested a model to evaluate the hypothesis that growth and fixed represent separate latent factors within mindset and universal and nonuniversal represent separate latent factors within universality. We retained the relationship between the mindset dimensions and the universality dimensions by including higher order latent factors. This model included 5 first-order factors: 1) growth, 2) fixed, 3) universal, 4) non-universal, and 5) brilliance, as well as two higher order factors: 1) mindset (consisting of growth and fixed latent factors) and 2) universality (consisting of universal and nonuniversal latent factors). Finally, we fit one final model to thoroughly check for conceptual overlap between brilliance and the non-universal belief specifically. The 2-factor model collapses all of universality with brilliance, but the brilliance belief most closely resembles nonuniversality specifically. Universal belief does not fit well with brilliance beliefs because it is the belief that all individuals have the same potential, whereas brilliance requires believing some individuals have more potential than others. Thus, brilliance may have overlap with nonuniversal belief but not the universal belief. To test this hypothesis, we fit a four first-order, two second-order factor model where growth, fixed, and nonuniversal items loaded onto their respective first-level latent factors, and nonuniversal and brilliance items loaded onto a fourth first-order latent factor. Growth and fixed loaded on a second-order latent factor (mindset) and universality and nonuniversality/brilliance loaded onto a second-order factor (universality).

CFA models were fit with the `cfa()` function in the `lavaan` package using robust maximum likelihood estimation. We evaluated whether a nested structure was necessary by examining Intraclass Correlation Coefficients for all items and found that a nested structure was not required to model our data².

CFA Results (Research Question 1). Testing alternative models with confirmatory factor analyses allowed us to address research question 1, evaluating the extent of conceptual and empirical overlap among the lay theories. The model data fit metrics are presented in Table 2. The model representing all lay theories as conceptually distinct was the best fit to the data. All models representing conceptual overlap among the lay theories exhibited worse fit to the data, as judged by observing changes in fit metrics. For more details about model estimation methods and fit interpretation, see Supplemental Material section 2.

²The mean ICC ranged from 0.006 to 0.070, with a mean of 0.030 and median of 0.028. For most (92%) of the items, $ICC \leq 0.050$ (and 88% were below), which is a common cutoff for ICCs in determining the appropriateness of multilevel analyses. Further, attempts to fit even single-factor models (e.g., Brilliance alone) did not converge even with adjustments to the criterion, which can happen when there is insufficient between-cluster variability relative to within-cluster variability. We also fit our final model with the Huber-White “Sandwich” estimator to account for clustering in the standard errors. Doing so resulted in identical model-data fit as the restricted maximum likelihood solution but also resulted in a not-positive definite matrix in the parameter estimates (likely due to a lack of between-cluster variance relative to within-cluster variability).

We additionally assessed the fit of the best fitting model (five first-order factors and two second-order factors) under the framework of equivalence testing, which differs from null-hypothesis significance testing by aiming to endorse a model under a null hypothesis rather than reject it (Yuan *et al.*, 2016). To test equivalence, we followed methods proposed by Yuan *et al.* (2016) to calculate the T-size adjusted RMSEA of our best model, 0.063, which indicated acceptable but not great fit.

Because fit metrics indicated less than ideal fit, we further assessed the fit of the model by examining the standardized residual covariances to identify misspecification. There were five item pairs with residual covariances greater than $|0.20|$. We examined these item pairs and deleted three items that had substantial residual covariances with other items to attempt to improve model fit. Further details about these items and the decisions about item deletion are provided in Supplemental Material Section 3. After removing these items, we refit the CFA and found improved model fit: RMSEA = 0.055, t-size adjusted RMSEA = 0.057, SRMSR = 0.063, CFI = 0.89, and TLI = 0.89.

Research Question 1 Results. The best supported model indicates that mindset, universality, and brilliance are distinct and empirically discriminable constructs.

Testing for potential measurement bias.

We tested for potential group level bias by conducting measurement invariance and differential item functioning analyses. Both analyses suggested that the measure functions equivalently across gender, race/ethnicity, first generation status, disability status, first language, institution type, and course enrolled (i.e., Biology, Chemistry, Math, or Physics). See Supplemental Material section 4 for more details on measurement invariance analysis methods and results and Supplemental Material section 5 for more details on differential item functioning analyses and results.

Recommended short form of the ULTrA. We conducted further analyses to examine each item’s psychometric properties so we could select a subset of items that could measure these lay theories without undue participant burden. We estimated item response theory models to guide decisions about which items to select as a short form of the ULTrA. We estimated a Graded Response Model for each first-order factor using the `mirt()` function within the `mirt` package. We evaluated item quality and functioning by inspecting alpha (discrimination) and beta (item difficulty) parameters and item information curves (see Supplemental Material section 6 for more information). Item information curves show how much “information” an item provides at different points along the latent trait. Curves are highest at the points where the item has the highest ability to discriminate. Thus, a set of items that provides the highest “information” across the largest range of the latent trait (θ) is ideal. Item information curves for all of the draft items as well as the items selected for the short form of the measure are presented in Figure 1. Details on the methods of the IRT analyses, fit metrics, and alpha and beta parameter values are presented in Supplemental Material section 6.

To generate the recommended short form of the ULTrA measure (Table 3), we selected five items for each factor (25 total) based the primary goals of maximizing: the range of item

TABLE 2. Model fit statistics for confirmatory factor models.

Model	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR	Chi ²	df
1 factor	0.540	0.520	0.101	0.100–0.102	0.108	14402	1175
2 factors	0.653	0.638	0.088	0.087–0.089	0.092	11199	1174
3 factors	0.726	0.713	0.078	0.077–0.080	0.093	9125	1172
five first-order, two second-order factors	0.858	0.851	0.056	0.055–0.058	0.067	5288	1168
four first-order, two second-order factors	0.827	0.819	0.062	0.061–0.064	0.065	6184	1170

difficulty (beta parameter), the discrimination of the items (alpha parameter), the test information (sum of item information curves), and conceptual range (first author's judgments). In selecting the items, the first author carefully considered the conceptual range of the shortform items to ensure they covered the full conceptual range of each construct and avoid construct underrepresentation.

Phase 4: Addressing Research Question 2 and Evaluating Evidence Based on Relations to Other Variables

After determining that mindset, universality, and brilliance are distinct constructs, we sought to explore how they relate to psychosocial and academic outcomes. In the process, we collected validity evidence for the ULTrA Survey based on relations to other variables. To evaluate this, we surveyed a second national sample of Science and Math undergraduates with measures of theoretically related constructs and outcomes. Specifically, we selected as psychosocial outcomes, sense of belonging, evaluative concern, achievement goals, and self-handicapping. We selected as academic outcomes course grades and intent to persist in Science. The rationales for these selections are described next and the hypothesized relationships are summarized in Table 5.

Sense of Belonging. Sense of belonging refers to the universal human need to have a sense of affiliation and identification in their environment (Hoffman *et al.*, 2002). A sense of belonging is important and relevant in many contexts, and is a particularly important factor for undergraduate retention in college (Tinto, 1988; Hoffman *et al.*, 2002). Universality and brilliance beliefs both center around who has the capacity to be successful in Science and Math and thus who belongs in these fields. Those who hold nonuniversal and brilliance beliefs believe that only some people have the capacity to succeed, and thus belong, in Science and Math. Rattan and colleagues (2018) conducted a series of studies investigating meta-universality beliefs, or what students believe that others believe about universality. They found with remarkable consistency that meta-universal beliefs predicted greater sense of belonging (Rattan *et al.*, 2018). Based on this theoretical basis and related empirical evidence, we propose that nonuniversal and brilliance beliefs will be negatively related to sense of belonging. In contrast, universal beliefs posit that everyone is capable of achieving high levels of intellectual ability, and thus these pose no barriers to who belongs. Therefore, we hypothesize that universal belief should be positively related to sense of belonging. Mindsets are theoretically related to belonging only indirectly, and there is little empirical work examining this link. One recent study with undergraduate Engineering students found that a growth mindset predicted a greater sense of belonging because it shaped how students perceived their interactions with their peers and instructors

(Williams *et al.*, 2021). Therefore, we predict weak but significant relationships between sense of belonging and growth mindset (positive) and fixed mindset (negative).

Evaluative Concerns. Evaluative concern is a form of psychological threat in which students are worried about saying or doing something wrong in class and being negatively evaluated (Muenks *et al.*, 2020). While there is less empirical research directly connecting lay theories and evaluative concerns, evaluative concern is the most theoretically closely related outcome from these lay theories. Because students with a fixed mindset believe that their abilities are an inherent part of them, they should be more concerned about being negatively evaluated based on their ability if they make a mistake in class. Conversely, students with a growth mindset believe that their abilities can be improved, so should be less concerned about being judged on the basis of mistakes. At least one study suggests that students who perceive their instructor holds a growth mindset experience more evaluative concern (Muenks *et al.*, 2020). Similarly, students who hold brilliance and/or nonuniversal beliefs are likely to be concerned with being judged as one of the students who “has what it takes.” However, students with a universal belief would not hold this concern because they don’t believe that there are students who don’t have the potential.

Achievement Goal Orientation. Achievement goal orientation refers to the types of goals that students prioritize in relation to their academic endeavors (Elliott and Dweck, 1988). Achievement goals are defined by a 2 × 2 framework: performance versus mastery and approach versus avoid. Performance goals focus on demonstrating competence relative to others, such as aiming to receive a high grade whereas mastery goals focus on learning and improvement (Elliot, 1999; Pintrich, 2000; Elliot and McGregor, 2001). Both performance and mastery goals can be further defined as either approach, striving toward a goal, or avoid, striving to avoid failing (Elliot and McGregor, 2001). For example, a performance-approach goal would be to earn an A whereas a performance-avoid goal would be to avoid appearing unintelligent. A mastery-approach goal would be to gain a new skill whereas a mastery-avoid goal would be to avoid holding misconceptions or forgetting material.

Theory and prior research suggests that students with different mindset beliefs prioritize different achievement goals (Burnette *et al.*, 2013). For mastery goals, theory and evidence suggests that mastery-approach goals should be positively correlated with growth and negative correlated with fixed beliefs. Students who believe they can improve their mastery will be motivated to do so, whereas students who do not believe it is possible will not endorse mastery goals. This pattern has been repeatedly empirically observed (Burnette *et al.*, 2013; De Castella and Byrne, 2015; Smiley *et al.*, 2016; Cook *et al.*, 2017;

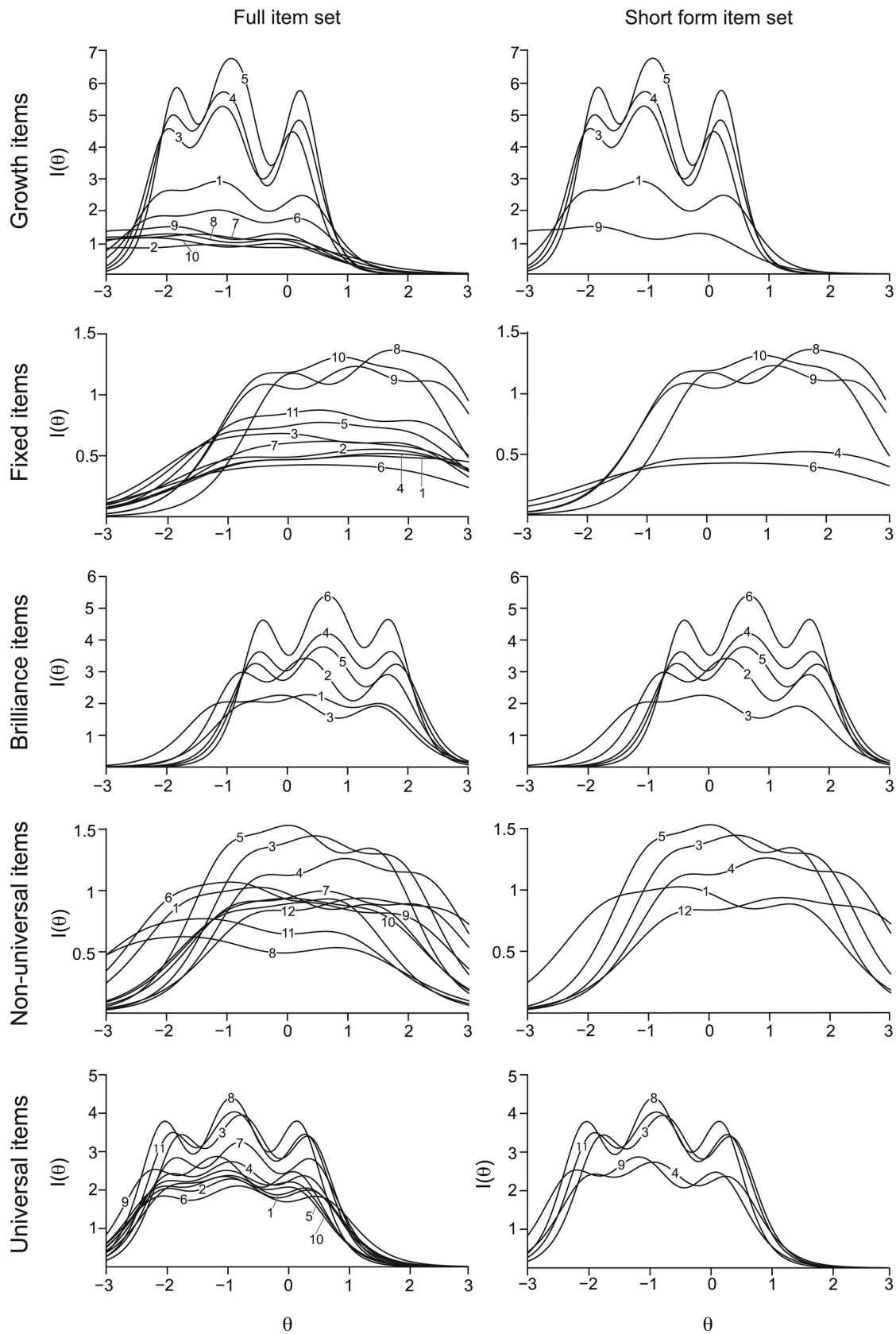


FIGURE 1. Item information curves. Information curves for the full item set (left) and items selected for the recommended short form of the measure (right).

TABLE 3. Items in the recommended shortform of the ULTrA Survey. See the full item set in Supplemental Materials Section 6, Supplemental Table 3.

Dimension	Item
Growth mindset	<ol style="list-style-type: none"> 1. I can become as good at analyzing information as highly successful STEM professionals if I try hard enough. 2. If I want to, I can become as effective at applying knowledge as STEM experts. 3. I can become excellent at applying knowledge to solve challenging problems. 4. If I try, I can become as effective at learning as STEM experts. 5. I could improve my intellectual abilities to the same level as successful STEM professionals.
Fixed mindset	<ol style="list-style-type: none"> 1. At the end of college, my ability to analyze information will be at about the same level that it is now. 2. How well I learn is something that I cannot change very much. 3. My ability to apply knowledge will change very little over time. 4. I will never be able to reach the highest level of intellectual ability. 5. It would be very difficult for me to improve how well I can apply knowledge.
Brilliance belief	<ol style="list-style-type: none"> 1. Excelling in STEM requires natural talent. 2. People who are highly successful in STEM have a natural talent for it. 3. Becoming a top student in STEM requires an innate talent that just can't be taught. 4. People have to be naturally talented to excel in challenging STEM courses. 5. Being a highly successful STEM professional requires natural talent that just can't be taught.
Universal belief	<ol style="list-style-type: none"> 1. With enough hard work, anyone could become as good at analyzing information as highly successful STEM professionals. 2. Anyone who tries could become as good at applying knowledge as STEM experts. 3. Anyone could become as effective at learning as highly successful STEM students. 4. Everyone has the intellectual ability to become a successful STEM professional if they want to. 5. With enough motivation, anyone can become as good at applying knowledge as high achieving STEM students.
Nonuniversal belief	<ol style="list-style-type: none"> 1. Even if they try, some people could never become as effective at analyzing information as their peers. 2. Only people with a natural talent can become good enough at applying knowledge to solve the most difficult problems. 3. Only people with a natural talent can become excellent at analyzing information. 4. Some people will always be less effective at learning than those who have a natural talent for it. 5. Only some people have the intellectual ability to become a successful STEM professional.

The prompt presented before the items is: Please indicate the extent to which you agree or disagree with the following statements. There are no correct answers, we want to understand how you think about these ideas. Note that STEM stands for Science, Technology, Engineering, and Mathematics. STEM professionals are individuals in a career in a STEM field, such as scientists, engineers, medical doctors, and other healthcare professionals.

Yan and Wang, 2021). Conversely, it is theoretically and empirically unclear how mastery-avoid goals should relate to mindset, and is an ongoing area of theoretical development (Burnette *et al.*, 2013; Cook *et al.*, 2017; Henry *et al.*, 2019).

For performance goals, theory, and evidence suggests that performance-avoid goals should be positively related with fixed and negatively related with growth mindsets. Students with a fixed mindset should hold particularly strong performance-avoid goals because they interpret failure as an indictment of low, fixed ability, which poses a threat to their self-concept whereas those with a growth mindset do not find this as threatening because they believe they could improve. This pattern has also been repeatedly empirically observed (Burnette *et al.*, 2013; Smiley *et al.*, 2016; Cook *et al.*, 2017; Yan and Wang, 2021). Conversely, students with both mindsets may be strongly motivated to achieve performance goals, which hold practical importance for obtaining jobs and admittance to postgraduate programs (Dweck and Leggett, 1988; Burnette *et al.*, 2013).

We are not aware of prior studies relating goal orientations to universality or brilliance beliefs. Therefore, we cannot make *a priori* predictions.

Self-Handicapping. Self-handicapping is a maladaptive behavior in which students sabotage their own performance so that failure can be blamed on the self-imposed obstacle rather

than their ability (Jones and Berglas, 1978). Examples of self-handicapping include spending insufficient time studying, staying up late before an exam, or procrastinating so that there is not enough time to complete an assignment well. Self-handicapping serves to protect a student's self-esteem in the face of a threatening achievement context because the failure could be blamed on the handicap and not the student's low ability (Jones and Berglas, 1978; Henry *et al.*, 2019). Thus, students with a fixed mindset, who are prone to view failure as a threat to their self-esteem, should be more likely to engage in self-handicapping behavior than students with a growth mindset, who do not view failure as diagnostic of immutable ability. Indeed, a meta-analysis showed that "helpless strategies," such as self-handicapping, were positively associated with a fixed mindset and negatively associated with a growth mindset (Burnette *et al.*, 2020). We expect students with a fixed mindset to engage in more self-handicapping than students with a growth mindset. We are not aware of any research connecting self-handicapping to universality or brilliance beliefs, so we are not able to make *a priori* predictions.

Intent to Persist and Course Grade. Through the direct effects predicted above, we hypothesize that each of these lay theories will predict two key academic outcomes: 1) intent to persist and 2) course grades. That is, by increasing sense of belonging

TABLE 4. Demographic information for phase 4 participants.

Descriptor	Number of respondents (% of respondents)
Gender	1365 participants (98%) reported their gender identity
Man	504 (37%)
Woman	859 (63%)
Nonbinary or transgender	two (0%): one agender, one gender fluid
Racial/Ethnic identity	1382 participants (99%) reported their racial/ethnic identity
White	908 (66%)
Hispanic or Latin(x)	154 (11%)
Black or African American	141 (10%)
Middle Eastern or North African	32 (2.3%)
East Asian	102 (7.4%)
South Asian	144 (10%)
Southeast Asian	34 (2.5%)
Native American or Alaskan Native	16 (1.2%)
Native Hawaiian or other Pacific Islander	seven (0.5%)
Generation in college	1387 participants (99%) reported their generation status
First generation	286 (21%)
Continuing generation	1101 (79%)
Institution type Carnegie Classification	Reported by all participants
Very High Research Activity	1221 (87%)
High Research Activity	0
Doctoral/Research	0
Master's granting	174 (12%)
Baccalaureate	0
2-yr institution	one (0%)

and mastery-approach goals, and decreasing performance-avoid goals, self-handicapping and evaluative concern, growth mindset and universal beliefs should promote intent to persist in Science and higher course grades. Conversely, fixed mindset, non-universal, and brilliance beliefs should result in lower intent to persist in Science and lower course grades. These predictions are based both on these indirectly theoretical links as well as much empirical evidence supporting these outcomes (e.g., Yeager *et al.*, 2016).

Data Collection Methods. We recruited participants by asking instructors of introductory Biology, Chemistry, Physics, and Math courses to send the survey to their students and through broad list-serve distributions at six institutions. We surveyed students during the fall 2021 semester and collected their Science/Math course grades from the instructor or institutional database after the semester ended. The full text of all items on the survey and a description of evidence of validity for the external measures used are available in Supplemental Material section 7. We collected participants' course grades in Science/Math courses in which they were recruited for the study from the course instructor after the semester ended. For a subset of our sample from one institution ($n = 299$), we accessed all of their Science and Math course grades from the institutional database and created the "course grade" variable by averaging all of their Science and Math course grades for the semester because we did not recruit them through a particular course. The methods for grade release were approved by a FERPA official at every institution where we collected data.

Participants. We received 1576 complete survey responses. Of these, 180 individuals failed the attention check questions and were removed from analyses, resulting in a sample of 1396. Demographic information of the sample is reported in Table 4. Participants were recruited from six institutions, with most participants coming from Very High Research Activity institutions (1221/1396; 87%). Five of the institutions were primarily white, one was an Hispanic-Serving Institution (299/1396; 21% of the sample).

Confirmatory Factor Analysis. First, we estimated a CFA model to reconfirm the factor structure we identified in phase three. Our model fit the data very well (CFI = 0.927; TLI = 0.918; RMSEA = 0.062; SRMR = 0.061, all factor loadings > 0.4), supporting our hypothesized factor structure and providing further confidence in our evidence of structural validity. The final factor model is presented in Figure 2.

Relationships Between Lay Theories and Other Variables. We examined correlations for each of our hypothesized relationships (Table 5; full correlation matrix in the Supplemental Table 6). Nearly every hypothesized relationship was observed as predicted (Table 5). Overall, growth and fixed mindset exhibited larger magnitude correlations with the psychosocial and academic outcomes than universality and brilliance. For mindset, the only observed relationship that did not match predictions was the nonsignificant and near-zero correlation between growth mindset and performance-avoid goals. It is not terribly surprising that even students with a growth mindset want to avoid negative

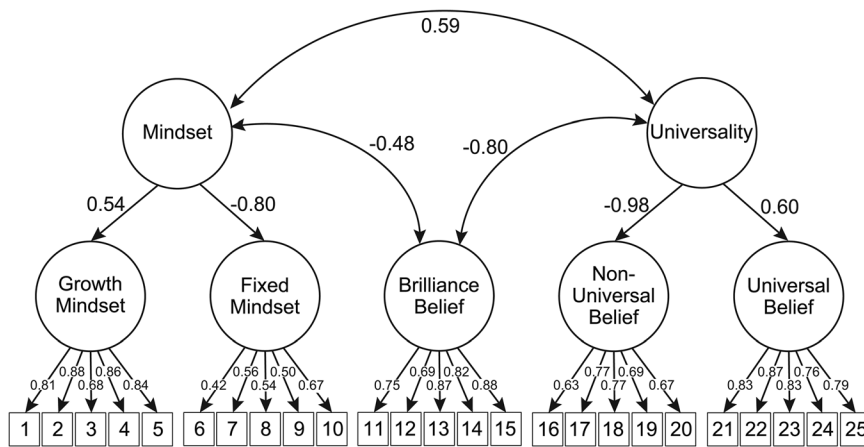


FIGURE 2. The factor model for the ULTrA measure. Squares represent observed variables (i.e., survey items). Circles represent first-order and second-order latent factors. Straight arrows represent factor loadings and curved arrows represent covariances.

performance outcomes, such as receiving a failing grade. These data constitute strong validity evidence based on relations to other variables for mindset.

Brilliance belief related as predicted to sense of belonging, evaluative concern, and intent to persist, but not course grade. Universal belief related as predicted to sense of belonging and intent to persist, but not evaluative concern or course grade. Nonuniversal belief related as predicted to sense of belonging, evaluative concern, and intent to persist, but not course grade. Accordingly, the overall universality belief related as expected to belonging and intent to persist, but not evaluative concern or course grade.

Considered together, universality and brilliance beliefs supported most, but not all of our hypothesized relationships with psychosocial and academic outcomes. This could be interpreted

as moderate validity evidence based on relations to other variables.

Structural Equation Modeling (SEM). We then examined the extent to which each lay theory uniquely related to each outcome by estimating a series of SEMs that controlled for other lay theories and socio-demographic variables. Our result that a second-order model is the best fitting model indicates that aggregating at both the first-order level and the second-order level is justified. Therefore, we estimated two sets of SEMs: one set using second-order factors (mindset, universality, and brilliance) as predictors, and the other set using the first-order factors (growth mindset, fixed mindset, universal belief, non-universal belief, and brilliance belief) as predictors. Each SEM included a target outcome (belonging, evaluative concern,

goal orientation, self-handicapping, intent to persist, and course grade) predicted by lay theories, gender, race/ethnic identity, and generation in college and the full measurement model for the relevant latent variables.

All of the SEMs using second-order factors as predictors fit acceptably or strongly (Table 6). Results of the SEMs using second-order factors as predictors are presented in Table 7. Due to the large number of analyses conducted, we adopted a more conservative critical value for interpreting significance of predictors ($p \leq 0.01$) to reduce risk of Type I error. Collectively, ULTrA second-order factors plus demographics predicted 13–25% of the variance in the outcomes, except for performance-avoid goals, which only reached 4% of the variance explained. All outcomes (except for performance-avoid goals) were significantly predicted by at least one of the lay theories

TABLE 5. Predicted and observed relationships between the lay theories and covariates/outcomes.

Covariates and outcomes	ULTrA dimensions						
	Growth mindset	Fixed mindset	Mindset	Brilliance belief	Universal belief	Non-universal belief	Universality
Sense of belonging	[+] 0.23***	[-] -0.28***	[+] 0.31***	[-] -0.11***	[+] 0.18***	[-] -0.16***	[+] 0.19***
Evaluative concerns	[-] -0.08**	[+] 0.20***	[-] -0.17***	[+] 0.13***	[-] 0.02 n.s.	[+] 0.08**	[-] -0.04 n.s.
Mastery-approach goals	[+] 0.27***	[-] -0.23***	[+] 0.30***	n.p.	n.p.	n.p.	n.p.
Performance-avoid goals	[-] -0.05 n.s.	[+] 0.11***	[-] -0.09***	n.p.	n.p.	n.p.	n.p.
Self-handicapping	[-] -0.15***	[+] 0.23***	[-] -0.23***	n.p.	n.p.	n.p.	n.p.
Intent to persist	[+] 0.28***	[-] -0.13***	[+] 0.25***	[-] -0.08**	[+] 0.08**	[-] -0.08**	[+] 0.09***
Course grade	[+] 0.15***	[-] -0.09**	[+] 0.14***	[-] -0.01 n.s.	[+] -0.06*	[-] 0.02 n.s.	[+] -0.05 n.s.

Each cell includes the predicted relationship direction (top line in brackets) and the observed Pearson correlation (bottom line). Observations that do not match predictions are emphasized in italics. Results are not shown for relationships that were not predicted a priori, but the full correlation table with these data are available in the Supplemental Table 6. n.p. = no prediction; n.s. = not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

TABLE 6. Model-data fit metrics for SEMs using second-order factors as predictors.

Model	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR	chi-sq	df
Belonging	0.891	0.885	0.050	0.049–0.052	0.071	5457	1359
Evaluative Concern	0.895	0.886	0.061	0.059–0.064	0.057	2807	482
Goal Orientation	0.932	0.926	0.049	0.046–0.051	0.051	2025	507
Self-Handicapping	0.922	0.915	0.049	0.047–0.052	0.052	2100	514
Persistence	0.922	0.915	0.052	0.050–0.055	0.053	1908	421
Course Grade	0.918	0.909	0.056	0.054–0.059	0.055	1611	365

with an effect size (standardized beta) greater than $|0.3|$ with a great degree of confidence that the effect is not a false positive (i.e., $p \leq 0.001$).

Sense of belonging, mastery-approach goals, and self-handicapping were each predicted by only mindset. In contrast to mastery-approach goals, performance-avoid goals were not predicted by any lay theory with a strong degree of confidence. Evaluative concern was predicted by all three lay theories: mindset, brilliance, and universality. Intent to persist and course grades were both predicted by both mindset and universality.

After controlling for gender, race/ethnicity, and generation, and the other lay theories, each lay theory contributed unique predictive power to at least one outcome.

- Mindset contributed unique predictive power to every outcome except for performance-avoid goals.
- Brilliance belief contributed unique predictive power to one outcome: evaluative concern.
- Universality contributed unique predictive power to evaluative concern, intent to persist, and course grade.

We additionally estimated SEMs that used first-order latent factors of the ULTrA (growth, fixed, universal, nonuniversal, and brilliance) as predictors. This approach enabled us to examine nuanced differences in the predictive efficacy of each lay theory dimension. Using first-order factors as predictors did not change model fit for the SEMs (Table 8).

Results using first-order factors as predictors are presented in Table 9. First-order factors tended to predict less variance in the outcomes than second-order factors, predicting 11–16% of the variance in the outcomes, except for performance-avoid goals, which only reached 6% of the variance explained. Accordingly, the effect size of the predictors (standardized betas) tended to be lower. The results did not change substantially: we continued to find that all outcomes were significantly predicted by at least one of the lay theories and that each of the lay theories contributed uniquely to predicting at least one outcome. We found intriguing differences in the predictive efficacy of growth compared with fixed and universal compared with nonuniversal.

Growth and fixed mindset each contributed unique predictive value to four outcomes, but these were not necessarily the same outcomes. Both growth and fixed contributed to sense of belonging and mastery-approach goals. However, growth, but not fixed, uniquely predicted the academic outcomes, intent to persist and course grade. In contrast, fixed, but not growth, uniquely predicted evaluative concern and self-handicapping.

Similarly, universal and nonuniversal beliefs contributed to predicting outcomes in different ways. They both uniquely pre-

dicted self-handicapping. However, universal belief also predicted four other outcomes that nonuniversal belief did not: 1) evaluative concern, 2) mastery-approach goal, 3) performance-avoid goal, and 4) course grade.

DISCUSSION

In this study, we explored the relationships among three lay theories that have been studied primarily in isolation of each other: mindset, universality, and brilliance beliefs. We sought to understand the extent of conceptual overlap among these three beliefs and how they relate to relevant student outcomes. Our data indicate that students' mindset, universality, and brilliance beliefs are conceptually distinct and empirically discriminable. Our results also demonstrate that each belief contributes unique predictive value for at least one outcome, and each outcome is predicted by a unique combination of beliefs. Our results suggest that examining the collection of beliefs should be useful for understanding students' psychosocial experiences and academic outcomes. Furthermore, we developed and collected extensive validity evidence for the ULTrA measure. We built this validity argument through four phases in which we collected evidence based on response processes, content, internal structure, and relations to other variables.

Conceptual Distinctions among Mindset, Universality, and Brilliance

Our results align with Rattan and colleagues' prior work (2012b) finding that mindset and universality are related but distinct. A student can believe that their intellectual abilities can be improved (growth mindset), but that some people have potential for more improvement than others (nonuniversal belief). Conversely, a student can believe that intellectual abilities cannot be changed (fixed mindset), but that everyone has sufficient intellectual abilities and there no is meaningful variation (universal belief).

Our results also corroborate prior research indicating that mindset and brilliance are conceptually distinct. Brilliance was initially defined as the belief that success requires raw talent that is fixed, or could not be taught (Leslie *et al.*, 2015; Meyer *et al.*, 2015). However, later work relaxed the assumption that malleability (i.e., mindset) was a key element of the brilliance belief. The original authors manipulated messages about the malleability of brilliance and found that this manipulation did not moderate the effect of brilliance messages on women's (vs. men's) interest and self-efficacy (Bian *et al.*, 2018b; experiment 6). It is possible for a student to believe that their intellectual abilities cannot change (i.e., fixed mindset), and that only those with higher innate abilities can become very successful

TABLE 7. Results of SEMs using second-order factors as predictors.

Lay theories		Sense of Belonging	Evaluative Concern	Mastery Approach Goals	Performance Avoid Goals	Self-Handicapping	Intent to Persist	Course Grade
Lay theories	Mindset	<i>b</i> (SE)	0.67*** (0.11)	0.81*** (0.13)	-0.34* (0.16)	-0.89*** (0.16)	1.11*** (0.19)	0.82*** (0.18)
		β	0.51***	0.54***	-0.13*	-0.37***	0.51***	0.31***
	Brilliance	<i>b</i> (SE)	0.11* (0.04)	0.11* (0.05)	0.15 (0.08)	-0.04 (0.07)	-0.12 (0.08)	-0.12 (0.08)
Universality		β	0.17*	0.15*	0.12	-0.04	-0.11	-0.10
		<i>b</i> (SE)	0.02 (0.10)	0.03 (0.12)	0.09 (0.17)	0.01 (0.14)	-0.61*** (0.19)	-0.60*** (0.18)
		β	0.02	0.02	0.04	0.00	-0.31***	-0.29***
Sociodemographics	Woman	<i>b</i> (SE)	0.04 (0.03)	0.10* (0.04)	0.20** (0.06)	0.03 (0.05)	0.26*** (0.06)	-0.16** (0.06)
		β	0.04	0.08*	0.09**	0.01	0.13***	-0.08**
	URM	<i>b</i> (SE)	-0.15*** (0.04)	-0.04 (0.05)	-0.02 (0.08)	0.08 (0.06)	-0.12 (0.07)	-0.23*** (0.07)
First-gen		β	-0.12***	-0.02	-0.01	0.04	-0.06	-0.10***
		<i>b</i> (SE)	-0.07 (0.04)	0.08 (0.05)	0.07 (0.08)	0.18** (0.06)	-0.22** (0.07)	-0.50*** (0.08)
		β	-0.05	0.05	0.03	0.08**	-0.10**	-0.20***
Variance predicted		$R^2 = 0.23$	$R^2 = 0.15$	$R^2 = 0.25$	$R^2 = 0.04$	$R^2 = 0.13$	$R^2 = 0.21$	$R^2 = 0.13$

Predictors significant at the $p \leq 0.01$ threshold are bolded for emphasis. *b* = unstandardized beta estimate, SE = standard error of the unstandardized beta estimate, β = standardized beta estimate. * $p \leq 0.05$; ** $p \leq 0.01$, *** $p \leq 0.001$.

(i.e., brilliance). However, it is also possible for a student to believe that their intellectual abilities can improve (i.e., growth mindset) and that reaching a high ability level is necessary to become successful (i.e., brilliance).

Our results suggest that universality and brilliance beliefs are also distinct, which to our knowledge has not been shown before. Universality centers on beliefs about the distribution for potential to reach high levels of ability whereas brilliance centers on beliefs about what it takes to be successful. Individuals with a universal belief assert that everyone has the potential to reach high ability levels, but acknowledge that individuals' circumstances may prevent that. Thus, someone could believe that all people hold equal potential (i.e., universal belief), but that only those whose circumstances allow them to realize high ability levels become successful (i.e., brilliance belief). Conversely someone could believe that only some people have the potential to reach high ability levels (i.e., nonuniversal belief) and that these are the only people with the potential to become successful (i.e., brilliance belief). An important distinction is that mindset and universality beliefs focus on individuals' abilities whereas brilliance focuses on beliefs about success in a field.

Relationship Between Growth/Fixed Mindset and Universal/Nonuniversal Beliefs

Our finding that a model including second-order factors is the best fit to the data implies that these lay theories can be conceptualized at both the first-order level (as growth and fixed factors separately) and at the second-order level (as an overall mindset score). In our SEMs, second-order factors predicted more variance in both psychosocial and academic outcomes than first-order factors. Using first-order factors as predictors revealed nuanced differences between growth/fixed and universal/nonuniversal beliefs. Thus, we recommend that researchers interested in understanding the extent of the relation between lay theories generally and outcomes use second-order factors as predictors, while researchers interested in exploring nuanced differences between different lay theories use first-order factors as predictors.

Growth and fixed mindsets are logical opposites of each other, as are universal and nonuniversal beliefs. Yet, the factor structure that best fit the data placed each of these in separate first-order factors, related by loading onto common second-order factors (Figure 2). This structure indicates that the mindset factors and the universality factors are not simply opposite ends of a spectrum. Rather, students can hold growth and fixed mindset beliefs and universal and nonuniversal beliefs simultaneously. Further, the raw correlations between growth/fixed and universal/nonuniversal are not as high as would be expected if they were opposite ends of a spectrum (growth and fixed $r = -0.3$; universal and nonuniversal $r = -0.5$; Supplemental Table 6).

Despite this evidence, it is puzzling how growth/fixed mindset and universal/nonuniversal beliefs could be conceptually distinct given that they are logical opposites of each other. We interpret these findings as the result of peoples' well-documented tendency to be able to hold conflicting beliefs at the same time (Festinger, 1962). This interpretation is further strengthened by the results that growth and fixed and universal and nonuniversal differed in how they related to outcomes and

TABLE 8. Model-data fit metrics for SEMs using first-order factors as predictors.

Model	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR	chi-sq	df
Belonging	0.891	0.885	0.050	0.049–0.052	0.071	5456	1357
Evaluative Concern	0.896	0.886	0.061	0.059–0.064	0.053	2789	480
Goal Orientation	0.934	0.927	0.048	0.046–0.050	0.049	1989	503
Self-Handicapping	0.923	0.916	0.049	0.047–0.051	0.051	2087	512
Persistence	0.924	0.916	0.052	0.050–0.054	0.054	1871	419
Course Grade	0.919	0.910	0.056	0.053–0.059	0.056	1600	363

that each contributed unique predictive power to some of the outcomes, even while controlling for all other beliefs (Table 9). For example, when all first-order ULTrA variables and demographics are included as predictors in SEMs, growth mindset contributes uniquely to predicting variance to intent to persist whereas fixed mindset does not; conversely, fixed mindset contributes uniquely to predicting evaluative concern whereas growth mindset does not (Table 9). Our results corroborate prior studies on other mindset measures which have concluded that growth and fixed mindsets are not opposite ends of a single spectrum, but rather distinct, moderately correlated constructs (Cook *et al.*, 2017; Scherer and Campos, 2022). The evidence that growth and fixed as well as universal and nonuniversal beliefs are distinct implies that it would be appropriate to retain both factors in future studies and consider how growth, fixed, universal, and nonuniversal beliefs all contribute unique explanatory value to relevant academic covariates and outcomes.

We acknowledge an alternative interpretation that responses to fixed mindset items correlated to each other more strongly and thus factored out separately as an artifact of peoples' well-documented tendency to prefer to agree to items and reluctance to disagree (Danner *et al.*, 2015). Measurement experts have long recognized that negatively worded items can often factor together as an artifact of being worded negatively rather than being driven by an underlying construct. However, if this factor structure were an artifact of measurement, we would not expect each factor to contribute uniquely to explaining variance in outcomes as we observe. Thus, in our judgment, it is more likely that our respondents genuinely hold conflicting ideas, which the ULTrA survey can detect.

Implications for Future Research in Lay Theories

Our results revealed that each lay theory contributes unique explanatory power towards understanding students' academic trajectories and outcomes, predicting unique variance to at least one outcome after controlling for students' other beliefs and sociodemographics. This finding corroborates prior work. Porter and Cimpian (2023, study 2) found that brilliance beliefs predicted intellectual humility, even when mindset beliefs were controlled for. Thus, brilliance beliefs explained variance in an outcome above and beyond what mindset could explain.

This suggests that future research seeking to understand the psychosocial factors that influence student outcomes should include all five of these lay theories. Therefore, future research involving mindset interventions in undergraduate educational contexts may benefit from measuring not only students' mindsets but also their universality and brilliance beliefs. This approach offers the potential to detect impacts of the intervention and increase power to predict outcomes.

Future research should further investigate which lay theories relate to which types of outcomes and explore the mechanisms that produce these patterns. For example, our results suggest that growth and fixed mindsets relate more strongly to mastery approach goals whereas universal belief relates more strongly to performance-avoid goals (Table 9). Future studies could investigate whether this difference is caused by the mastery versus performance dimension, or the approach versus avoid dimension, and why mindset and universality beliefs related differently to these different facets of goal orientation.

Another trend that could be explored is whether the positive or negative valence of beliefs influences which outcomes they relate to and how. For example, our data indicated that course grades were predicted by both positively valenced beliefs (growth and universal beliefs), but not the negatively valenced beliefs (fixed, nonuniversal, and brilliance; Table 9).

The ULTrA Survey

An important outcome of this work is the development of a high quality measure of mindset, universality, and brilliance beliefs for Math and Science undergraduates (Table 3). This tool makes it possible for future research to tease apart the effects of lay theories and address the future research directions proposed above. Based on our extensive validity evidence, we suggest the ULTrA may be more suitable for measuring lay theories among undergraduate Science and Math students than other measures that have been developed and tested in precollege or other populations.

Limitations

Our study and measurement development focused on undergraduate Science and Math majors in the United States. Thus, it is not clear whether the ULTrA survey would be a useful measure for studying lay theories among undergraduates in other disciplines or nationalities or with individuals at different levels of education (e.g., with faculty or secondary school students). Collecting evidence of validity for using this instrument to measure lay theories in these other contexts and with these other populations would be a fruitful avenue for further research.

CONCLUSION

Our results indicate that mindset, universality, and brilliance beliefs are distinct and empirically discriminable constructs and that each belief uniquely predicts salient undergraduate educational outcomes. We have developed and collected extensive validity evidence for the ULTrA survey, a concise (25 items total, five items per dimension) measure that future researchers can use to effectively and reliably measure American Science and Math undergraduates' lay theories.

TABLE 9. Results of SEMs using first-order factors as predictors.

Lay theories		Sense of Belonging	Evaluative Concern	Mastery Approach Goals	Performance Avoid Goals	Self-Handicapping	Intent to Persist	Course Grade
Lay theories	Growth	<i>b</i> (SE)	0.08** (0.03)	0.16*** (0.03)	-0.03 (0.05)	-0.09* (0.04)	0.42*** (0.05)	0.20*** (0.05)
		β	0.12**	0.19***	-0.02	-0.08*	0.32***	0.15***
	Fixed	<i>b</i> (SE)	-0.34*** (0.06)	-0.29*** (0.07)	0.22* (0.11)	0.46*** (0.10)	-0.03 (0.10)	-0.25* (0.12)
		β	-0.29***	-0.21***	0.09*	0.23***	-0.01	-0.10*
Brilliance		<i>b</i> (SE)	0.09* (0.04)	0.06 (0.04)	0.09 (0.07)	-0.06 (0.06)	-0.11 (0.07)	-0.07 (0.07)
		β	0.14*	0.08	0.07	-0.05	-0.09	-0.06
	Universal	<i>b</i> (SE)	0.03 (0.03)	0.11*** (0.03)	0.25*** (0.05)	0.16*** (0.04)	-0.05 (0.05)	-0.14** (0.05)
		β	0.05	0.14***	0.19***	0.15***	-0.04	-0.11**
Nonuniversal		<i>b</i> (SE)	-0.06 (0.05)	-0.03 (0.06)	0.21* (0.09)	0.21** (0.08)	0.08 (0.08)	0.10 (0.09)
		β	-0.08	-0.04	0.16*	0.19**	0.06	0.08
Sociodemographics	Woman	<i>b</i> (SE)	0.04 (0.03)	0.10* (0.04)	0.19** (0.06)	0.01 (0.05)	0.28*** (0.06)	-0.15* (0.06)
		β	0.04	0.08*	0.08**	0.01	0.14***	-0.07*
	URM	<i>b</i> (SE)	-0.15*** (0.04)	-0.05 (0.05)	-0.04 (0.08)	0.06 (0.06)	-0.14* (0.07)	-0.24*** (0.07)
		β	-0.12***	-0.03	-0.02	0.03	-0.06*	-0.10***
First-gen		<i>b</i> (SE)	-0.07 (0.04)	0.07 (0.05)	0.06 (0.08)	0.17** (0.06)	-0.24*** (0.07)	-0.50*** (0.08)
		β	-0.06	0.04	0.02	0.08**	-0.10***	-0.20***
Variance predicted		$R^2 = 0.16$	$R^2 = 0.14$	$R^2 = 0.16$	$R^2 = 0.06$	$R^2 = 0.12$	$R^2 = 0.14$	$R^2 = 0.11$

Predictors significant at the $p \leq 0.01$ threshold are bolded for emphasis. *b* = unstandardized beta estimate, SE = standard error of the unstandardized beta estimate, β = standardized beta estimate. * $p \leq 0.05$; ** $p \leq 0.01$, *** $p \leq 0.001$.

ACKNOWLEDGMENTS

We are extremely grateful to a large number of scholars who provided feedback that greatly strengthened this work. The authors of the theories this work was based on for their generosity in meeting with us and discussing ideas: Carol Dweck, Aneeta Rattan, and Andrei Cimpian. Experts who provided feedback on the draft items: Connie Barroso Garcia, Jeni Burnette, Michael Barger, Andrei Cimpian, Lisa Corwin, Caitlyn Jones, Sandhya Krishnan, Katie Kroeper, Paul O'Keefe, Katie Muenks, Heidi Williams, Mary Murphy, and Aneeta Rattan. Advisory board members who provided extensive feedback throughout the project: Debbi Bandalos, Bill Boone, Allan Cohen, and Mary Murphy. Finally, thank you to two anonymous reviewers and monitoring editor Ross Nehm who carefully reviewed earlier versions of this manuscript and provided extensive, helpful feedback. The work was supported in part by the National Science Foundation's Building Capacity in STEM Education Research Program (#1937684 & #2200485) and the Research Experiences for Undergraduates program (#2012362 & #1659423). Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation. No funding agencies had any role in the design of this study.

REFERENCES

- AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. Lanham, MD: American Educational Research Association.
- Bian, L., Leslie, S.-J., & Cimpian, A. (2018a). Evidence of bias against girls and women in contexts that emphasize intellectual ability. *American Psychologist*, 73(9), 1139–1153. <https://doi.org/10.1037/amp0000427>
- Bian, L., Leslie, S.-J., Murphy, M. C., & Cimpian, A. (2018b). Messages about brilliance undermine women's interest in educational and professional opportunities. *Journal of Experimental Social Psychology*, 76, 404–420. <https://doi.org/10.1016/j.jesp.2017.11.006>
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78(1), 246–263. <https://doi.org/10.1111/j.1467-8624.2007.00995.x>
- Burnette, J. L., Hoyt, C. L., Russell, V. M., Lawson, B., Dweck, C. S., & Finkel, E. (2020). A growth mind-set intervention improves interest but not academic performance in the field of computer science. *Social Psychological and Personality Science*, 11(1), 107–116. <https://doi.org/10.1177/1948550619841631>
- Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, 139(3), 655–701. <https://doi.org/10.1037/a0029531>
- Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48(6), 1–29. <https://doi.org/10.18637/jss.v048.i06>
- Cook, D. A., Castillo, R. M., Gas, B., & Artino, A. R. (2017). Measuring achievement goal motivation, mindsets and cognitive load: Validation of three instruments' scores. *Medical Education*, 51(10), 1061–1074. <https://doi.org/10.1111/medu.13405>
- Crooks, T. J., Kane, M., & Cohen, A. S. (1996). Threats to the valid use of assessments. *Assessment in Education: Principles, Policy & Practice*, 3(3), 265–286. <https://doi.org/10.1080/0969594960030302>
- Danner, D., Aichholzer, J., & Rammstedt, B. (2015). Acquiescence in personality questionnaires: Relevance, domain specificity, and stability. *Journal of Research in Personality*, 57, 119–130. <https://doi.org/10.1016/j.jrp.2015.05.004>
- De Castella, K., & Byrne, D. (2015). My intelligence may be more malleable than yours: The revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *European Journal of Psychology of Education*, 30(3), 245–267. <https://doi.org/10.1007/s10212-015-0244-y>

- Desimone, L. M., & Le Floch, K. C. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, 26(1), 1–22. <https://doi.org/10.3102/01623737026001001>
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. London, England: Psychology Press.
- Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, 6(4), 267–285. https://doi.org/10.1207/s15327965pli0604_1
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256–273.
- Elliott, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34(3), 169–189. https://doi.org/10.1207/s15326985ep3403_3
- Elliott, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, 80, 501–519. <https://doi.org/10.1037/0022-3514.80.3.501>
- Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54(1), 5–12. <https://doi.org/10.1037/0022-3514.54.1.5>
- Festinger, L. (1962). Cognitive Dissonance. *Scientific American*, 207(4), 93–106.
- Gonzalez, O., MacKinnon, D. P., & Muniz, F. B. (2021). Extrinsic Convergent Validity Evidence to Prevent Jingle and Jangle Fallacies. *Multivariate Behavioral Research*, 56(1), 3–19. <https://doi.org/10.1080/00273171.2019.1707061>
- Henry, M. A., Shorter, S., Charkoudian, L., Heemstra, J. M., & Corwin, L. A. (2019). FAIL Is not a four-letter word: A theoretical framework for exploring undergraduate students' approaches to academic challenge and responses to failure in STEM learning environments. *CBE—Life Sciences Education*, 18(1), ar11. <https://doi.org/10.1187/cbe.18-06-0108>
- Hoffman, M., Richmond, J., Morrow, J., & Salomone, K. (2002). Investigating "sense of belonging" in first-year college students. *Journal of College Student Retention: Research, Theory & Practice*, 4(3), 227–256. <https://doi.org/10.2190/DRYC-CXQ9-JQ8V-HT4V>
- Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M. S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Psychology and Social Psychology*, 77(3), 588–599.
- Jones, E. E., & Berglas, S. (1978). Control of attributions about the self through self-handicapping strategies: The appeal of alcohol and the role of underachievement. *Personality and Social Psychology Bulletin*, 4(2), 200–206. <https://doi.org/10.1177/014616727800400205>
- Kane, M. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112(3), 527.
- Kelley, T. L. (1927). *Interpretation of educational measurements*. (pp. 1–353). Yonkers-on-Hudson, NY: World Book.
- Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, 347(6219), 262–265. <https://doi.org/10.1126/science.1261375>
- Limeri, L. B., Carter, N. T., Choe, J., Harper, H. G., Martin, H. R., Benton, A., & Dolan, E. L. (2020). Growing a growth mindset: Characterizing how and why undergraduate students' mindsets change. *International Journal of STEM Education*, 7(35), 1–19. <https://doi.org/10.1186/s40594-020-00227-2>
- Limeri, L. B., Choe, J., Harper, H. G., Martin, H. R., Benton, A., & Dolan, E. L. (2020). Knowledge or abilities? How undergraduates define intelligence. *CBE—Life Sciences Education*, 19(1), ar5. <https://doi.org/10.1187/cbe.19-09-0169>
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Meyer, M., Cimpian, A., & Leslie, S.-J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Frontiers in Psychology*, 6, <https://doi.org/10.3389/fpsyg.2015.00235>
- Muenks, K., Canning, E. A., LaCrosse, J., Green, D. J., Zirkel, S., Garcia, J. A., & Murphy, M. C. (2020). Does my professor think my ability can change? Students' perceptions of their STEM professors' mindset beliefs predict their psychological vulnerability, engagement, and performance in class. *Journal of Experimental Psychology: General*, <https://doi.org/10.1037/xge0000763>
- Muradoglu, M., Horne, Z., Hammond, M. D., Leslie, S.-J., & Cimpian, A. (2021). Women—Particularly underrepresented minority women—And early-career academics feel like impostors in fields that value brilliance. *Journal of Educational Psychology*, <https://doi.org/10.1037/edu0000669>
- Nahm, A. Y., Rao, S. S., Solis-Galvan, L. E., & Ragu-Nathan, T. S. (2002). The Q-sort method: Assessing reliability and construct validity of questionnaire items at a pre-testing stage. *Journal of Modern Applied Statistical Methods*, 1(1), 114–125. <https://doi.org/10.22237/jmasm/1020255360>
- National Center for Education Statistics. (2019). *Digest of Educational Statistics: 2019*. Washington, DC: U.S. Department of Education. Retrieved December 1, 2020, from https://nces.ed.gov/programs/digest/2019menu_tables.asp
- Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92(3), 544–555. <https://doi.org/10.1037/0022-0663.92.3.544>
- Porter, T., & Cimpian, A. (2023). A Context's Emphasis on Intellectual Ability Discourages Expression of Intellectual Humility. *Motivation Science*, in press.
- R Core Team. (2021). *R: A language and environment for statistical computing*. Ames, IA: R Foundation for Statistical Computing. Retrieved December 1, 2020, from <https://www.R-project.org/>
- Rattan, A., Good, C., & Dweck, C. S. (2012a). "It's ok — Not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology*, 48(3), 731–737. <https://doi.org/10.1016/j.jesp.2011.12.012>
- Rattan, A., Savani, K., Komarraju, M., Morrison, M. M., Boggs, C., & Ambady, N. (2018). Meta-lay theories of scientific potential drive underrepresented students' sense of belonging to Science, Technology, Engineering, and Mathematics (STEM). *Journal of Personality and Social Psychology*, 115(1), 54–75.
- Rattan, A., Savani, K., Naidu, N. V. R., & Dweck, C. S. (2012b). Can everyone become highly intelligent? Cultural differences in and societal consequences of beliefs about the universal potential for intelligence. *Journal of Personality and Social Psychology*, 103(5), 787–803. <https://doi.org/10.1037/a0029263>
- Revelle, W. (2022). *psych: Procedures for Psychological, Psychometric, and Personality Research (2.2.3) [R package]*. Evanston, IL: Northwestern University.
- Rickert, N. P., Meras, I. L., & Witkow, M. R. (2014). Theories of intelligence and students' daily self-handicapping behaviors. *Learning and Individual Differences*, 36, 1–8. <https://doi.org/10.1016/j.lindif.2014.08.002>
- Rossee, Y. (2012). *lavaan: An R Package for Structural Equation Modeling*. *Journal of Statistical Software*, 48(2), 1–36.
- Scherer, R., & Campos, D. G. (2022). *Measuring those who have their minds set: An item-level meta-analysis of the implicit theories of intelligence scale in education*. *Educational Research Review*, 37, 100479.
- Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., & Macnamara, B. N. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses. *Psychological Science*, 29(4), 549–571. <https://doi.org/10.1177/0956797617739704>
- Smiley, P. A., Butitta, K. V., Chung, S. Y., Dubon, V. X., & Chang, L. K. (2016). Mediation models of implicit theories and achievement goals predict planning and withdrawal after failure. *Motivation and Emotion*, 40(6), 878–894. <https://doi.org/10.1007/s11031-016-9575-5>
- Storage, D., Charlesworth, T. E. S., Banaji, M. R., & Cimpian, A. (2020). Adults and children implicitly associate brilliance with men more than women. *Journal of Experimental Social Psychology*, 90, 104020. <https://doi.org/10.1016/j.jesp.2020.104020>
- Sun, X., Nancekivell, S., Gelman, S. A., & Shah, P. (2021). Perceptions of the malleability of fluid and crystallized intelligence. *Journal of Experimental Psychology: General*, 150(5), 815–827. <https://doi.org/10.1037/xge0000980>
- Tinto, V. (1988). Stages of Student Departure: Reflections on the Longitudinal Character of Student Leaving. *The Journal of Higher Education*, 59(4), 438–455. <https://doi.org/10.2307/1981920>
- Vial, A. C., Muradoglu, M., Newman, G. E., & Cimpian, A. (2022). An Emphasis on Brilliance Fosters Masculinity-Contest Cultures. *Psychological Science*, 095679762110441. <https://doi.org/10.1177/09567976211044133>

- Williams, C. L., Hirschi, Q., Hulleman, C. S., & Roksa, J. (2021). Belonging in STEM: Growth Mindset as a Filter of Contextual Cues. *International Journal of Community Well-Being*. <https://doi.org/10.1007/s42413-021-00111-z>
- Yan, V. X., & Wang, L. (2021). What predicts quality of learners' study efforts? Implicit beliefs and interest are related to mastery goals but not to use of effective study strategies. *Frontiers in Education*, 6, 643421. <https://doi.org/10.3389/feduc.2021.643421>
- Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, 47(4), 302–314. <https://doi.org/10.1080/00461520.2012.722805>
- Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., ... & Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, 573, 364–369. <https://doi.org/10.1038/s41586-019-1466-y>
- Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, 81(2), 267–301.
- Yeager, D. S., Walton, G. M., Brady, S. T., Akcinar, E. N., Paunesku, D., Keane, L., ... & Dweck, C. S. (2016). Teaching a lay theory before college narrows achievement gaps at scale. *Proceedings of the National Academy of Sciences*, 113(24), E3341–E3348. <https://doi.org/10.1073/pnas.1524360113>
- Yuan, K.-H., Chan, W., Marcoulides, G. A., & Bentler, P. M. (2016). Assessing Structural Equation Models by Equivalence Testing With Adjusted Fit Indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(3), 319–330. <https://doi.org/10.1080/10705511.2015.1065414>