"It's More Of A Me-Thing Than An Evolution Thing": Exploring The Validity Of Evolution Acceptance Measures Using Student Interviews

Taya Misheva,[†] Sara E. Brownell,[†#*] and M. Elizabeth Barnes[‡§]

[†]Biology Education Research Lab, Research for Inclusive STEM Education Center, School of Life Sciences, Arizona State University, Tempe, AZ 85287; [‡]Social Perceptions of Science Lab, Department of Biology, Middle Tennessee State University, Murfreesboro, TN 37132

ABSTRACT

Researchers who study student acceptance of evolution rely on surveys that are designed to measure evolution acceptance. It is important for these surveys to measure evolution acceptance accurately and in isolation from other constructs, so that researchers can accurately determine what leads to low acceptance. The Inventory of Student Evolution Acceptance (I-SEA) and the Generalized Acceptance of EvolutioN Evaluation (GAENE) are two surveys that were developed to improve upon the limitations of earlier surveys. Yet neither survey has been extensively tested for response process validity, which can assess the extent to which students use constructs other than their acceptance of evolution to answer survey items. In this study, we examined the response-process validity of the I-SEA and GAENE by conducting cognitive interviews with 60 undergraduate students. Interviews revealed that both surveys retain certain response-process issues. The I-SEA conflated knowledge about and acceptance of evolution for a subset of students. The GAENE measured evolution acceptance inconsistently because students interpreted "evolution" in different ways; it also measured willingness to advocate for evolution in addition to acceptance. Researchers can use these findings to better inform their survey choice when designing future studies, and to further improve the measurement of evolution acceptance.

INTRODUCTION

What is evolution acceptance and why does it matter?

Evolution is "the central organizing principle of modern biology" (National Academy of Sciences, 2008), and as such has been identified as one of the five core concepts for Biology education (American Association for the Advancement of Science, 2011; Brownell *et al.*, 2014). Despite its central role in biology, evolution remains socially controversial among both the general public (Gallup Inc, 2019 Pew Research Center, 2019; Miller *et al.*, 2022) and among college biology students in the United States (Dunk and Wiles, 2018; Siciliano-Martina and Martina, 2020; Ferguson and Jensen, 2021; Laidlaw *et al.*, 2022). It is important for students to not only understand evolution, but to also accept it as the best scientific explanation for the unity and diversity of life on Earth (Smith, 2010; Sickel and Friedrichsen, 2013; Nadelson and Hardy, 2015). Individuals who reject one or more aspects of evolution are unlikely to apply evolutionary concepts to solve biology-related problems once they leave the classroom (Smith, 2010; Smith and Siegel, 2016). While most undergraduate biology students do not go on to become academic researchers (U.S. Bureau of Labor Statistics, 2019), the applications of evolution are not limited to evolutionary biology research. Both evolutionary processes and evolutionary history are relevant to biomedical research (e.g., selecting appropriate animal models, impact of evolutionary history on patterns of human health and disease), public health (e.g., risks of zoonotic disease transfer),

Erika Offerdahl, *Monitoring Editor*

Submitted Jan 31, 2023; Revised Jul 17, 2023; Accepted Aug 2, 2023

CBE Life Sci Educ December 1, 2023 22:ar41

DOI:10.1187/cbe.23-01-0022

[§]Brownell S. E. and Barnes M. E. are co-senior authors.

*Address correspondence to: Sara E. Brownell (sbrownel@asu.edu).

agriculture (e.g., importance of genetic diversity in crops), and teaching biology to the next generation of students (Rühli and Henneberg, 2013; Sickel and Friedrichsen, 2013; Preuss and Robert, 2014; Nadelson and Hardy, 2015; Grunspan *et al.*, 2018).

While the precise definition of what it means for a student to "accept evolution" has been debated within evolution education research (Smith and Scharmann, 1999; Southerland *et al.*, 2001; Sinatra *et al.*, 2003; Ingram and Nelson, 2006; Smith, 2010; Wiles, 2014; Smith and Siegel, 2016), some common themes and similar definitions have emerged. One common theme is that acceptance of evolution is distinct from understanding of evolution in that understanding of evolution pertains to one's awareness of factual information about evolution, while acceptance is about whether one agrees that the theory of evolution is the best available explanation for the development of life on Earth (Southerland *et al.*, 2001; Sinatra *et al.*, 2003; Ingram and Nelson, 2006; Wiles, 2014; Kuschmierz *et al.*, 2020b). As such, studies have found that some students can accept the theory of evolution as being generally true despite displaying misconceptions about how evolution works (Sinatra *et al.*, 2003; Kuschmierz *et al.*, 2021), while other students can score well on measures of evolution understanding despite rejecting the veracity of evolutionary theory (Sinatra *et al.*, 2003; Ingram ad Nelson, 2006).

Researchers also make a distinction between acceptance of evolution and understanding of the nature of science (NOS). NOS encompasses ideas within the philosophy of science about how to distinguish between scientific vs. nonscientific questions, what constitutes scientific evidence, and the tentative NOS (Smith and Scharmann, 1999; Smith, 2010). While a student's understanding of NOS could potentially influence their acceptance of evolution, these are two conceptually distinct constructs. Most recently, a diverse group of 16 researchers convened to discuss the measurement of evolution acceptance; one product of the meeting was a consensus definition which defines evolution acceptances as, "agreement that evolution is valid and the best explanation from science for the unity and diversity of life on Earth, which includes speciation, the common ancestry of life, and that humans evolved from nonhuman ancestors" (submitted manuscript).

One of the primary uses of measuring student acceptance of evolution is to identify whether and how evolution acceptance is associated with understanding of evolution, understanding of NOS, and other potentially related constructs. Information about such correlations can be used to inform the design of instructional strategies that aim to increase students' evolution acceptance. However, how evolution acceptance is measured may influence results and lead to inconsistencies across studies. For example, studies conducted with preservice and in-service science teachers and university science students have found moderate to strong correlations between understanding and acceptance of evolution (Rutledge and Warden, 2000; Trani, 2004; Stanisavljevic *et al.*, 2013; Dunk *et al.*, 2017), while other studies with preservice science teachers, secondary school students, and university students in various majors found weak correlations (Deniz *et al.*, 2008; Cavallo *et al.*, 2011; Rissler *et al.*, 2014). Some studies with in-service science teachers and university students in various majors found no significant correlation between acceptance and understanding of evolution (Brem *et al.*, 2003; Sinatra *et al.*, 2003; Athanasiou *et al.*, 2016; Gefaell *et al.*, 2020). Additionally, research shows that understanding and acceptance of evolution may display greater correlation in populations with a higher overall understanding of evolution (e.g. university students) than populations with generally lower levels of understanding (e.g. middle-grade students) (Kuschmierz *et al.*, 2020a; Kuschmierz *et al.*, 2020b). Similarly, research conducted with preservice science teachers and university science students indicates that evolution-acceptance is correlated with understanding of NOS (Rutledge and Warden, 2000; Dunk *et al.*, 2017), whereas other studies with preservice and in-service science teachers have failed to increase acceptance by increasing understanding of NOS (Coleman *et al.*, 2015; Cofré *et al.*, 2017). Such inconsistencies can make it difficult for researchers to identify the most promising approaches for increasing evolution acceptance.

How can measurement of evolution acceptance influence study results?

Population differences between studies can contribute to differences in evolution acceptance research findings, but inconsistency in the measurement of evolution acceptance also plays a role. Before the 2000s, evolution education researchers had no standardized measure of evolution acceptance. Researchers developed unique survey tools that were often used in a single study; these surveys differed in item wording, number of items, and range of answer choices for each item (Lawson, 1983; Zimmerman, 1987; Koevering and Stiehl, 1989). These differences in measurement limited researchers' ability to compare findings across studies. The 1999 publication of the Measure of Acceptance of the Theory of Evolution (MATE) was an important step forward in evolution education research because it provided the education research community with a peer-reviewed, publicly available survey tool supported by multiple forms of validity evidence (Rutledge and Warden, 1999). The MATE proved to be a popular tool, as usage of the MATE in education research rapidly grew following its publication (Mead *et al.*, 2019; Kuschmierz *et al.*, 2020b; Barnes *et al.*, 2022).

After the MATE was published, researchers started to voice concerns about how accurately it measured students' acceptance of evolution (Barnes *et al.*, 2019; Smith, 2010; Nadelson and Southerland, 2012; Sickel and Friedrichsen, 2013; Romine *et al.*, 2018). These concerns led to the development of other measures of evolution acceptance, namely the Inventory of Student Evolution Acceptance (I-SEA; Nadelson and Southerland, 2012) and the Generalized Acceptance of EvolutioN Evaluation (GAENE; Smith *et al.*, 2016). Like the MATE, the I-SEA, and the GAENE are multi-item survey tools that employ a five-point Likert scale ranging from "strongly agree" to "strongly disagree." Both were designed for general use by instructors and education researchers, and published in stand-alone articles with accompanying validity evidence.

The I-SEA sought to improve upon the original MATE by dividing up the measurement of evolution acceptance into three subscales for microevolution, macroevolution, and human evolution (Nadelson and Southerland, 2012). The authors made these distinctions based on prior research showing that students perceive differences between micro- and macroevolution, and between human and nonhuman evolution, even though these concepts are not biologically distinct (Reznick and Ricklefs, 2009;

Nehm and Ha, 2011). Smith and colleagues (2016) then sought to improve upon both the MATE and the I-SEA by creating the GAENE 2.1, a new measure based on an explicit definition of evolution acceptance as, "The mental act or policy of deeming, positing, or postulating that the current theory of evolution is the best current available scientific explanation of the origin of new species from preexisting species." Four years after its publication, the GAENE 2.1 was updated into the GAENE 3.0 in an effort to make the instrument more psychometrically robust by eliciting a broader distribution of scores (Glaze *et al.*, 2020).

Although the development of new instruments can lead to improvements in the measurement of evolution acceptance, it can also create novel challenges. Though the MATE, I-SEA, and GAENE 3.0 share similarities, these survey tools each consist of a unique set of items that focus on slightly different aspects of evolution acceptance (Rutledge and Warden, 1999; Nadelson and Southerland, 2012; Smith *et al.*, 2016; Glaze *et al.*, 2020). Even though these survey tools are intended to only measure evolution-acceptance, there is emerging evidence that some items conflate evolution acceptance and other constructs such as understanding of evolution and understanding of NOS (Southerland *et al.*, 2001; Smith, 2010; Dunk *et al.*, 2017). Two of our prior studies suggest that this could be happening. In one study, we found that the MATE frequently measures not only evolution-acceptance, but also other things like understanding of evolution, understanding of NOS, and perception of scientists' views on evolution (Barnes *et al.*, 2022). In another study, we found that results on the MATE, I-SEA, and GAENE 2.1 differ from each other even when the same students take all three surveys (Barnes *et al.*, 2019). This suggests there may be systematic differences in: (1 how the surveys conflate evolution acceptance with other constructs, (2 the extent to which they conflate acceptance with other constructs, and (3 which other constructs each survey measures. This is a cause for concern because such differences can bias results in studies that seek to examine the relationship between evolution acceptance and other potentially related constructs. These are, in short, questions of survey validity.

What is survey validity, and how can gaps in validity evidence influence the measurement of evolution-acceptance?

Survey validity addresses the question of whether a survey truly measures what researchers intend it to measure. Validity is not a static property of the survey. Rather, "validity" refers to the accuracy of the inferences that can be drawn from the survey results when used within certain contexts and populations (Messick, 1995; American Educational Research Association *et al.*, 2014). For the I-SEA and GAENE (2.1 and 3.0), each survey's validity is the extent to which the survey scores enable researchers to make accurate inferences about the evolution-acceptance of the students who completed the survey. Validity is context-dependent; for example, responses on the I-SEA would presumably be more valid for university students in the U.S. (a population for which it was developed) than for secondary students in a non-English speaking country. There is even the potential problem of consequences of testing that can happen if a survey is deployed in a sufficiently different context with cultural factors that may interfere with drawing accurate interpretations from the responses (American Educational Research Association *et al.*, 2014; Beniermann *et al.*, 2022).

Validation evidence is gathered by assessing content validity, internal-structure validity, external-structure validity, and response-process validity (Campbell and Nehm, 2013; American Educational Research Association *et al.*, 2014; Artino *et al.*, 2014; Mead *et al.*, 2019). Content validity addresses the extent to which a survey presents a complete and accurate representation of the relevant knowledge domain; it is typically evaluated using expert review (American Educational Research Association *et al.*, 2014). In this case, content validity would refer to the scientific accuracy and theoretical relevance of survey items. Internal-structure validity addresses the extent to which individual items on the survey are all measuring the same construct; it is typically evaluated using quantitative methods such as factor analysis or Rasch analysis (Campbell and Nehm, 2013; American Educational Research Association *et al.*, 2014). In this case, internal-structure validity would assess whether all survey items measure evolution acceptance – or the intended subparts of evolution acceptance – or if some items primarily measure some other construct. External-structure validity addresses the extent to which a survey displays the expected relationships with other measures; one way to evaluate it is to look for correlations with other surveys designed to measure the same or similar constructs (Campbell and Nehm, 2013; Mead *et al.*, 2019). In this case, external-structure validity can be assessed by administering the GAENE or the I-SEA together with the MATE to determine the extent of correlation between the surveys' scores. These three forms of validity evidence have been gathered for the I-SEA, GAENE 2.1, and GAENE 3.0 both during their initial development (Nadelson and Southerland, 2012; Smith *et al.*, 2016; Glaze *et al.*, 2020), and in later validation studies (Romine *et al.*, 2018; Sbeglia and Nehm, 2018, 2019; Barnes *et al.*, 2019).

Response-process validity addresses the extent to which participants interpret survey items in the way that researchers intended. Process validity is violated when a participant selects answers based on reasons other than what the researchers intended, and can indicate that an item is measuring extraneous information other than the targeted construct (American Educational Research Association *et al.*, 2014; Artino *et al.*, 2014). Response-process validity is assessed through cognitive interviews, in which participants "think aloud" as they reason through why they answered a survey item in a particular way (Willis, 2004; García, 2011). The GAENE 2.1 received a limited assessment of response-process validity during its development; the initial draft of the survey received written feedback from 26 high school students, as well as verbal feedback from five high school and four university students during interviews. Feedback and discussion largely focused on various aspects of item clarity, and qualitative results from this feedback was not reported in the publication (Smith *et al.*, 2016). While data from student interviews was used to generate items for the I-SEA, neither the I-SEA nor the new items on the GAENE 3.0 have been assessed for response-process validity using cognitive interviews (Nadelson and Southerland, 2012; Glaze *et al.*, 2020).

Our recent study on the original MATE demonstrated that cognitive interviews with students can reveal substantial issues with measuring extraneous constructs that other forms of validation have been unable to clearly detect. We found that even though the MATE was designed to only measure acceptance of

evolution, it measured understanding of evolution, understanding of NOS, and students' perceptions of scientist views on evolution in addition to their levels of personal acceptance of evolution. We also found that interpretations of the word "evolution" varied between students whenever items use the term without specifying species or context (Barnes *et al.*, 2022). Educational research standards and our own recent findings with the MATE thus indicate that assessing response-process validity is an essential component in gathering validity evidence for the I-SEA and the GAENE, if researchers want to continue to use these instruments to measure evolution acceptance.

Current Study

Given that cognitive interviews can identify response process issues that other forms of validation testing are less suited to detect (Willis, 2004; García, 2011; American Educational Research Association *et al.*, 2014), we contend that process validity testing via cognitive interviews with students is an essential step in establishing validity evidence for any measure of evolution acceptance. Thus, the goal of the current study was to explore the process-validity of two recent instruments that have been developed to measure evolution acceptance; 1) the I-SEA and 2) the GAENE (2.1 and 3.0) – using student interviews in order to identify what process validity issues (if any) are present on these instruments.

METHODS

The goal of this study was to explore the response process validity of the I-SEA and the GAENE (2.1 and 3.0) using cognitive interviews with undergraduate students. Each interview consisted of a one-on-one cognitive interview using one of the survey tools, followed by several open-ended questions about the student's views on evolution and a brief demographic survey (Supplemental Materials). We conducted three separate rounds of interviews, with different students in each round (see Table 1 for summary). The first round occurred in Fall 2020; all participants were recruited from an upper-level Biology course at a single institution and took either the I-SEA or the GAENE 2.1. The second round occurred in Spring 2021; participants were recruited from several public universities across the U.S. and took either the I-SEA or the GAENE 2.1. The third round occurred in Fall 2021; participants were recruited from an upper-level Biology course at one institution and an introductory level Biology course at another institution; these students were interviewed on the new items added to the GAENE 3.0. Below we provide more detailed descriptions of the measures used, recruitment methods, interview protocols, and data analysis.

Survey Tools

The I-SEA is a 24-item survey that is partitioned into three subscales for; 1) macroevolution, 2) microevolution, and 3) human evolution. Each subscale consists of eight forward- and reverse-coded items scored on a five-point Likert scale ranging from "strongly agree" to "strongly disagree" (Nadelson and Southerland, 2012). An item is "forward-coded" when agreement indicates evolution acceptance (strongly agree = five) and "reverse-coded" when agreement indicates evolution rejection (strongly agree = zero). The subscale categories were not labeled on the student version of the survey.

TABLE 1. Summary of data collection, including number and timing of student interviews.

Instrument	Number of Items	Semester collected	Interviews
I-SEA	24	Fall 2020, Spring 2021	22
GAENE 2.1	13	Fall 2020, Spring 2021	17
GAENE 3.0 (new items)	10	Fall 2021	21

The GAENE 2.1 is a 13-item survey intended to measure evolution acceptance as a single construct. The GAENE 2.1 is also scored on a five-point Likert scale; however, all the items are forward-coded (Smith *et al.*, 2016). The GAENE 3.0 is a 22-item survey. It contains 12 items from the GAENE 2.1, one item that was removed during the initial development of the GAENE 2.1 but added back in for the GAENE 3.0, and nine entirely new items (Glaze *et al.*, 2020). The first two rounds of interviews used the GAENE 2.1 because the GAENE 3.0 was not yet published when data collection began. The third round of data collection focused exclusively on the items that were on the GAENE 3.0 but not the GAENE 2.1.

From here on, we will report participants' composite scores on whichever survey they completed. A composite score is the participant's average answer across items scored on a Likert scale; for these measures, 5.0 indicates maximum evolution acceptance and 1.0 indicates minimum acceptance. Unlike a total score, a composite score enables simple direct comparison between surveys with different numbers of items.

Recruitment

In the first round of data collection, we recruited 17 students from an upper-level Biology course for majors at a research-intensive public university in the southwestern U.S. during the Fall 2020 semester. These participants received extra credit worth one daily assignment grade in the course as an incentive for participation. Though religiously diverse, the majority of students recruited from this course exhibited high levels of evolution acceptance and had taken five or more college-level Biology courses, so they had fairly strong Biology and evolution backgrounds (see Supplemental Materials for demographic survey).

In the second round of data collection, we sought to expand the diversity of our sample by sending individual emails to students who received low scores on other measures of evolution acceptance as part of a separate study exploring students' levels of evolution acceptance. This second set of 22 additional participants came from a nationwide sample of students at public universities. Students recruited in this manner received a $15 Amazon gift card for participation.

We conducted the third round of data collection to assess the response-process validity of items that were present on the GAENE 3.0 but not on the GAENE 2.1. During the Fall 2021 semester, we recruited 21 students from two courses: 1) an upper-level Biology course at the first-round institution, and 2) an introductory-level Biology course at a public R2 university in the southeast. Students who participated were offered either extra credit worth one daily assignment grade or a $10 Amazon gift card.

Cognitive Interviews

We conducted 22 interviews with the I-SEA, 17 with the GAENE 2.1, and 21 with the new items on the GAENE 3.0 (Table 1). During the Fall 2020 and Spring 2021 rounds of data collection, interviews alternated between the GAENE 2.1 and the I-SEA, such that each instrument was tested in both rounds. During the cognitive interviews, participants read each item from the given instrument out loud, selected an answer out loud, and explained why they selected the answer that they chose as opposed to the other answers available to them (García, 2011). At the end of each interview, the interviewer also asked a set of free-response questions that addressed the student's acceptance of various aspects of evolution, including macroevolution and human evolution (see Supplemental Materials for questions). The purpose of these free-response questions was to give participants the opportunity to describe their views on evolution in their own words and potentially clarify any inconsistencies across their interview. T.M. conducted and recorded all interviews virtually via Zoom.

Students were asked to fill out a brief demographic survey after the interview (See Supplemental Materials for a copy of the survey). Given the qualitative nature of this study, the purpose of collecting demographic information was not to use it for data analysis, but to track the diversity of our sample. The survey contained questions on religiosity and religious affiliation to help us include students with a variety of religious perspectives. To check whether the sample contained students with different levels of evolution-education, the survey also asked how many college-level Biology courses they had taken and whether any of these courses had been primarily about evolution. This was not intended to be a direct measure of students' knowledge about evolution, but a proxy of their prior exposure.

Data Analysis

To identify any response process issues with how students answer items on the GAENE and the I-SEA, we qualitatively analyzed the cognitive interviews using a combination of deductive and inductive coding (Cho and Lee, 2014; Krippendorff, 2018). To enable direct comparison between evolution acceptance instruments, student responses were initially coded using a deductively developed, relatively broad codebook with codes based on prior critiques of evolution acceptance instruments (Smith, 2010; Nadelson and Southerland, 2012; Sickel and Friedrichsen, 2013; Romine *et al.*, 2018; Barnes *et al.*, 2019) and process issues that we previously identified on the MATE (Barnes *et al.*, 2022). The deductive codebook included codes to be applied whenever a student answered an item based on either (1 their factual understanding of evolution, (2 their understanding of NOS, (3 defining "evolution" in a way that explicitly excludes human or macroevolution when the item itself does NOT specify microevolution alone, (4 their perception of scientists' views on evolution (rather than their own views), or (5 misalignment between an item that assumes Biblical creationism and their own religion's creation account. We included these five codes in the initial codebook because these were the response-process issues that we had previously found when we conducted cognitive interviews on the original MATE. However, we understood that some of these response-process issues may not arise in the present study, and included an "other" code for any novel-process issues. This codebook was used by the interviewer as a tool for structuring notes during the interview process.

Because we wanted to identify all potential-process issues, we proceeded to inductively code the interview data after the interviews were complete. We developed an inductively derived codebook by listening to each interview recording, assigning a detailed new code whenever a student made a novel process issue, and conducting a constant-comparison analysis in which each student's process issues were compared with existing codes to determine whether an existing code is applicable or whether a new code is warranted (Cho and Lee, 2014). During this process, student responses were analyzed and further broken down into inductively derived subcodes of the existing deductive codes, and entirely new codes were developed for novel process issues initially labeled as "other." For example, researchers inductively analyzed student responses coded as "understanding of evolution" to identify the subcodes "low certainty about own understanding of evolution" and "misconceptions about evolution" (See Supplemental Material for the final full codebook).

After T.M. coded all interviews a second researcher used the codebook to independently code 10% of the interviews. A comparison of the codes assigned by the two researchers yielded an acceptable level of interrater agreement (Cohen's kappa = 0.77).

The following results include quotes from students in the study; names have been changed to protect identity and some quotes have been lightly edited for clarity. The Institutional Review Board of Arizona State University approved the procedures for this study (ASU IRB #00010903).

RESULTS

Participants

We interviewed a total of 60 students for this study. Table 2 displays a summary of our participants in terms of gender, race/ethnicity, religious affiliation, academic year, major, and prior-evolution exposure. Prior-evolution exposure was classified as "high" for those who have taken a course focused primarily on evolution, "medium" for those who have taken ≥ three Biology courses but no course on evolution, and "low" for those who had taken ≤ two Biology courses and no course on evolution. For gender, nonbinary and fill-in-the-blank options were provided but not selected by any of the participants. The "GAENE" column shows a combined sample of students who interviewed on the GAENE 2.1 and students who interviewed on the new items from the GAENE 3.0. Three participants did not fill out the demographic form.

I-SEA Finding 1: Students struggle to answer items on the I-SEA when they lack knowledge about evolution.

Students struggling to answer items due to their limited understanding of evolution was the single-most common process validity issue on the I-SEA. While this process issue arose on at least one item for most (91%) participants, it was largely clustered with particular students and on particular items.

We found that four students – Anemone, Marinus, Rio, and Tethys – struggled with understanding of evolution across the entire I-SEA instrument. Students were categorized as "struggling with understanding of evolution" when they had knowledge-related process issues on six or more items, while all other students had this process issue on three items or less. On

average, these four students had knowledge-related process issues on over a third (39%) of their answers on the I-SEA. It is worth noting that three of the four had low prior exposure to college-level evolution instruction, while the fourth did not fill out the demographic questionnaire.

Marinus' responses were a good example of how a student's limited understanding of evolution can impact survey validity. He described fully accepting evolution in the open-ended response and received average composite scores of 4.1 (macro), 4.9 (micro), and 4.6 (human). However, Marinus' responses were at times affected by misconceptions about evolution. For example, his answer for Item 18: Although humans may adapt, humans have not/do not evolve revealed the misconception that evolution occurs at the level of individuals rather than populations, and that individuals that fail to reproduce play no part in the evolution of a species. This misconception directly influenced his answer choice:

> Marinus (disagree): "I would disagree with this simply because our ancestors were used to living in caves, which was vastly different compared with now. (I didn't pick strongly disagree) because if the ones that do not evolve and eventually die out, then technically those individuals don't get to evolve and reproduce."

Another student with many knowledge-related process issues was Tethys, who likewise expressed full acceptance of evolution during the open-ended interview and received average composite scores of 3.9 (macro), 4.0 (micro), and 4.1 (human). They chose "undecided" rather than "strongly agree" for Item 4: I think all complex organisms evolved from single celled organisms:

> Tethys (undecided): "I'm (going to) say undecided because I'm not too informed on single-celled organisms to know whether (complex organisms) evolved from them or not. Just (without) knowing I would think so, but then again, I'm not sure."

While a majority of the knowledge-related process issues were concentrated in the four students discussed above, other students' one or two knowledge-related process issues were concentrated in a subset of items. Namely, of the remaining 18 students who were generally confident in their knowledge about evolution, four students (22%) had knowledge-related process issues on Item 14 and six students (33%) had these process issues on Item 24.

Item 14: I think there is an abundance of observable evidence to support the theory describing how variations within a species can happen.

This item revealed uncertainty about the extent of the "abundance of observable evidence", even for students who were otherwise confident in their knowledge. For example, Azure described accepting microevolution and speciation within closely related taxa and provided answers that were largely consistent with this view (composite score of 4.9 for microevolution). The one exception was Item 14, for which she said:

> Azure (agree): "(I don't strongly agree because) we haven't gotten too far into this topic in my Biology class. I feel like I'd need to see a bit more observable evidence to strongly agree. I need a little bit more in depth research about it. I'd need to learn more evidence in my class."

Item 24: Physical variations in humans (i.e., eye color, skin color) were derived from the same processes that produced variation in other groups of organisms.

Item 24 revealed low confidence in personal knowledge, misconceptions, and conceptual mix-ups among participants. One example of a mix-up comes from Ariel, who had an average composite score of 3.0 on the human evolution subscale despite saying that she believes God created humans in their present form. Part of the reason for Ariel's unexpectedly high score comes from her misinterpretation of what the item is referring to:

> Ariel (strongly agree): "Yes, I'd strongly agree. I'm thinking how eye color and skin color (are) all genetic. So, I'm thinking it's within the DNA and I don't think it has anything to do with evolution per se, because it's not like we're evolving blue eyes. That stuff we inherit from our parents."

Ariel interpreted Item 24 as genetic processes related to inheritance and the central dogma, rather than to evolutionary processes such as natural selection or genetic drift. This indicates that the overall scientific complexity of this item can leave students confused about what "processes" the item refers to, which leads to an answer that reflects the student's views on a different topic.

I-SEA Finding 2: Several items on the I-SEA consistently produce process issues due to unclear wording of the items

Validity issues with individual items arose when multiple students with varying views and social identities answered items based on difficulties related to item wording, rather than on their own acceptance or rejection of evolution. Below, we describe items that exhibited process-validity errors for at least 15% of participants. Given the typically small sample sizes in cognitive interview studies, 15% has been used as a standard cutoff for differentiating between potential-validity issues vs. random errors (Nápoles-Springer *et al.*, 2006).

Item six: There is little or no observable evidence to support the theory that describes how one species of organism evolves from a different ancestral form.

In addition to the several students who expressed low certainty about their knowledge of the observable evidence, an additional four students (18%) struggled to interpret the phrase "different ancestral form." Two examples of this come from Moana, who said that she fully accepts evolution, and Ocean, who said that she accepts human macroevolution but believes that higher taxa such as mammals and insects do not share a common ancestor.

> Moana (undecided): "This question doesn't make much sense. What do you mean by 'different ancestral form'? Different from what?"

> Ocean (undecided): "I don't really understand this question.… Maybe you mean an ancestral form that looks different from (the species in question)? Like a fish and a tiger? Or it could

TABLE 2. Participant demographics, evolution education, and evolution acceptance scores (n = 60).

Demographic Variable	GAENE 2.1 + 3.0	I-SEA	Demographic Variable	GAENE 2.1 + 3.0	I-SEA
Gender Identity			**Religious Affiliation**		
Man	26.3%	36.3%	None	31.6%	22.7%
Woman	68.4%	59.1%	Christian	47.4%	50%
No answer[1]	5.3%	4.5%	Hindu	2.6%	9.1%
Race/Ethnicity			Muslim	5.3%	9.1%
Asian/Asian American	15.8%	22.7%	Other	5.3%	4.5%
Black/African American	15.8%	13.6%	No answer	7.9%	4.5%
Hispanic/Latinx	7.9%	18.2%	**Major**		
Native American	2.6%	0%	Biology	65.8%	45.4%
White	42.1%	22.7%	Other STEM	21.1%	40.9%
More than one	10.5%	9.1%	Other non-STEM	7.9%	9.1%
No answer	5.3%	13.6%	No answer	5.3%	4.5%
Academic Year			**Prior Evolution Exposure**		
Lower-level	52.6%	36.3%	High	28.9%	36.3%
Upper-level	39.5%	76.5%	Medium	34.2%	22.7%
No answer	5.3%	4.5%	Low	31.6%	36.3%
			No answer	5.3%	4.5%

[1]Gender is a spectrum, but no participants identified as nonbinary.

mean like, the connection between humans and apes? I don't really know what is being asked."

Item seven: The forms and diversity of organisms have changed dramatically over time, and
Item 16: There is overwhelming evidence supporting the theory of evolution to explain how variations in a species develop over time.

For Items seven and 16, four students (18%) on each item avoided the "agree" or "strongly agree" options due to these items' use of strong adjectives such as "dramatically" and "overwhelming."

Two examples come from Triton, who said that he's undecided about his views on human evolution and the shared ancestry of all life, and Anemone, who said that she fully accepts evolution.

> Triton (Item seven, undecided): "I'm not sure. Like, are we talking about when the earth started, or a thousand years? Is it asking if an organism has the ability to change dramatically? I'm not too sure what is supposed to be changing dramatically."

> Anemone (Item 16, agree): "I would say only 'agree' and not 'strongly agree' just because of the word 'overwhelming.' Like, I don't know, just certain words…I'm not really sure what is meant by the word 'overwhelming' in this case."

Overall, these findings suggest that while process-validity issues are neither especially numerous nor especially frequent on the I-SEA, there is nevertheless room for improvement in terms of scientific complexity and overall item clarity.

GAENE Finding 1: The GAENE can overestimate evolution-acceptance for students who use an incomplete definition of evolution

While the I-SEA delineates between micro, macro, and human evolution, most items on the GAENE use the terms "evolution" or "evolutionary theory" without the same specification. This leaves room for survey takers to potentially use an incomplete definition of evolution that excludes more controversial concepts such as our shared ancestry with primates (human macroevolution) and the shared ancestry of distantly related higher taxa (such as mammals and cephalopods). An examination of student explanations in response to both the individual items and the open-ended questions reveals that 71% of students who described not fully accepting evolution used an incomplete definition of evolution at least once. For students who used an incomplete definition of evolution at least once, doing so influenced their answer choices on approximately one third (34%) of the items, on average.

One example of this trend comes from Ariadne, a Baha'i student who described accepting speciation but rejecting both human macroevolution and the shared ancestry of higher taxa:

> Ariadne (open-ended interview questions): "I think there were multiple starting species, which had babies and evolved. We definitely didn't start out with one (ancestral species). I'm thinking there were probably hundreds of starting species. I'm thinking maybe like, the big cats evolved from a similar ancestor and maybe a lot of rodents (share a common ancestor)."

Yet despite these self-described views, Ariadne selected agree or strongly agree for eight items on the GAENE 2.1, resulting in a composite score of 3.5. We can see an example of how this occurred in Ariadne's reasoning for the item, "Nothing in Biology makes sense without evolution," in which she described how excluding most forms of macroevolution from her definition of "evolution" led her to agree with this item:

> Ariadne (agree): "(When the item refers to 'evolution') I'm thinking of something on the border between micro and macro. I think the majority of it just stems back to natural selection. I guess that's more of a microevolution idea."

Another example comes from Cassandra, a Southern Baptist student who likewise described rejecting both human macroevolution and the shared ancestry of higher taxa. Cassandra selected strongly agree for the new GAENE 3.0 item, "All

evidence supports the claim that evolution is true," with the following explanation (emphasis ours):

> Cassandra (strongly agree): "I do believe that we all have to come from something or someone. (Interviewer: What concepts are you associating with the word "evolution?") I do believe that some species of animals have evolved from other species of animals, and I do believe that humans have evolved since the stone age. Species themselves evolve, but I don't believe that we all evolved from one species."

The quote above demonstrates that when answering this item, Cassandra strongly agreed because she was using a definition of evolution that includes microevolution and some macroevolution for a subset of species, yet excludes the shared ancestry of life. Yet this is one of the new GAENE items that was designed to differentiate between those with a "high" vs. a "very high" level of acceptance, which suggests that limiting "evolution" to microevolution and limited macroevolution was not the survey authors' intent.

As these examples show, students who do not fully accept evolution tend to define "evolution" in a way that includes the ideas that they agree with and excludes the ideas that they disagree with. Doing so, results in a score that may be higher than we would expect for students who do not accept human evolution and/or macroevolution.

GAENE Finding 2: Several items on the GAENE consistently produced process issues because students frequently used constructs other than their own acceptance of evolution

Validity issues with individual items arose when multiple students with varying views and social identities answered particular items based on factors other than their own acceptance or rejection of evolution. These factors can be categorized as emotions and behaviors pertaining to evolution, such as willingness to advocate for evolution and emotional attachment to evolution. We describe how these factors influenced students' answers for particular items below.

Item 1: Everyone should understand evolution.

Half (50%) of the students who said that they fully accept evolution stated that while they personally accept evolution, they do not think that it is essential for evolution to be taught to those who do not wish to learn it. Two students who expressed this view were Helen and Jocasta. Both said that they fully accept evolution and had composite scores of 4.7 and 4.3, respectively, yet both selected disagree for this item.

> Helen (Item one, disagree): "Although I personally think it's an important topic, I do not think it should be required of everyone and do not think that people who do not hold scientific views, but instead hold the religious views, I don't think that they should need to understand evolution."

> Jocasta (Item one, disagree): "In my personal opinion, yeah, I strongly agree. But in a more realistic manner, (given) that some people are more religious, I'm gonna say disagree because that (would be like) somebody who's majoring in physics telling me that I have to understand physics."

As these quotes demonstrate, a low score on this item does not necessarily indicate that a student personally rejects evolution. Rather, some students are simply less insistent than others when it comes to teaching evolution to reluctant individuals.

Item 6: I would be willing to argue in favor of evolutionary theory in a public forum such as a school club, church group, or meeting of public-school parents.

A similar conflation between evolution acceptance and a person's behaviors and emotions can be seen for Item six. We found that about 59% of all students (and 75% of those who fully accept evolution) took into account the potential for social stress in the scenario when selecting their answer, which tended to reduce agreement. This stands in contrast to Item 10 (I would be willing to argue in favor of evolution in a small group of friends), for which no student indicated that social stress would impact their answer. Penelope's responses to these two items illustrate the unique features of Item six (emphasis ours):

> Penelope (Item six, strongly disagree): "The word 'argue' kind of gave me a negative connotation. I wouldn't necessarily argue in favor of evolutionary theory, but I would slightly debate, (though) not in a violent or harsh way. It seems a little controversial to talk about in a church or meeting of public-school parents."

> Penelope (Item 10, agree): "(This scenario is) a small group of friends so it makes it more realistic. Also, I would argue in favor of evolution in front of my friends because it seems like a safe space to talk about evolution and kind of just learn from them, as opposed to the church group."

As we can see from the contrast between Penelope's two answers, her choice to strongly disagree with Item six was not based on the extent of her personal acceptance of evolution, or even on her willingness to present an argument in favor of evolution under certain circumstances. Instead, her answer choice reflected the perceived hostility of the public forum scenario. These findings indicate that Item six may consistently underestimate students' acceptance of evolution, particularly for students who prefer to avoid confrontation or feel unsafe in certain spaces.

Item 17: Evolution is the most important theory devised by man.

100% of students who fully accept evolution stated that while the theory of evolution is both important and true, there are other major scientific theories that are equally or more important for science and/or society. In fact, only two students from the entire sample selected agree for this item, and not a single person selected strongly agree. The responses of Pandora and Eurydice, two agnostic students who fully accept evolution, show why:

> Eurydice (strongly disagree): "I can't think of any other theories that are extremely important, but personally I don't think that learning how we came to be is the most important thing we've ever done. I'm sure there have been other theories in Science, particularly in medicine, that have helped us more."

> Pandora (disagree): "I think that it's important to understand our origins, but I don't think that's the most important theory. I would say bigger overarching theories, like the theory of

> relativity, that help us understand (the universe) as a whole (are more important)."

These quotes show that our participants' near-universal lack of agreement with this item does not reflect a lack of agreement with any evolutionary concept; rather, it indicates that these students do not think that the theory of evolution is inherently more important than other well-established scientific theories.

Item 18: I would bet my life on the claim that evolution is true.

Just over half (56%) of students who fully accept evolution said that it is simply not in their personality to bet their life on any theory, no matter how well supported. The responses of Callisto and Antigone illustrate this mindset (emphasis ours):

> Callisto (undecided): "I'll have to go with undecided because I'm not going to bet my life on anything. But I do believe that evolution is true."

> Antigone (disagree): "I'm going to disagree with that one only because I would bet my life on very little. It's probably more of a me thing than an evolution thing."

Both Callisto and Antigone explicitly stated that their answer choices do not reflect their views on evolution but instead reflect their risk-averse personalities. The notion of betting their life on any idea gave them pause.

Item 19: Understanding evolution has changed my life.

Just under half (44%) of students who reported that they fully accept evolution said that their understanding of it has not changed their life as a whole. This trend took two main forms. One subset of students said that while they recognize the importance of evolution within Science, this theory has had little impact on their personal, day-to-day life outside of Science. One such student was Callisto, a nonreligious student who described fully accepting evolution:

> Callisto (disagree): "I believe that (evolution) is important and true, but it hasn't impacted my life in any type of way."

Another subset of students said that they have never deliberately rejected evolution, so learning about it did not produce a change in their acceptance of evolution – they simply moved from a state of ignorance to a state of knowledge about the topic. One such student was Alcyone, a spiritual student who also fully accepts evolution:

> Alcyone (disagree): "I grew up in a religious household, so evolution wasn't something that I thought about much. So, when I learned about it (in) school, I just kind of took it for what it was instead of (experiencing) cognitive dissonance."

GAENE Finding three: Several items on the GAENE were consistently impacted by students' knowledge about evolution and the NOS

We found that several items on the GAENE consistently test students' factual knowledge about evolution and NOS, such that their answers may not reflect their general acceptance of the evolutionary concepts in question.

Item 12: Evolution is a scientific fact and Item 16: Evolution is a fact.

For students who described fully accepting evolution, 63% of responses on Item 12 and 19% of responses on Item 16 were affected by students' understanding of NOS, namely the knowledge that scientific theories are categorically different from scientific facts. Evolution includes both fact and theory. The existence of evolutionary change is an observable fact, while the theory of evolution explains the mechanisms that drive this change (Branch and Mead, 2008; Lenski, 2020). However, students frequently either did not know or forgot about evolutionary change as a scientific fact, and focused exclusively on the theory of evolution when answering these items. Examples of this come from Jocasta and Circe:

> Jocasta (Item 12, strongly disagree): "From what I've learned so far in every Biology class I've taken, there is no such thing as a scientific fact."

> Circe (Item 16, undecided): "Evolution is a theory, so I feel like 'fact' is an inappropriate word for that. So maybe undecided? (Evolution) is factual and supported by evidence, but it's more of a theory because a fact is something that you can test in a lab, and you can't necessarily do that with evolution because it's a long-term process."

As these quotes show, Circe and Jocasta both interpreted these items as being only about the theory of evolution, and answered based on the understanding that a scientific theory is, by definition, different from a scientific fact. In particular, Circe's explanation shows that she is aware that evolutionary theory is based on factual evidence, but failed to label the factual evidence itself as "evolution." Yet both of these students described accepting the shared ancestry of all life in the free-response portion of the interview, and did not dismiss evolution as "just" a theory in the colloquial sense when answering these items. Further, Jocasta's response highlighted a potential misconception about there not being any facts in Science; although Science can change, there are some concepts that are so well established that they become factual (Gregory, 2008).

Item 14: All evidence supports the claim that evolution is true and Item 15: All species can be traced back to a single ancestor.

For both Items 14 and 15, just under half (44%) of students who fully accept evolution said that they do not have enough knowledge about evolution to be sure of the scientific accuracy of these statements. Two examples of this trend come from Antigone and Dido, who both described fully accepting evolution despite themselves recognizing that they have some knowledge gaps on the topic (emphasis ours):

> Antigone (Item 14, undecided): "I'm not super educated on the in-depth details of evolution. As far as the basics go, I'd say that the concept as a whole is true, but I would want to do a little more research before I wholeheartedly agree with (this statement)."

> Dido (Item 15, undecided): "I'd probably say undecided just because of my lack of knowledge. Reading this question

> without any evolution or genetics background, I would think that there could be multiple ancestors.…I don't know if they all share one ancestor."

As we can see here, it is possible for students accept the shared ancestry of life in a broad sense without being factually aware of whether all of life traces back to a single ancestor vs. multiple ancestors. It is likewise possible for students to recognize that evolution is a very well-supported theory without being sure of what "all evidence" encompasses and whether all that evidence directly supports evolution. These knowledge gaps pose a validity issue for these items because students' answers reflect uncertainty about factual details, rather than uncertainty about the basic concepts of large-scale shared ancestry or the existence of abundant supporting evidence for evolution.

Item 20: Evolution explains how bacteria that are resistant to an antibiotic can arise in a population exposed to that antibiotic, and Item 21: Evolution explains how careful breeding can produce members of a species that look different from their ancestors.

For both Items 20 and 21, one third (33%) of all students were uncertain about what scenario the item refers to or how that scenario connects to evolution. Students who struggled with these items include Clytemnestra, an agnostic student who fully accepts evolution, and Hippolyta, a protestant student who rejects human evolution and believes that nonhuman species were created at higher taxonomic ranks such as class or kingdom:

> Clytemnestra (Item 20, undecided): "I don't know if it's the question that I don't understand, or if it's tying evolution into bacterial resistance. I would just say the question's unclear to me. (It's unclear how) evolution ties into antibiotic resistance."

> Hippolyta (Item 21, undecided): "I'm not sure what to put for this one because I'm not quite clear on 'careful breeding.'… when I think of 'careful breeding,' I think about a person breeding poodles for more desirable traits, and I'm not clear on how this connects to species looking different from their ancestors. (But) maybe 'careful breeding' means like in nature, where (colorful) male peacocks are more desirable to females."

These quotes show that about one third of students in our study were confused by Item 20 because they are unacquainted with antibiotic resistance as an example of natural selection, while another third of students were confused by Item 21 because they were uncertain about whether "careful breeding" refers to artificial selection or sexual selection, and/or how artificial selection relates to natural evolution.

DISCUSSION

This study sought to explore the process validity of two evolution-acceptance instruments: the I-SEA and the GAENE. Overall, we found several issues with the validity of these instruments that we recommend be addressed if we wish to continue using these instruments to measure student acceptance of evolution.

For the I-SEA, we found that the main response-process validity issue was students struggling with the survey due to their limited understanding of evolution. There were several students who described fully accepting evolution when asked to explain their views in their own words, yet repeatedly selected more neutral answers primarily because they were uncertain about whether certain items align with the scientific consensus, and not because they expressed personal doubt about the truthfulness of familiar scientific information. There were also two items that proved challenging for other students as well because they were either uncertain about what qualifies as "an abundance of observable evidence" (Item 14), or were confused about what biological concept the item is referring to (Item 24). Our finding for Item 14 aligns with a recent study that found this item to display differential item function relative to the rest of the I-SEA (Beniermann *et al.*, 2022). Additionally, there were several other items that may benefit from the use of fewer superlatives or simpler language. One previous study found that the human evolution subscale is not unidimensional, with Items 17, 20, 23, and 24 clustering separately from Items 18, 19, 21, and 22; the authors suggested that these clusters represent macroevolution and microevolution respectively (Sbeglia and Nehm, 2019). While the item set that seemingly represents microevolution did elicit more response-process issues in which students excluded human macroevolution from their interpretation (nine process issues across four "micro" items, vs. three process issues across four "macro" items), no individual item exceeded the 15% threshold for students exhibiting the same process issue on an item.

For the full GAENE 3.0, we found response-process validity issues on half of the items. These validity issues involve the measurement of several constructs other than evolution acceptance, including understanding of evolution (Items 14, 15, 20, and 21), understanding of NOS (Items 12 and 16), and several constructs pertaining to participants' priorities and personality traits, such as the perceived importance of evolution relative to other scientific theories (Item 17) and their willingness to engage in public debate (Item six). Some of the items we found to have validity issues have been flagged as potentially problematic by other studies (Romine *et al.*, 2018; Sbeglia and Nehm, 2018; Beniermann *et al.*, 2022). For example, Romine *et al.* (2018), found Item six to have poor-item fit and hypothesized that, "fear of public speaking is quite common, and it is straightforward to argue that one can display acceptance of evolution without extraversion" (Romine *et al.*, 2018, p. 17). For Item one, these authors likewise identified poor fit and hypothesized that, "the misfit is likely caused by responses from students who accept evolution, but nonetheless do not view it as a necessity for engaging in other courses of study or for advancing one's quality of life" (Romine *et al.*, 2018, p. 16). Meanwhile validity issues with other items are identified here for the first time. For instance, the lack of specificity in many items' use of the term "evolution" allows students who do not fully accept evolution to simply exclude the concepts that they reject from their interpretation of "evolution." This same process validity issue was also present on the original MATE (Barnes *et al.*, 2022). Furthermore, this may help explain findings from a recent study in which self-identified creationists on average received GAENE 2.1 scores that indicate intermediate – rather than low – levels of evolution acceptance (Beniermann *et al.*, 2022).

The GAENE 2.1 item "some parts of evolutionary theory could be true" was deleted from the GAENE 3.0 due to poor fit in the Rasch model (Glaze *et al.*, 2020). We found that two

students described the phrasing of this item as a "trick question" because the statement that only some parts of evolution (vs. all or most) merely could be true (vs. are true), suggests that neither the "agree" nor "disagree" options reflect full acceptance of evolution (likewise for broad rejection of evolution). Meanwhile six students either agreed or strongly agreed with the item based on the stated reasoning that they personally accept some parts of evolution but reject other parts, which makes this item an accurate reflection of their views. This finding supports the decision of Glaze and colleagues to not include this item in the GAENE 3.0.

Limitations

Though we tried to recruit participants with a diverse range of views about evolution, roughly half of the participants who completed the I-SEA and GAENE 2.1, and about three-quarters of the participants who completed the new GAENE 3.0 items expressed full acceptance of evolution when describing their views in their own words. As such, there may be some less-common process issues that are particular to students with a low acceptance of evolution that we failed to detect. Future research with populations with very low evolution acceptance could extend this work and illuminate additional process issues.

Recommendations for Instrument Use and Future Research

The prevalence of process-validity issues with the GAENE 3.0 in this study and original MATE in our prior study (Barnes *et al.*, 2022) illustrate the importance for measures of evolution acceptance to clearly define "evolution" for survey takers. When used without context, this term can be interpreted to include all aspects of evolution, everything except for human evolution, or microevolutionary processes alone. This allows students who do not fully accept evolution to exclude any concepts they reject from their personal definition of evolution, which causes the instrument to overestimate their acceptance. To potentially circumvent this issue, we recommend that instructors and researchers who wish to use the GAENE 3.0 provide students with the definition of evolution acceptance provided in the original publication, "Evolution acceptance is the mental act or policy of deeming, positing, or postulating that the current theory of evolution is the best currently available scientific explanation of the origin of new species from preexisting species" (Smith *et al.*, 2016, p. 1296). The I-SEA largely avoids this validity issue by describing specific evolutionary concepts (e.g., Item 20: "I think that humans and apes share an ancient ancestor.") instead of referring to "evolution" as a whole. The main drawback of describing specific evolutionary concepts is that items can easily become prone to scientific misinterpretation by students (e.g., confusing natural selection for trait heritability) or require the use of specific knowledge about evolutionary processes or history. While this appears to be less of a concern for students with prior exposure to college-level evolution instruction, we find that it can cause issues of conflation between understanding and acceptance for students who may not have learned about evolution recently. As such, we recommend that instructors and researchers consider the general level of evolutionary knowledge in their target population when deciding whether the I-SEA is the best instrument for their purposes.

This study further demonstrates that establishing process-validity evidence via cognitive interviews is an essential step in developing a reliable survey instrument. While asking several students to review an instrument for clarity is a step in that direction, issues with measuring constructs other than the intended construct are liable to slip through review unless systematic interviews are conducted with a diverse population of students. Establishing process validity is all the more important when adding items that are intended to be "very easy" or "very difficult" for survey takers to agree with, as these items may be particularly at risk for being easy or difficult for reasons other than a student's level of evolution-acceptance. Furthermore, we advise researchers who seek to further improve the measurement of student evolution-acceptance to closely examine the alignment between survey items and an instrument's measurement goals. This includes addressing questions such as, "What is our definition of evolution acceptance?", "Does each item align with this definition of evolution acceptance?", and "What views on evolution do we want the minimum and maximum scores on our instrument to reflect?". Only then will we be able to more accurately measure student acceptance of evolution.

ACKNOWLEDGMENTS

We thank the students who took the time to share their thoughts and interpretations of the questions. We thank Madisyn Valverde for her assistance in coding student interview data for this project. We also thank the Arizona State University Biology Education Research Lab for thoughtful discussions about this project. This project was supported by National Science Foundation grants IUSE 1818659 and RCN-UBE 2018558.

REFERENCES

American Association for the Advancement of Science. (2011). *Vision and Change in Undergraduate Biology Education: A call to action*. Washington, DC: American Association for the Advancement of Science.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *The Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Artino, A. R., La Rochelle, J. S., Dezee, K. J., & Gehlbach, H. (2014). Developing questionnaires for educational research: AMEE Guide No. 87. *Medical Teacher*, *36*(6), 463–474. https://doi.org/10.3109/0142159X.2014.889814

Athanasiou, K., Katakos, E., & Papadopoulou, P. (2016). Acceptance of evolution as one of the factors structuring the conceptual ecology of the evolution theory of Greek secondary school teachers. *Evolution: Education and Outreach*, *9*(1), 7. https://doi.org/10.1186/s12052-016-0058-7

Barnes, M. E., Dunlop, H. M., Holt, E. A., Zheng, Y., & Brownell, S. E. (2019). Different evolution acceptance instruments lead to different research findings. *Evolution: Education and Outreach*, *12*(1), 4. https://doi.org/10.1186/s12052-019-0096-z

Barnes, M. E., Misheva, T., Supriya, K., Rutledge, M., & Brownell, S. E. (2022). A Revised Measure of Acceptance of the Theory of Evolution: Introducing the MATE 2.0. *CBE—Life Sciences Education*, *21*(1), ar10. https://doi.org/10.1187/cbe.21-05-0127

Beniermann, A., Moormann, A., & Fiedler, D. (2022). Validity aspects in measuring evolution acceptance: Evidence from surveys of preservice biology teachers and creationists. *Journal of Research in Science Teaching*, *n/a*(n/a). https://doi.org/10.1002/tea.21830

Branch, G., & Mead, L. S. (2008). "Theory" in Theory and Practice. *Evolution: Education and Outreach*, *1*(3). Article 3. https://doi.org/10.1007/s12052-008-0056-5

Brem, S. K., Ranney, M., & Schindel, J. (2003). Perceived consequences of evolution: College students perceive negative personal and social

impact in evolutionary theory. *Science Education*, *87*(2), 181–206. https://doi.org/10.1002/sce.10105

Brownell, S. E., Freeman, S., Wenderoth, M. P., & Crowe, A. J. (2014). BioCore Guide: A Tool for Interpreting the Core Concepts of Vision and Change for Biology Majors. *CBE—Life Sciences Education*, *13*(2), 200–211. https://doi.org/10.1187/cbe.13-12-0233

Campbell, C. E., & Nehm, R. H. (2013). A Critical Analysis of Assessment Quality in Genomics and Bioinformatics Education Research. *CBE—Life Sciences Education*, *12*(3), 530–541. https://doi.org/10.1187/cbe.12-06-0073

Cavallo, A. M. L., White, K. J., & McCall, D. (2011). The Mismatch Among Students' Views About Nature of Science, Acceptance of Evolution, and Evolutionary Science Understandings. *Science Education Review*, *10*(2), 6.

Cho, J., & Lee, E.-H. (2014). Reducing Confusion about Grounded Theory and Qualitative Content Analysis: Similarities and Differences. *The Qualitative Report*, *19*(32), 1–20. https://doi.org/10.46743/2160-3715/2014.1028

Cofré, H., Cuevas, E., & Becerra, B. (2017). The relationship between biology teachers' understanding of the nature of science and the understanding and acceptance of the theory of evolution. *International Journal of Science Education*, *39*(16), 2243–2260. https://doi.org/10.1080/09500693.2017.1373410

Coleman, J., Stears, M., & Dempster, E. (2015). Student teachers' understanding and acceptance of evolution and the nature of science. *South African Journal of Education*, *35*(2). Article 2. https://doi.org/10.4314/saje.v35i2

Deniz, H., Donnelly, L. A., & Yilmaz, I. (2008). Exploring the factors related to acceptance of evolutionary theory among Turkish preservice biology teachers: Toward a more informative conceptual ecology for biological evolution. *Journal of Research in Science Teaching*, *45*(4), 420–443. https://doi.org/10.1002/tea.20223

Dunk, R. D. P., Petto, A. J., Wiles, J. R., & Campbell, B. C. (2017). A multifactorial analysis of acceptance of evolution. *Evolution: Education and Outreach*, *10*(1), 4. https://doi.org/10.1186/s12052-017-0068-0

Dunk, R. D. P., & Wiles, J. R. (2018). Changes in Acceptance of Evolution and Associated Factors during a Year of Introductory Biology: The Shifting Impacts of Biology Knowledge, Politics, Religion, Demographics, and Understandings of the Nature of Science (p. 280479). *bioRxiv*, https://doi.org/10.1101/280479

Ferguson, D. G., & Jensen, J. L. (2021). Role models, compatibility, and knowledge lead to increased evolution acceptance. *Evolution: Education and Outreach*, *14*(1), 16. https://doi.org/10.1186/s12052-021-00155-x

Gallup Inc. (2019). *Evolution, Creationism, Intelligent Design*. *Gallup.Com*. Retrieved May 25, 2022, from https://news.gallup.com/poll/21814/Evolution-Creationism-Intelligent-Design.aspx

García, A. A. (2011). Cognitive Interviews to Test and Refine Questionnaires. *Public Health Nursing*, *28*(5), 444–450. https://doi.org/10.1111/j.1525-1446.2010.00938.x

Gefaell, J., Prieto, T., Abdelaziz, M., Álvarez, I., Antón, J., Arroyo, J., … & Rolán-Alvarez, E. (2020). Acceptance and knowledge of evolutionary theory among third-year university students in Spain. *PLOS ONE*, *15*(9), e0238345. https://doi.org/10.1371/journal.pone.0238345

Glaze, A. L., Snyder, S., Devereaux, R. A., & Smith, M. U. (2020). The Generalized Acceptance of EvolutioN Evaluation (GAENE) 3.0: Enhancement and Validation of a New Measure of Acceptance. *Journal of Higher Education Theory and Practice*, *20*(7). 128–148. https://doi.org/10.33423/jhetpv20i7.3157

Gregory, T. R. (2008). Evolution as Fact, Theory, and Path. *Evolution: Education and Outreach*, *1*(1). 46–52. https://doi.org/10.1007/s12052-007-0001-z

Grunspan, D. Z., Nesse, R. M., Barnes, M. E., & Brownell, S. E. (2018). Core principles of evolutionary medicine: A Delphi study. *Evolution, Medicine, and Public Health*, *2018*(1), 13–23. https://doi.org/10.1093/emph/eox025

Ingram, E. L., & Nelson, C. E. (2006). Relationship between achievement and students' acceptance of evolution or creation in an upper-level evolution course. *Journal of Research in Science Teaching*, *43*(1), 7–24. https://doi.org/10.1002/tea.20093

Krippendorff, K. (2018). *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: SAGE Publications.

Kuschmierz, P., Beniermann, A., Bergmann, A., Pinxten, R., Aivelo, T., Berniak-Woźny, J., … & Graf, D. (2021). European first-year university students accept evolution but lack substantial knowledge about it: A standardized European cross-country assessment. *Evolution: Education and Outreach*, *14*(1), 17. https://doi.org/10.1186/s12052-021-00158-8

Kuschmierz, P., Beniermann, A., & Graf, D. (2020a). Development and evaluation of the knowledge about evolution 2.0 instrument (KAEVO 2.0). *International Journal of Science Education*, *42*(15), 2601–2629. https://doi.org/10.1080/09500693.2020.1822561

Kuschmierz, P., Meneganzin, A., Pinxten, R., Pievani, T., Cvetković, D., Mavrikaki, E., … & Beniermann, A. (2020b). Towards common ground in measuring acceptance of evolution and knowledge about evolution across Europe: A systematic review of the state of research. *Evolution: Education and Outreach*, *13*(1), 18. https://doi.org/10.1186/s12052-020-00132-w

Laidlaw, C. T., Bybee, S. M., Shumway, S., Ogden, T. H., Peck, S., & Jensen, J. L. (2022). An Examination of Constructivism, Active Learning, and Reflexive Journaling and Their Independent and Combined Effects on Student Acceptance of Biological Evolution. *Journal of College Science Teaching*, *51*(3), 88–98.

Lawson, A. E. (1983). Predicting science achievement: The role of developmental level, disembedding ability, mental capacity, prior knowledge, and beliefs. *Journal of Research in Science Teaching*, *20*(2), 117–129. https://doi.org/10.1002/tea.3660200204

Lenski, R. E. (2020). *Evolution: Fact and Theory*. *National Center for Science Education*. https://ncse.ngo/evolution-fact-and-theory

Mead, L. S., Kohn, C., Warwick, A., & Schwartz, K. (2019). Applying measurement standards to evolution education assessment instruments. *Evolution: Education and Outreach*, *12*(1), 5. https://doi.org/10.1186/s12052-019-0097-y

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Miller, J. D., Scott, E. C., Ackerman, M. S., Laspra, B., Branch, G., Polino, C., & Huffaker, J. S. (2022). Public acceptance of evolution in the United States, 1985–2020. *Public Understanding of Science*, *31*(2), 223–238. https://doi.org/10.1177/09636625211035919

Nadelson, L. S., & Hardy, K. K. (2015). Trust in science and scientists and the acceptance of evolution. *Evolution: Education and Outreach*, *8*(1), 9. https://doi.org/10.1186/s12052-015-0037-4

Nadelson, L. S., & Southerland, S. (2012). A More Fine-Grained Measure of Students' Acceptance of Evolution: Development of the Inventory of Student Evolution Acceptance—I-SEA. *International Journal of Science Education*, *34*(11), 1637–1666. https://doi.org/10.1080/09500693.2012.702235

Nápoles-Springer, A. M., Santoyo-Olsson, J., O'Brien, H., & Stewart, A. L. (2006). Using cognitive interviews to develop surveys in diverse populations. *Medical Care*, *44*(11 Suppl 3), S21–30. https://doi.org/10.1097/01.mlr.0000245425.65905.1d

National Academy of Sciences. (2008). *Science, Evolution, and Creationism*. Washington, DC: National Academies Press. https://doi.org/10.17226/11876

Nehm, R. H., & Ha, M. (2011). Item feature effects in evolution assessment. *Journal of Research in Science Teaching*, *48*(3), 237–256. https://doi.org/10.1002/tea.20400

Pew Research Center. (2019,. February 6). The Evolution of Pew Research Center's Survey Questions About the Origins and Development of Life on Earth. *Pew Research Center's Religion & Public Life Project*. Retrieved May 25, 2022, from https://www.pewresearch.org/religion/2019/02/06/the-evolution-of-pew-research-centers-survey-questions-about-the-origins-and-development-of-life-on-earth/

Preuss, T. M., & Robert, J. S. (2014). Animal Models of the Human Brain: Repairing the Paradigm. In Gazzaniga, M. S., & Mangun, G. R. (Eds.), *The Cognitive Neurosciences*, 5th ed. (pp. 59–66). Cambridge, MA: MIT Press.

Reznick, D. N., & Ricklefs, R. E. (2009). Darwin's bridge between microevolution and macroevolution. *Nature*, *457*(7231). 837–842. https://doi.org/10.1038/nature07894

Rissler, L. J., Duncan, S. I., & Caruso, N. M. (2014). The relative importance of religion and education on university students' views of evolution in the Deep South and state science standards across the United States. *Evolution: Education and Outreach*, *7*(1), 24. https://doi.org/10.1186/s12052-014-0024-1

Romine, W. L., Todd, A. N., & Walter, E. M. (2018). A closer look at the items within three measures of evolution acceptance: Analysis of the MATE,

I-SEA, and GAENE as a single corpus of items. *Evolution: Education and Outreach*, *11*(1), 17. https://doi.org/10.1186/s12052-018-0093-7

Rühli, F. J., & Henneberg, M. (2013). New perspectives on evolutionary medicine: The relevance of microevolution for human health and disease. *BMC Medicine*, *11*(1), 115. https://doi.org/10.1186/1741-7015-11-115

Rutledge, M. L., & Warden, M. A. (1999). The Development and Validation of the Measure of Acceptance of the Theory of Evolution Instrument. *School Science and Mathematics*, *99*(1), 13–18. https://doi.org/10.1111/j.1949-8594.1999.tb17441.x

Rutledge, M. L., & Warden, M. A. (2000). Evolutionary Theory, the Nature of Science & High School Biology Teachers: Critical Relationships. *The American Biology Teacher*, *62*(1), 23–31. https://doi.org/10.2307/4450822

Sbeglia, G. C., & Nehm, R. H. (2018). Measuring evolution acceptance using the GAENE: Influences of gender, race, degree-plan, and instruction. *Evolution: Education and Outreach*, *11*(1), 18. https://doi.org/10.1186/s12052-018-0091-9

Sbeglia, G. C., & Nehm, R. H. (2019). Do you see what I-SEA? A Rasch analysis of the psychometric properties of the Inventory of Student Evolution Acceptance. *Science Education*, *103*(2), 287–316. https://doi.org/10.1002/sce.21494

Siciliano-Martina, L., & Martina, J. P. (2020). Shifting barriers to the acceptance of evolution in an underrepresented student group. *International Journal of Science Education*, *42*(13), 2205–2223. https://doi.org/10.1080/09500693.2020.1815247

Sickel, A. J., & Friedrichsen, P. (2013). Examining the evolution education literature with a focus on teachers: Major findings, goals for teacher preparation, and directions for future research. *Evolution: Education and Outreach*, *6*(1), 23. https://doi.org/10.1186/1936-6434-6-23

Sinatra, G. M., Southerland, S. A., McConaughy, F., & Demastes, J. W. (2003). Intentions and beliefs in students' understanding and acceptance of biological evolution. *Journal of Research in Science Teaching*, *40*(5), 510–528. https://doi.org/10.1002/tea.10087

Smith, M. U. (2010). Current Status of Research in Teaching and Learning Evolution: I. Philosophical/Epistemological Issues. *Science & Education*, *19*(6), 523–538. https://doi.org/10.1007/s11191-009-9215-5

Smith, M. U., & Scharmann, L. C. (1999). Defining versus describing the nature of science: A pragmatic analysis for classroom teachers and science educators. *Science Education*, *83*(4), 493–509. https://doi.org/10.1002/(SICI)1098-237X(199907)83:4<493::AID-SCE6>3.0.CO;2-U

Smith, M. U., & Siegel, H. (2016). On the Relationship Between Belief and Acceptance of Evolution as Goals of Evolution Education. *Science & Education*, *25*(5), 473–496. https://doi.org/10.1007/s11191-016-9836-4

Smith, M. U., Snyder, S. W., & Devereaux, R. S. (2016). The GAENE—Generalized Acceptance of EvolutioN Evaluation: Development of a new measure of evolution acceptance. *Journal of Research in Science Teaching*, *53*(9), 1289–1315. https://doi.org/10.1002/tea.21328

Southerland, S. A., Sinatra, G. M., & Matthews, M. R. (2001). Belief, Knowledge, and Science Education. *Educational Psychology Review*, *13*(4), 325–351. https://doi.org/10.1023/A:1011913813847

Stanisavljevic, J., Papadopoulou, P., & Djuric, D. (2013). Relationship between Acceptance and Understanding of the Evolution Theory by Various Groups of Teachers. *Croatian Journal of Education: Hrvatski Časopis Za Odgoj i Obrazovanje*, *15*(3), 693–714.

Trani, R. (2004). I Won't Teach Evolution; It's against My Religion. And Now for the Rest of the Story... *The American Biology Teacher*, *66*(6), 419–427. https://doi.org/10.2307/4451708

U.S. Bureau of Labor Statistics. (2019). *Field of degree: Biology: Occupational Outlook Handbook*. Washington, DC: Office of Occupational Statistics and Employment Projections. Retrieved August 29, 2022, from https://www.bls.gov/ooh/field-of-degree/biology/biology-field-of-degree.htm

Van Koevering, T. E., & Stiehl, R. (1989). Evolution, creation & wisconsin biology teachers. *The American Biology Teacher*, *51*(4), 200–202. https://doi.org/10.2307/4448903

Wiles, J. R. (2014). Gifted students' perceptions of their acceptance of evolution, changes in acceptance, and factors involved therein. *Evolution: Education and Outreach*, *7*(1), 4. https://doi.org/10.1186/s12052-014-0004-5

Willis, G. B. (2004). *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA: SAGE Publications, Inc. https://us.sagepub.com/en-us/nam/cognitive-interviewing/book225856

Zimmerman, M. (1987). The Evolution-Creation Controversy: Opinions of Ohio High School Biology Teachers. *The Ohio Journal of Science*, *87*(4), 115–125.