Received: 11 September 2020 | Accepted: 4 November 2020

DOI: 10.1002/nse2.20031

Natural Sciences Education

ORIGINAL RESEARCH ARTICLE

Undergraduate Education

Longitudinal measurement invariance and stability of individual interest across a 16-week introductory animal sciences course

MaryGrace Erickson[1] | **Michel A. Wattiaux**[1] | **Elizabeth L. Karcher**[2]

[1] Dep. of Animal and Dairy Sciences, Univ. of Wisconsin–Madison, 1675 Observatory Dr., Madison, WI 53706, USA

[2] Dep. of Animal Sciences, Purdue Univ., 270 S. Russell St., West Lafayette, IN 47906, USA

Correspondence
Elizabeth L. Karcher, Dep. of Animal Sciences, Purdue Univ., 270 S. Russell St., West Lafayette, IN 47906, USA.
Email: ekarcher@purdue.edu

Assigned to Associate Editor Maja Krzic.

Abstract

Educators often compare quantitative scores on motivational traits across time (e.g., pre- and post-semester), yet few studies have examined the longitudinal measurement equivalence of such traits or described typical trajectories. Our research explored the longitudinal measurement of individual interest in two cohorts of an introductory animal sciences course across four measurement occasions during a 16-week semester. First, we modified an existing individual interest scale and validated it within our population (CFI = 1.00, SRMR = .02, RMSEA = .05). Second, we established partial scalar invariance across measurement occasions through nested model comparisons. Third, we described the trajectory of individual interest with latent growth curve models (LGCM). Individual interest started high for both cohorts (intercepts = 65.02, 61.06 on a scale from 0 [low] to 70 [high] for Fall 2018 and Spring 2019, respectively). Individual interest followed a curvilinear pattern in Fall 2018; however, no significant shape trends described Spring 2019 data. Overall, our results show that individual interest can be measured equivalently across a semester; however, it follows heterogenous trajectories. Further research is needed to improve the sensitivity of individual interest scales within high-interest populations and relate heterogenous interest trajectories to classroom experiences.

1 | INTRODUCTION

Previous authors have described introductory science courses as a volatile landscape for individual interest development (Suresh, 2006). On the one hand, introductory courses may lead students to deepening and formalizing their interest in a particular subject matter (Harackiewicz, Smith, & Priniski, 2016; Kyndt et al., 2015). On the other hand, they often function to "weed out" potential students who—for better or worse—find the course and its topics uninspiring (Sithole et al., 2017). Indeed, studies have shown that processes of motivation and achievement differ substantially across students with heterogenous personal, psychosocial, and socioeconomic profiles (De Clercq, Galand, & Frenay, 2020; Martens & Metzger, 2017). Although recent work has advanced theoretical understanding of interest in science undergraduates (O'Keefe, Dweck, & Walton, 2018), operationalized science-specific forms of interest as measurable constructs (Knetka, Rowland, Corwin, & Eddy, 2020; Lamb, Annetta, Meldrum, & Vallett, 2012), and related interest to

Abbreviations: AFI, alternative fit index; BIC, Bayesian information criterion; CFA, confirmatory factor analysis; CFI, comparative fit index; FIML, full-information maximum likelihood; GRR, growth rate reliability; IIQ, individual interest questionnaire; LGCM, latent growth curve model; MFI, McDonald's non-centrality index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.

science introductory course experiences (Erickson, Marks, & Karcher, 2020), limited existing research addresses issues associated with measuring interest. For example, many experimental and observational studies are premised on the implied assumption that interest can be measured equivalently across sampling periods (e.g., pre- and post-semester) and/or student characteristics. Given that interest has been described as a relational, dynamic construct at the intersection of person and environment, it theoretically may shift form amid varying personal or contextual characteristics (Ainley, 2017). Sources of longitudinal interest variation include conceptual recalibration (beta change), reconceptualization (gamma change), and true quantitative changes in interest (alpha change), each of which has different implications for measuring and managing interest in the classroom (Golembiewski, Billingsley, & Yeager, 1976). Additional work exploring the quantitative measurement of interest is needed to elucidate its dynamics in science educational settings and facilitate efforts to target it as a learning outcome. The present research explores the longitudinal measurement of individual interest in introductory animal science undergraduates across a semester-long course.

1.1 | Theoretical framework

Interest is a powerful motivator of learning, engagement, and achievement that takes several forms (Harackiewicz et al., 2016). In their influential Four-Phase Model of Interest Development, Hidi and Renninger (2006) distinguish two main functional forms of interest: situational and individual. Situational interest is a dynamic, multidimensional construct triggered by the immediate features of a task or environment (Chen, Darst, & Pangrazi, 2001; Erickson et al., 2020). In contrast, the present study deals with individual interest. Individual interest is a relatively stable disposition to re-engage with a particular subject matter that develops slowly, representing interest's crystallization as an aspect of identity and personality (Harackiewicz et al., 2016).

Recent work has empirically confirmed theoretical relations between individual and situational interest. For example, Rotgans and Schmidt (2017) showed that repeated experiences of situational interest can lead to the development of individual interest. Knogler, Harackiewicz, Gegenfurtner, and Lewalter (2015) used a latent state-trait model to show that situational interest is context-specific. Within the context of animal sciences, past work has shown that interest remains relatively stable across introductory course experiences (Erickson, Guberman, Zhu, & Karcher, 2019) and that various active learning modalities can prompt varying levels of situational interest (Erickson et al., 2020). However, to our knowledge, no prior work has explored the longitudinal measurement invariance and stability of individual interest across an introductory course.

Core Ideas

- Educators often compare quantitative scores on traits across time.
- Results showed partial scalar invariance of an individual interest scale across a semester.
- Our introductory course students began study with high interest.
- Interest followed diverse trajectories during the semester.

1.2 | Context in animal sciences

Individual interest has assumed particular importance as an assessment outcome in undergraduate animal science education in recent years as the demographics of enrollees changes (Erickson et al., 2019; Lyvers-Peffer, 2011; Reiling, Marshall, Brendemuhl, Mcquagge, & Umphrey, 2003). As fewer Americans experience animal agriculture during childhood and adolescence, fewer opportunities exist for developing individual interest in the subject prior to collegiate study (Dimitri, Effland, & Conklin, 2005). This creates barriers to successful adaptation to animal science industry and academic careers, where practitioners share a strong social identity (Erickson, Guberman, & Karcher, 2020). Accurate quantification of individual interest development during undergraduate animal science education therefore has implications for improving equitable access to our discipline through a variety of pathways.

1.3 | Purpose

This study investigates the measurement properties of Linnenbrink-Garcia et al. (2010) individual interest scale and describes individual interest development within undergraduate animal science students during a 16-week course. The following three objectives guided our research: (a) validate the unidimensional Individual Interest Questionnaire (IIQ) proposed by Linnenbrink-Garcia et al. (2010) within introductory animal science students using a confirmatory factor analysis, (b) test the temporal measurement invariance of individual interest over four timepoints during the semester, and (c) describe the trajectory of individual interest over four timepoints during two cohorts of an introductory course.

TABLE 1 Demographic characteristics of $n = 246$ students (86.6% response rate) and their $n = 491$ reported parents/guardians

Category	Count
Parent/guardian highest education[a]	
Bachelor's degree (BS)	167
High school diploma or equivalency (GED)	147
Master's degree (MS)	76
Associate degree (AS)	47
Professional (MD, JD, DDS, etc.)	18
Doctorate degree (PhD)	13
Other (please specify)	13
None of the above (less than high school)	10
Student gender identification	
Female	193
Male	51
Other/prefer not to respond	2
Student racial and ethnic identity[b]	
White	218
Hispanic or Latino	21
Asian	10
Black or African American	10
Native Hawaiian or Other Pacific Islander	2
Other	2
American Indian or Alaska Native	1

[a]Highest education denotes the final formal schooling received.
[b]Students self-described racial and ethnic identity through a "SELECT ALL" question.

1.4 | Context and participants

This research focuses on students in ANSC 10200, Introduction to Animal Agriculture, a 16-week medium-to-large-enrollment introductory course taught at a midwestern land-grant university. Historically, the course has consisted primarily of first-year and pre-veterinary students with relatively little experience in animal sciences (Erickson et al., 2019). Course sessions include twice-weekly 50-minute lectures and a weekly 110-minute laboratory session. The total sample of participants in this study includes approximately 284 undergraduate students from two introductory animal sciences courses taught by two separate instructors at a large midwestern university. The sample consisted primarily of females and the majority of students were Caucasian (Tables 1 and 2).

2 | METHODS

2.1 | Research design

The Institutional Review Board approved all study procedures. A graduate laboratory coordinator and a total of 11 undergraduate TAs (two to three per laboratory section) facilitated the course's five laboratory sessions and carried out experimental procedures for this study. At the beginning of the semester (Period 0) students completed an outside-of-class pre-questionnaire. At three timepoints during the semester (Period 1–3), students completed a survey assessing their individual interest at the beginning of the course laboratory session. During the final week of the semester, students completed a survey with individual interest and demographic questions outside of class (Period 4). Table 2 shows the exact timing of survey administration, which varied slightly from the spring to fall semester but remained aligned with similar course laboratory topics.

2.2 | Instrumentation

To measure individual interest, we selected a questionnaire conceptually aligned with our operationalization of interest development at the academic discipline level: Linnenbrink-Garcia and colleagues' (2010) theoretically unidimensional eight-item Individual Interest Questionnaire (New Table 3). Linnenbrink-Garcia et al. (2010) adapted the IIQ from the task value scale of Pintrich, Smith, Garcia, and Mckeachie's (1993) Motivated Strategies for Learning Questionnaire, observing it to be highly reliable ($\alpha = .90$) and empirically distinct from situational interest in undergraduate psychology. We adapted wording to make items specific to our context (e.g., "Animal sciences is practical for me to know"). Students responded to all items via a Qualtrics form using a sliding scale from 0 (strongly disagree) to 70 (strongly agree) during experimentation (Qualtrics). Because the pre-test used a Likert scale with radio buttons 1 to 7, responses were excluded from measurement invariance testing but rescaled and incorporated into latent growth analysis. Descriptive statistics across each period are presented in Tables 4 and 5).

2.3 | Statistical analysis

We conducted all statistical analyses in R (R Core Team, 2019) using primarily the lavaan and semTools packages (Jorgensen, Pornprasertmanit, Schoemann, & Rosseel, 2019; Rosseel, 2012) and declaring significance at $p < .05$. To handle missing data, we used the Full Information Maximum Likelihood (FIML) Estimator suggested by Arbuckle (1996). Compared with listwise deletion, FIML has been shown to produce less-biased parameter estimates (Ferro, 2014). Similar to Linnenbrink-Garcia et al. (2010), our IIQ responses violated multivariate normality assumptions (Curran, West, & Finch, 1996). To accommodate, we used maximum likelihood estimation with robust Huber-sandwich estimation of standard errors (Huber, 1967; White, 1980). Likewise, we

TABLE 2 Experimental schedule, response count, and response rate for the Fall 2018 and Spring 2019 semesters

	Cohort					
	Fall 2018			Spring 2019		
Period	Week	*n*	%	Week	*n*	%
0	1	147	79.0%	1	91	92.3%
1	4	164	88.2%	6	87	88.8%
2	7	158	84.9%	7	92	93.9%
3	9	159	85.5%	12	93	94.9%
4	15	157	84.4%	15	78	79.6%
Total unique		186	100.0%		98	100.0%

TABLE 3 Individual interest questionnaire

Individual interest questionnaire (IIQ)	
II_1	Animal sciences is practical for me to know.
II_2	Animal sciences helps me in my daily life outside of school.
II_3	It is important to me to be a person who thinks like an animal scientist.
II_4	Thinking like an animal scientist is an important part of who I am.
II_5	I enjoy the subject of animal sciences.
II_6	I like animal sciences.
II_7	I enjoy doing animal sciences activities.
II_8	Animal sciences is exciting to me.

TABLE 4 Frequencies of missing data patterns across the four experimental periods. $N = 284$ students

Period					
Pre	1	2	3	4	*N*
1	1	1	1	1	183
1	1	1	1	0	24
0	1	1	1	1	4
0	1	1	1	0	18
1	1	0	1	1	7
1	1	0	1	0	1
0	1	0	1	1	1
0	1	0	1	0	1
1	0	1	1	1	9
1	0	1	1	0	2
0	0	1	1	0	1
1	0	0	1	1	1
1	1	1	0	1	8
1	1	0	0	1	2
1	1	0	0	0	1
0	1	0	0	1	1
0	0	1	0	0	1
0	0	0	0	1	19
46	33	34	32	49	

Note. Pattern expressed as 1 = present, 0 = missing for Periods 1, 2, 3, and 4.

TABLE 5 Raw descriptive statistics item-by-item for the individual interest questionnaire (IIQ) for each of the four experimental periods

Period	1		2		3		4		All	
Item	M	SD	M	SD	M	SD	M	SD	Skew.	Kurt.
II_1	63.2	14.6	64.4	10.9	61.9	13.4	62.2	13.2	–2.7	0.4
II_2	52.4	18.4	57.7	15.6	57.2	16.7	54.3	17.2	–1.3	0.5
II_3	60.6	14.5	61.5	12.8	60.2	14.6	58.8	15.5	–1.9	0.4
II_4	56.8	17.4	59.6	14.5	59.2	15.1	56.9	17.3	–1.6	0.5
II_5	65.1	11.7	64.7	10.1	63.1	12.3	62.8	12.4	–2.9	0.4
II_6	65.8	10.4	65.0	10.0	63.5	11.8	63.3	12.9	–3.0	0.4
II_7	65.1	10.8	64.4	10.2	62.9	12.5	62.0	13.3	–2.7	0.4
II_8	64.3	12.3	64.6	10.0	62.1	13.9	61.8	13.7	–2.6	0.4

Note: M, mean; SD, standard deviation; skew., skewness; kurt., kurtosis. Table 1 presents response rate and count for each period.

compared nested models using the scaled difference in the log likelihood chi squared statistic ($\Delta\chi^2$) and absolute differences in alternative fit indices (AFIs). The $\Delta\chi^2$ represents the standard approach, whereas AFIs are modern alternatives that have the advantage of being independent of sample size. Fit indices included the scaled root mean squared error of approximation (RMSEA; Steiger, 1989), the scaled comparative fit index (CFI; Hu & Bentler, 1999), the standardized root mean square residual (SRMR; Hu & Bentler, 1999), gamma hat ($\hat{\gamma}$; West, Taylor, & Wu, 2012), and McDonald's noncentrality index (MFI; McDonald, 1989). Per the recommendations of Cheung and Rensvold (2002), we considered $\Delta\text{CFI} \geq .01$, $\Delta\hat{\gamma} \geq .001$, and $\Delta\text{MFI} \geq .02$ as indicative of significant differences in nested model fit. We constructed diagrams with the semPlot package (Epskamp, 2019).

3 | RESULTS

3.1 | Single-group confirmatory factor analysis

To validate the measurement model within our population and check for problematic items, we fitted preliminary CFA models on a subset of data from experimental Period 2 (Figure 1; Table 6). We selected Period 2 due to the high response rate and the synchronicity of its timing across semesters. The initial model fit the data poorly and modification indices revealed two problematic indicators, "II.2" and "II.4" (Awang, 2012). Both indicators were significantly and highly correlated with other indicators on the scale but had higher error variance and lower loadings. Because ancillary parallel analyses suggested a one-factor solution and only two indicators were problematic, we removed items and proceeded with a reduced unidimensional scale rather than attempting to fit two- or three-factor models. The revised model showed excellent fit (Hu & Bentler, 1999; Tabachnick & Fidell, 2007).

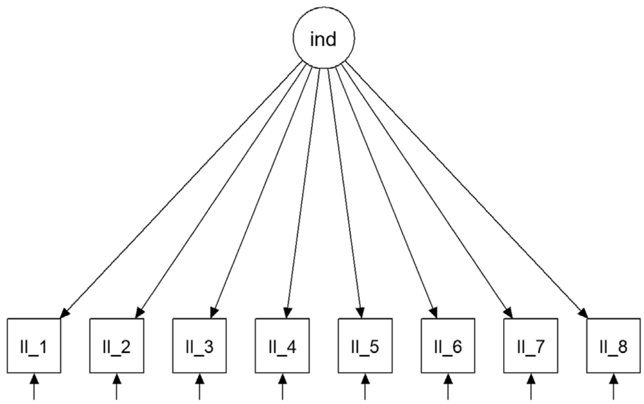

FIGURE 1 Proposed measurement model for individual interest

3.2 | Longitudinal measurement invariance

Next, we began testing the measurement invariance across time. First, we examined a configural model that simultaneously estimated four correlated factors representing each of the measurement periods. We standardized factor means and variances to 0 and 1, respectively, to identify the model (Reise, Widaman, & Pugh, 1993). We freely estimated loadings, intercepts, and the residual covariances between each indicator across experimental periods. The configural model showed adequate fit (Table 7) and inspection of modification indices did not reveal logical opportunities for re-specification. Consequently, we concluded that configural invariance requirements were satisfied and moved forward with constraining additional parameters.

Then we constrained each item's factor loadings to equality across experimental periods to test metric invariance. In this model, we again fixed the latent factor means to 0 and the variance of the factor representing the first experimental period to 1. However, we freely estimated variances for subsequent experimental periods (Johnson, Meade, & DuVernet, 2009). The metric invariance model showed adequate fit (Table 7) and did not differ significantly from the configural model

TABLE 6 Standardized results of initial and revised CFA models for the Individual Interest Questionnaire on experimental Period 2. $N = 254$ students

Initial model				Revised model		
	Est.	**SE**	***P***	**Est.**	**SE**	***P***
Factor loadings						
II.1	.85	.13	<.001	.84	.13	<.001
II.2	.59	.08	<.001			
II.3	.77	.09	<.001	.75	.09	<.001
II.4	.72	.09	<.001			
II.5	.89	.12	<.001	.90	.12	<.001
II.6	.93	.13	<.001	.94	.13	<.001
II.7	.93	.12	<.001	.94	.12	<.001
II.8	.92	.11	<.001	.93	.11	<.001
Residual variances						
II.1	.28	.06	<.001	.30	.07	<.001
II.2	.65	.10	<.001			
II.3	.40	.08	<.001	.44	.09	<.001
II.4	.48	.08	<.001			
II.5	.20	.10	.049	.19	.10	.054
II.6	.13	.04	.005	.11	.04	.005
II.7	.13	.03	<.001	.12	.03	<.001
II.8	.14	.04	<.001	.14	.04	.001
Fit indices						
χ^2	111.21***			10.65		
df	20			9		
CFI	.88			1.00		
SRMR	.08			.02		
RMSEA	.22			.05		

Note. Est., estimate; SE, standard error; CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual. Scaled values are presented for CFI, RMSEA, and χ^2.
***Significance at the .001 level.

TABLE 7 Measurement invariance models for the Individual Interest Questionnaire (IIQ) of $n = 284$ students across four experimental periods throughout the semester

	Fit Index							
	df	**RMSEA**	**CFI**	**SRMR**	$\hat{\gamma}$	**MFI**	χ^2	**Δdf**
Configural	210	.07	.91a[a]	.09	.913a	.274a	556.04A	
Metric	225	.07	.90a	.09	.913a	.256a	572.00A	15
Scalar	240	.07	.90a	.09	.908b	.248a	608.84B	15
Partial scalar (Item-wise)	234	.07	.90a	.09	.911b	.255a	589.61A	9
Partial scalar (Period-wise)	234	.07	.90a	.09	.908b	.248a	597.56B	9

Note. RMSEA, root mean square error of approximation; CFI, comparative fit index; SRMR, standardized root mean square residual; MFI, McDonald's non-centrality index.
[a]Capital letters represent pairwise differences at $p < .05$. Lowercase letters indicate pairwise differences based on Cheung and Rensvold (2002) proposed cutoffs. Scaled values are presented for df, CFI, RMSEA, and χ^2.

TABLE 8 Descriptive statistics for individual interest of $n = 284$ students before and during four experimental periods throughout the semester

Period	Cohort: Fall 2018 (n = 186)	Spring 2019 (n = 98)
Pre-test	64.7 (60.0–69.6)	61.8 (58.0–69.6)
1	66.4 (65.0–70.0)	61.8 (59.0–70.0)
2	66.1 (65.5–70.0)	61.5 (57.2–70.0)
3	65.0 (63.5–70.0)	57.2 (51.7–70.0)
4	62.3 (59.8–70.0)	59.9 (55.4–70.0)

Note. Mean (first quartile–third quartile). Rated on a continuous scale from 0 (strongly disagree) to 70 (strongly agree) corresponding to low and high individual interest, respectively.

on the basis of the $\Delta\chi^2$ (Counsell, Cribbie, & Flora, 2020). This indicated metric invariance (i.e., the equivalence of the contribution of each indicator to the latent factor over time). We subsequently proceeded to fitting a scalar invariance model.

Finally, we tested a scalar invariance model by constraining indicator intercepts to equality across each experimental period. In this model, we fixed the mean and variance of the latent factor representing the first experimental period to 0 and 1, respectively, and freely estimated means and variances for factors representing experimental Periods 2, 3, and 4. This model constrained factor loadings and intercepts for all indicators to equality across time but allowed free estimation of residual variances. The $\Delta\chi^2$ indicated significantly poorer fit of the scalar compared with the metric invariance model. Comparing model alternative fit indices showed mixed results, with ΔCFI and ΔMFI indicating scalar invariance and $\Delta\hat{\gamma}$ suggesting scalar non-invariance. As a result, we sequentially removed constraints on equality of intercepts for the most non-invariant items to test a partial scalar invariance model (Partial Scalar Item-wise). Releasing the constraints on items 1 and 3 produced a model that satisfied not only ΔCFI and ΔMFI requirements but also the $\Delta\chi^2$ test. We therefore concluded that the IIQ was at least partially temporally invariant in our sample (i.e., that the majority of indicators functioned similarly over the four measurement periods).

Because our data were not missing completely at random, we examined the possibility that unique and missing responses at Period 4 had altered scalar invariance results. To do so, we released equality constraints and freely estimated intercepts for all indicators on Period 4. All other model parameters were estimated as described in the initial scalar invariance model above. The model (Partial Scalar Period-wise) did not show substantially improved fit on the basis of $\Delta\chi^2$ and ΔAFIs compared with the metric model. Because this model had the same degrees of freedom as the Item-wise Partial Scalar model, we therefore concluded that non-invariance in means of items 1 and 3 throughout the semester contributed more substantially to total scale non-invariance than did non-invariance in means of all items on experimental Period 4.

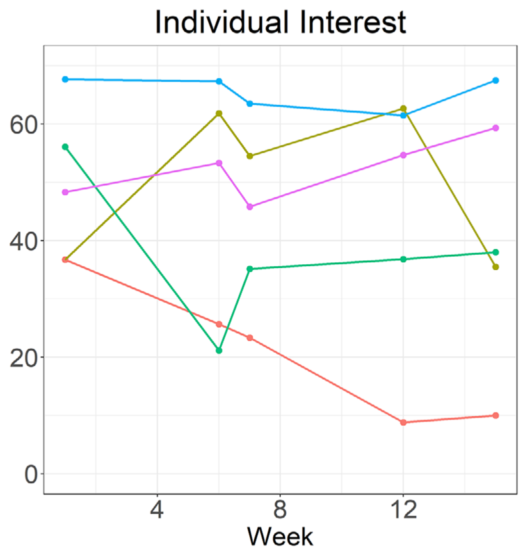

FIGURE 2 Semester-long individual interest profiles of selected students, demonstrating heterogeneous growth patterns ($n = 5$)

3.3 | Longitudinal descriptive statistics

Having established partial measurement invariance, we next explored the trajectory of individual interest throughout the semester in both our fall and spring classes (Table 8). Responses were skewed toward the upper end of the scale throughout the semester for both cohorts. Although graphical visualization revealed substantial between-subjects variation in growth trajectories (e.g., Figure 2), slight negative trends appeared present for both cohorts. To explore whether trends represented statistically significant group-level phenomena, we next fitted latent growth curve models.

3.4 | Latent growth curve model

To explore group-level patterns in individual interest across time, we estimated latent growth curve models (LGCM) of the standard form described by Duncan and Duncan (2009) in which the intercepts, variances, and covariance of latent intercept and slope parameters are freely estimated whereas

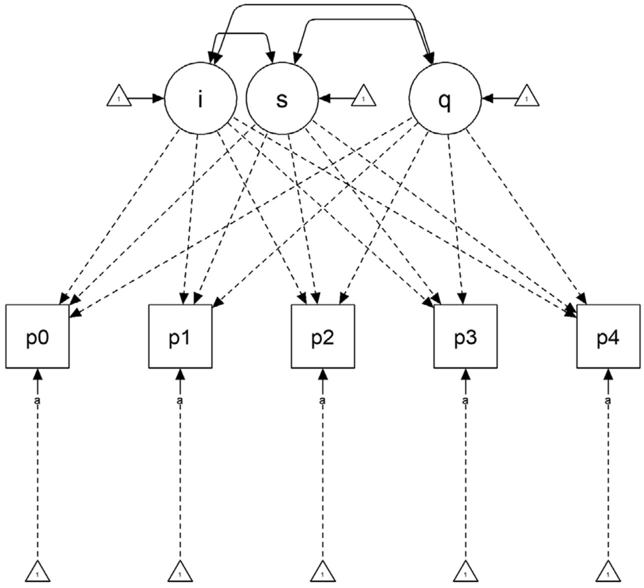

FIGURE 3 General form of latent growth curve models for individual interest during the semester. Latent linear growth, quadratic growth, and intercepts are denoted by s, q, and i, respectively

intercepts of manifest variables (individual interest at each sampling) are fixed to zero and factor loadings on the latent variables are fixed to numeric weights (Figure 3; McArdle & Bell, 2000). Because sampling timing differed for Fall 2018 and Spring 2019 cohorts, we estimated two separate models and interpolated linear and quadratic slope weights scaling time in 1-week units. We again accounted for missing data by employing the FIML estimator (Arbuckle, 1996). In each cohort, our sample size and observations per individual satisfied the recommended minima proposed by Curran, Obeidat, and Losardo (2010). However, we computed Willett's (1989) growth rate reliability (GRR) to estimate statistical power to detect slope variance for each LGCM.

For both cohorts, we explored a series of nested models estimating: (a) a latent linear growth term and intercept, and (b) latent linear and quadratic growth terms and intercept. To identify the second model, we assumed homoscedasticity of time-specific residual variances and constrained them to equality. Inspection of model estimates and BIC indicated that the second form of model better described the data from both cohorts.

A LGCM with both linear and quadratic growth terms described the Fall 2018 data well (Table 9). Model results imply that individual interest started high on average (intercept of latent intercept) albeit with significant variation around the starting point (variance of latent intercept). Individual interest growth during the semester was significantly predicted by positive linear and slight negative quadratic time terms (linear and quadratic intercepts). However, significant latent variances for both growth terms signaled heterogeneity in growth paths between individual students. Neither linear nor quadratic growth terms covaried significantly with the intercept, indicating that the initial level of individual interest had no bearing on growth observed. Willett's (1989) GRR was > .99 for each Fall 2018 measurement occasion suggesting excellent reliability. Taken together, these results suggested that individual interest in the Fall 2018 cohort started high and followed a curvilinear trend, unaffected by starting level but varying significantly between individuals.

The LGCM approximated the Spring 2019 data poorly. A significant intercept implied that individual intercept started at the upper range of the scale. However, we discovered no significant shape trends. Residual variances and global fit indices showed poor model fit with the data. In addition, Willett's (1989) GRR ranged from .30 to .70 for each measurement occasion, indicating poor reliability of the LGCM. Although the sample size of the Spring 2019 cohort was roughly half that of the Fall 2018 cohort, the large fraction of residual variance indicated that misfit derived from misspecification—the linear and quadratic shape parameters did not describe the data well. Although it is possible to specify more complex shape parameters in LGCM, this approach extended beyond our objectives and was not suggested by our data (Ning & Luo, 2017). As a result, we concluded that Spring 2019 individual interest showed no consistent shape patterns, and its trajectory was obscured substantially by both within- and between-individual variation.

4 | LIMITATIONS

Before findings are discussed in detail, several limitations of our research warrant discussion. First, our research examined a sample with limited diversity from one introductory course at one university. As such, the generalizability of our results is restricted to similar populations. Second, we assumed data were missing at random and used a model-based estimator to account for missing responses. Although missing data and FIML estimation are both standard in longitudinal studies, non-monotone missingness can bias parameter estimates (Ferro, 2014). Third, our experiment assessed narrow aims related to measurement invariance over the course of the semester. We had insufficient power to test measurement invariance between semesters or along other dimensions such as culture and gender (Kline, 2005). Future studies with larger, more diverse samples or employing different experimental design (e.g., staggered survey administration) may more precisely model other possible sources of non-invariance. A further limitation of our study is its reliance on self-report data. Although self-report measurement is standard in interest literature, it nonetheless relies on participants' metacognitive awareness of their interest and is susceptible to certain testing effects (Renninger & Su, 2012). Future studies examining multiple sources of information from multiple data collection

TABLE 9 Latent growth curve models of individual interest in two cohorts of an introductory animal sciences course (Fall, $N = 186$; Spring, $N = 98$)

	Fall 2018			Spring 2019		
	Est.	SE	*p*	Est.	SE	*p*
Factor loadings (linear)						
Pre	0.00[a]			0.00[a]		
1	3.00[a]			5.00[a]		
2	6.00[a]			6.00[a]		
3	8.00[a]			11.00[a]		
4	14.00[a]			14.00[a]		
Factor loadings (quadratic)						
Pre	0.00[a]			0.00[a]		
1	9.00[a]			25.00[a]		
2	36.00[a]			36.00[a]		
3	64.00[a]			121.00[a]		
4	196.00[a]			196.00[a]		
Residual variance						
Each	11.16	1.99	<.01	52.72	17.64	<.01
Intercepts						
Intercept	65.02	0.46	<.01	61.56	0.89	<.01
Linear	0.47	0.13	<.01	−0.13	0.27	.62
Quadratic	−0.05	0.01	<.01	−0.01	0.02	.77
Variances						
Intercept	24.77	6.25	<.01	22.91	20.52	.26
Linear	1.43	0.41	<.01	0.94	1.79	.60
Quadratic	0.01	0.00	<.01	0.00	0.01	.91
Covariances						
Intercept with linear	−0.47	1.65	.78	9.51	5.55	.09
Intercept with quadratic	0.06	0.11	.58	−0.54	0.34	.11
Linear with quadratic	−0.09	0.03	<.01	−0.03	0.13	.81
Fit indices						
χ^2 (scaled)	10		<.01	10		<.01
CFI	0.95			0.89		
TLI	0.95			0.89		
RMSEA	0.09			0.09		

Note. Est., estimate; SE, standard error; RMSEA, root mean square error of approximation; CFI, comparative fit index; TLI, Tucker–Lewis index.
[a]Fixed parameter.

methods may enhance the trustworthiness of results from the revised IIQ.

Additionally, our research raised serious questions regarding the sensitivity of the IIQ and the treatment of individual interest scores as continuous responses. Our data were negatively skewed with a large proportion of observations at the upper limit. Censoring of responses at the upper end of the scale reduces both the content validity of the IIQ and its sensitivity within highly interested populations. Additional scale development work may improve researchers' ability to differentiate students at high levels of individual interest. For existing IIQ data, it may be sensible to treat responses as ordinal and fit structural models with Muthén and Muthén's (2007) weighted least squares with means and variances adjusted estimator (Li, 2016). When considering longitudinal IIQ data, which likely involves not only censored responses but also between-person heterogeneity of growth trajectories, latent class growth analysis, growth mixture models, and Tobit growth models may provide superior predictive accuracy and reduced bias compared with traditional LGCM (Feng, Hancock, & Harring, 2019; Ram & Grimm, 2009; McArdle & Anderson, 1990).

5 | DISCUSSION

Our research examined the measurement of individual interest over four testing periods during the semester of an introductory animal science course. Although interest has assumed importance as an educational outcome and teaching quality indicator in animal sciences, limited research has assessed issues related to its measurement. Our research addressed three such aims: (a) validity and reliability of a modified individual interest scale within the animal science context, (b) measurement equivalence of the individual interest scale across a 16-week introductory course, and (c) patterns in mean individual interest (i.e., "alpha change") over time.

First, our initial single timepoint confirmatory factor analyses extended past research (Linnenbrink-Garcia et al., 2010) to a new population—showing the revised IIQ to be a valid and reliable measure of individual interest in animal science undergraduates after removing two problematic items. Second, similar to past research, which has demonstrated the temporal measurement equivalence of closely related constructs such as intrinsic and extrinsic motivation throughout the first year of university study (Brahm, Jenert, & Wagner, 2017), we found that the revised IIQ achieved partial scalar invariance over four measurements during the semester. This indicates that the measurement properties of most revised IIQ indicators do not change substantially over time and composite scores can capture true changes in individual interest across measurement occasions (i.e., alpha change; Brown, 2006; Golembiewski et al., 1976).

However, pursuant to our first and second objectives, CFA and measurement invariance analyses discovered several items that were inconsistent within (II.2 and II.4) and between (II.1 and II.3) time periods, which we addressed by reducing the scale and freeing model constraints. However, more research is warranted to understand why these items produced erratic responses and to determine if replacement or additional items can improve the validity and reliability of the IIQ. Notably, the inconsistent items discovered in the initial CFA (II.2 and II.4) align conceptually with the non-invariant items (II.1 and II.3) discovered in subsequent analysis: all focus on individual interest's value and identification dimensions (Harackiewicz & Hulleman, 2010). In contrast to the enjoyment-focused items (II.5–II.8) that appear relatively consistent, these value- and identification-focused items functioned inconsistently across students and time.

This differential functioning, in part, may be traceable to discipline-based differences. For example, the immediate practical relevance of animal science may appear uncertain for students who leave family farms or other animal-based employment to pursue undergraduate degrees, or those with relatively little prior animal experience (Fraser, 2010). Similarly, social identification with animal sciences likely differs from that with psychology, the disciplinary context in which the IIQ was developed. Animal science is a broad, heterogenous discipline in which subgroup membership may be more salient to identity—in other words, a student might identify more strongly as a swine scientist than an animal scientist (Stets & Burke, 2000). Indeed, Knogler et al. (2015) highlighted the narrowness of content breadth in individual interest as potentially incongruous with acknowledging its development from situational sources. Operationalizing individual interest as broader and more tied to action (e.g., Lawless & Kulikowich, 2006) may improve test sensitivity. These issues draw attention to a need to develop individual interest measures that accurately reflect the complex social and personal value animal science holds for its diverse undergraduate constituents.

Finally, our third objective involved profiling individual interest across a semester-long introductory course through LGCM. Past research has shown growth in the individual interest of primary school students over 4 weeks when situational interest is repeatedly stimulated (Rotgans & Schmidt, 2011). In contrast, Frenzel, Goetz, Pekrun, and Watt (2010) described curvilinear declines in interest as middle school students studied mathematics over 4 years. Our results showed that individual interest was fairly stable and generally concentrated at the upper end of the scale but varied substantially between individuals. Additionally, we detected significant linear and quadratic shape trends for the Fall 2018 but not Spring 2019 cohorts. One potential source of this discrepancy may be differences in teacher, student, and course characteristics. Fall 2018 and Spring 2019 cohorts were taught by different instructional teams, leading to slight variation in the course content and teaching style. Different students comprised each cohort, and class sizes were smaller in Spring 2019 (∼20 vs. ∼37 per laboratory). Contextual factors likely also differed between cohorts. For example, many fall semester enrollees took the course during their first semester of undergraduate study as they transitioned to college life. In contrast, most spring semester enrollees entered the course with at least one semester of prior experience.

Our study is not the first to describe heterogeneous, non-linear motivational growth in first-year students. For example, in a 25-month study spanning the transition from secondary to higher education across a range of disciplines, Kyndt et al. (2015) fit models describing slopes for autonomous motivation growth that differed across five timepoints. De Clercq et al. (2020) and Martens and Metzger (2017) illustrated that the development of academic motivation differs substantially across student groups of varying personal, psychosocial, and socioeconomic characteristics. As Hofer (2010) points out, individual interest forms amid a network of diverse, often conflicting goals and personal experiences. It is possible that major life events such as the adjustment to college influence the trajectory of individual interest independently or in concert with course experiences.

6 | CONCLUSIONS

A revised version of the individual interest instrument from Linnenbrink-Garcia et al. (2010) demonstrated excellent reliability and validity with animal science introductory course students after removing two problematic items. Additionally, the revised IIQ exhibited partial scalar invariance over four measurement occasions throughout a semester of instruction, indicating that mean differences can capture true changes in individual interest over time. However, further research on the psychometric properties of individual interest scales is warranted to improve sensitivity in high-interest populations and measurement validity of value/identification dimensions. With respect to interest trajectories, LGCM showed that group mean individual interest during a Fall 2018 cohort started high and followed a curvilinear trend. However, LGCM of a Spring 2019 cohort detected no significant linear or quadratic shape trends across subjects and showed substantial between-subjects variation in growth trajectories. Interest development during undergraduate study likely follows heterogenous paths, which can represent both spurious between-person differences and structural forces. More longitudinal research is warranted to understand the complexity of individual interest's trajectory during undergraduate study and its interconnections with other developmental processes (e.g., knowledge and skill development), personal characteristics (e.g., achievement-related beliefs), behavior patterns (e.g., course-taking, performance-approach goals), and achievement outcomes (e.g., graduation, GPA).

ACKNOWLEDGMENTS

We are grateful to the student survey respondents who made this work possible. Additionally, we thank the reviewers and editorial team for their helpful suggestions on an earlier draft.

AUTHOR CONTRIBUTIONS

MaryGrace Erickson: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Visualization; Writing-original draft; Writing-review & editing. Michel A. Wattiaux: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing-review & editing. Elizabeth L. Karcher: Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Writing-review & editing.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

MaryGrace Erickson https://orcid.org/0000-0002-8919-2664
Michel A. Wattiaux https://orcid.org/0000-0001-8713-1641
Elizabeth L. Karcher https://orcid.org/0000-0002-1432-2727

REFERENCES

Ainley, M. (2017). Interest: Knowns, unknowns, and basic processes. In P. A. O'Keefe, & J. M. Harackiewicz (Eds.), *The science of interest* (pp. 3–24). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-55509-6_1

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 243–277). Mahwah, NJ: Lawrence Erlbaum.

Awang, Z. (2012). *A handbook on SEM, "Structure equation modeling* (4th ed., Vol. *4*). University Technology MARA Kelantan: Malaysia Press.

Brahm, T., Jenert, T., & Wagner, D. (2017). The crucial first year: A longitudinal study of students' motivational development at a Swiss business school. *Higher Education*, *73*(3), 459–478. https://doi.org/10.1007/s10734-016-0095-8

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Chen, A., Darst, P. W., & Pangrazi, R. P. (2001). An examination of situational interest and its sources in physical education. *British Journal of Educational Psychology*, *71*(3), 383–400.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Counsell, A., Cribbie, R. A., & Flora, D. B. (2020). Evaluating equivalence testing methods for measurement invariance. *Multivariate Behavioral Research*, *55*(2), 312–328. https://doi.org/10.1080/00273171.2019.1633617

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, *11*(2), 121–136. https://doi.org/10.1080/15248371003699969

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29. https://doi.org/10.1037/1082-989X.1.1.16

De Clercq, M., Galand, B., & Frenay, M. (2020). One goal, different pathways: Capturing diversity in processes leading to first-year students' achievement. *Learning and Individual Differences*, *81*, 101908. https://doi.org/10.1016/j.lindif.2020.101908

Dimitri, C., Effland, A., & Conklin, N. (2005). *The 20th century transformation of U.S. agriculture and farm policy* (Econ. Inf. Bull. 3).

Washington, DC: USDA-ERS. Retrieved from https://www.ers.usda.gov/webdocs/publications/44197/13566_eib3_1_pdf

Duncan, T. E., & Duncan, S. C. (2009). The ABCs of LGM: An introductory guide to latent variable growth curve modeling. *Social and Personality Psychology Compass*, *3*(6), 979–991. https://doi.org/10.1111/j.1751-9004.2009.00224.x

Epskamp, S. (2019). *semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output*. R package version 1.1.2. Retrieved from https://CRAN.R-project.org/package=semPlot

Erickson, M., Marks, D., & Karcher, E. (2020). Characterizing student engagement with hands-on, problem-based, and lecture activities in an introductory college course. *Teaching & Learning Inquiry*, *8*(1), 138–153. https://doi.org/10.20343/teachlearninqu.8.1.10

Erickson, M. G., Guberman, D., & Karcher, E. L. (2020). Undergraduates' experiences of transculturation toward engaged pedagogy through a partnership program in animal sciences. *NACTA Journal*, *64*(1), 214–223. Retrieved from https://www.nactateachers.org/index.php/volume-64-2020/3017-undergraduates-experiences-of-transculturation-toward-engaged-pedagogy-through-a-partnership-program-in-animal-sciences

Erickson, M. G., Guberman, D., Zhu, H., & Karcher, E. L. (2019). Interest and active learning techniques in an introductory animal science course. *NACTA Journal*, *63*(1), 293–298. Retrieved from https://www.nactateachers.org/attachments/article/2872/34%20NACTA%20Journal%20MS2018_0062.pdf

Feng, Y., Hancock, G. R., & Harring, J. R. (2019). Latent growth models with floors, ceilings, and random knots. *Multivariate Behavioral Research*, *54*(5), 751–770. https://doi.org/10.1080/00273171.2019.1580556

Ferro, M. A. (2014). Missing data in longitudinal studies: Cross-sectional multiple imputation provides similar estimates to full-information maximum likelihood. *Annals of Epidemiology*, *24*(1), 75–77. https://doi.org/10.1016/j.annepidem.2013.10.007

Fraser, D. (2010). Farm animal production: Changing agriculture in a changing culture. *Journal of Applied Animal Welfare Science*, *4*(3), 175–190. https://doi.org/10.1207/s15327604jaws0403_02

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. G. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence*, *20*(2), 507–537. https://doi.org/10.1111/j.1532-7795.2010.00645.x

Golembiewski, R. T., Billingsley, K., & Yeager, S. (1976). Measuring change and persistence in human affairs: Types of change generated by OD designs. *The Journal of Applied Behavioral Science*, *12*(2), 133–157. https://doi.org/10.1177/002188637601200201

Harackiewicz, J. M., & Hulleman, C. S. (2010). The importance of interest: The role of achievement goals and task values in promoting the development of interest. *Social and Personality Psychology Compass*, *4*(1), 42–52. https://doi.org/10.1111/j.1751-9004.2009.00207.x

Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters. *Policy Insights from the Behavioral and Brain Sciences*, *3*(2), 220–227. https://doi.org/10.1177/2372732216655542

Hidi, S., & Renninger, A. K. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hofer, M. (2010). Adolescents' development of individual interests: A product of multiple goal regulation? *Educational Psychologist*, *45*(3), 149–166. https://doi.org/10.1080/00461520.2010.493469

Hu, L. T., & Bentler, M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (pp. 221–233). Berkeley: University of California Press.

Johnson, E. C., Meade, A. W., & DuVernet, A. M. (2009). The role of referent indicators in tests of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 642–657. https://doi.org/10.1080/10705510903206014

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2019). *semTools: Useful tools for structural equation modeling*. R package version 0.5-2. Retrieved from https://CRAN.R-project.org/package=semTools

Kline, R. B. (2005). *Methodology in the social sciences. Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Knetka, E., Rowland, A. A., Corwin, L. A., & Eddy, S. (2020). Measuring university students' interest in biology: Evaluation of an instrument targeting Hidi and Renninger's individual interest. *International Journal of STEM Education*, *7*, 23. https://doi.org/10.1186/s40594-020-00217-4

Knogler, M., Harackiewicz, J. M., Gegenfurtner, A., & Lewalter, D. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemporary Educational Psychology*, *43*, 39–50. https://doi.org/10.1016/j.cedpsych.2015.08.004

Kyndt, E., Coertjens, L., van Daal, T., Donche, V., Gijbels, D., & van Petegem, P. (2015). The development of students' motivation in the transition from secondary to higher education: A longitudinal study. *Learning and Individual Differences*, *39*, 114–123. https://doi.org/10.1016/j.lindif.2015.03.001

Lamb, R. L., Annetta, L., Meldrum, J., & Vallett, D. (2012). Measuring science interest: Rasch validation of the science interest survey. *International Journal of Science and Mathematics Education*, *10*, 643–668. https://doi.org/10.1007/s10763-011-9314-z

Lawless, K., & Kulikowich, J. (2006). Domain knowledge and individual interest: The effects of academic level and specialization in statistics and psychology. *Contemporary Educational Psychology*, *31*, 30–43. https://doi.org/10.1016/j.cedpsych.2005.01.002

Li, C. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*, 936–949. Retrieved from https://doi-org.ezproxy.library.wisc.edu/10.3758/s13428-015-0619-7

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement*, *70*(4), 647–671. https://doi.org/10.1177/0013164409355699

Lyvers-Peffer, P. A. (2011). Demographics of an undergraduate animal sciences course and influence of gender and major on course performance. *NACTA Journal*, *55*, 26–31.

Martens, T., & Metzger, C. (2017). Different transitions of learning at university: Exploring the heterogeneity of motivational processes. In E. Kyndt, V. Donche, K. Trigwell, & S. Lindblom-Ylänne (Eds.), *Higher education transitions: Theory and research* (1st ed., pp. 31–46). London: Routledge.

McArdle, J. J., & Anderson, E. (1990). Latent variable growth models for research on aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging (third edition)* (pp, 21–44). San Diego, CA: Academic Press. https://doi.org/10.1016/B978-0-12-101280-9.50008-5

McArdle, J. J., & Bell, R. Q. (2000). An introduction to latent growth models for developmental data analysis. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 69–107, 269–281). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

McDonald, R. P. (1989). An index of goodness-of-fit based on noncentrality. *Journal of Classification*, *6*, 97–103.

Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.

Ning, L., & Luo, W. (2017). Specifying turning point in piecewise growth curve models: Challenges and solutions. *Frontiers in Applied Mathematics and Statistics*, *3*, 19. https://doi.org/10.3389/fams.2017.00019

O'Keefe, P. A., Dweck, C. S., & Walton, G. M. (2018). Implicit theories of interest: Finding your passion or developing it? *Psychological Science*, *29*(10), 1653–1664. https://doi.org/10.1177/0956797618780643

Pintrich, P. R., Smith, D. A. F., Garcia, T., & Mckeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (Mslq). *Educational and Psychological Measurement*, *53*(3), 801–813. https://doi.org/10.1177/0013164493053003024

R Core Team (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Ram, N., & Grimm, K. J. (2009). Methods and measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development*, *33*(6), 565–576. https://doi.org/10.1177/0165025409343765

Reiling, B. A., Marshall, T. T., Brendemuhl, J. H., Mcquagge, J. A., & Umphrey, J. E. (2003). Experiential learning in the animal sciences: Development of a multispecies large-animal management and production practicum. *Journal of Animal Science*, *81*, 3202–3210.

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin*, *114*(3), 552–566. https://doi.org/10.1037/0033-2909.114.3.552

Renninger, K. A., & Su, S. (2012). Interest and its development. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 167–187). Oxford, UK: Oxford University Press.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/

Rotgans, J. I., & Schmidt, H. G. (2011). Situational interest and academic achievement in the active-learning classroom. *Learning and Instruction*, *21*(1), 58–67.

Rotgans, J. I., & Schmidt, H. G. (2017). How individual interest influences situational interest and how both are related to knowledge acquisition: A microanalytical investigation. *The Journal of Educational Research*, *111*, 1–11. https://doi.org/10.1080/00220671.2017.1310710

Sithole, A., Chiyaka, E. T., McCarthy, P., Mupinga, D. M., Bucklein, B. K., & Kibirige, J. (2017). Student attraction, persistence and retention in STEM programs: Successes and continuing challenges. *Higher Education Studies*, *7*(1), 46. https://doi.org/10.5539/hes.v7n1p46

Steiger, J. H. (1989). *Causal modeling: A supplementary module for SYSTAT and SYGRAPH*. Evanston, IL: SYSTAT.

Stets, J. E., & Burke, P. J. (2000). Identity theory and social identity theory. *Social Psychology Quarterly*, *63*(3), 224–237. https://doi.org/10.2307/2695870

Suresh, R. (2006). The relationship between barrier courses and persistence in engineering. *Journal of College Student Retention: Research, Theory & Practice*, *8*(2), 215–239. https://doi.org/10.2190/3QTU-6EEL-HQHF-XYF0

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education.

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). New York: Guilford Press.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*, 817–830.

Willett, J. B. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth. *Educational and Psychological Measurement*, *49*, 587–602. https://doi.org/10.1177/001316448904900309

How to cite this article: Erickson MG, Wattiaux MA, Karcher EL. Longitudinal measurement invariance and stability of individual interest across a 16-week introductory animal sciences course. *Nat Sci Educ*. 2020;49:e20031. https://doi.org/10.1002/nse2.20031