

REVIEW

Open Access



Evaluating the current state of evolution acceptance instruments: a research coordination network meeting report

M. Elizabeth Barnes^{1†}, Rahmi Q. Aini¹, James P. Collins², Ryan D. P. Dunk³, Emily A. Holt⁴, Jamie Jensen^{5†}, Joanna R. Klein⁶, Taya Misheva^{7†}, Louis S. Nadelson⁸, Michael J. Reiss⁹, William L. Romine¹⁰, Andrew Shtulman¹¹, Amanda L. Townley¹², Jason R. Wiles^{13*†}, Yi Zheng¹⁴ and Sara E. Brownell^{7†}

Abstract

Hundreds of studies have explored student evolution acceptance because evolution is a core concept of biology that many undergraduate biology students struggle to accept. However, this construct of “evolution acceptance” has been defined and measured in various ways, which has led to inconsistencies across studies and difficulties in comparing results from different studies. Many studies and essays have offered evaluations and perspectives of evolution acceptance instruments, but publications with a focus on consensus building across research teams is still needed. Further, little attention has been paid to how evolution acceptance instruments may be interpreted differently by students with varied religious backgrounds. Funded by a Research Coordination Network in Undergraduate Biology Education grant from the National Science Foundation, we gathered 16 experts from different disciplinary and religious backgrounds to review current evolution acceptance instruments and create a guide to the strengths and weaknesses of these instruments, including appropriate contexts for using these instruments and their potential weaknesses with different religious populations. Finally, in an attempt to move the field forward, we articulated a consensus definition of evolution acceptance that can be used to guide future instrument development.

[†]M. Elizabeth Barnes, Jamie Jensen, Taya Misheva, Jason R. Wiles and Sara E. Brownell contributed to this work equally.

*Correspondence:

Jason R. Wiles
jwiles01@syr.edu

¹ Social Perceptions of Science Lab, Biology Department, Middle Tennessee State University, Murfreesboro, TN 37132, USA

² School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA

³ Department of Biological Sciences, Auburn University, Auburn, AL 36849, USA

⁴ School of Biological Sciences, University of Northern Colorado, Greeley, CO 80639, USA

⁵ Department of Biology, Brigham Young University, Provo, UT, USA

⁶ Department of Biology and Biochemistry, University of Northwestern - St. Paul, St. Paul, MN 55113, USA

⁷ Center for Research in Inclusive Science Education, School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA

⁸ Department of Leadership Studies, University of Central Arkansas, Conway, AR 72035, USA

⁹ Faculty of Education and Society, University College London, London WC1H 0AL, UK

¹⁰ Department of Biological Sciences, Wright State University, 3640 Colonel Glenn Hwy, Dayton, OH 45435, USA

¹¹ Department of Psychology, Occidental College, 1600 Campus Road, Los Angeles, CA 90041, USA

¹² Department of Middle Grades and Secondary Education, Georgia Southern University, 3200 Southern Drive, Statesboro, GA 30458, USA

¹³ Department of Biology, Syracuse University, 107 College Place, Syracuse, NY 13244, USA

¹⁴ Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ 85287, USA



Introduction

To understand how to improve student evolution acceptance, evolution education researchers have extensively explored the extent to which students accept evolution (Barnes et al. 2022b; Dunk et al. 2019; Rice et al. 2011) as well as the variables and interventions associated with higher levels of evolution acceptance (Barnes et al. 2020; Bowen et al. 2022; Dunk et al. 2017; Fiedler et al. 2019; Glaze et al. 2014; Green and Delgado 2021; Lindsay et al. 2019; Rutledge and Sadler 2011; Wiles 2014; Wiles and Alters 2011). However, even though evolution acceptance is a key construct in evolution education research, methods and recommendations for measuring evolution acceptance vary widely (Mead et al. 2019; Nadelson and Southerland 2012; Sbeglia and Nehm 2018, 2019; Smith 2010; Smith et al. 2016), which can contribute to inconsistent results and conclusions from studies aimed at improving evolution acceptance (Barnes et al. 2019). Although several publications have evaluated and/or reviewed certain evolution acceptance instruments (Kuschmierz et al. 2020; Lloyd-Strovas and Bernal 2012; Mead et al. 2019; Metzger et al. 2018; Romine et al. 2018; Sbeglia and Nehm 2018, 2019; Smith 2010; Smith et al. 2016), a guide of the strengths and weaknesses of each instrument and the appropriate populations for their use does not yet exist, which limits the utility of these recommendations because it requires one to engage with multiple publications to contrast different evolution acceptance instruments. Further, current studies and recommendations for measurement in evolution acceptance research are written by individuals (Smith 2009), single research teams (Sbeglia and Nehm 2018, 2019), or smaller collaborative groups (Barnes et al. 2019; Beniermann et al. 2023; Glaze et al. 2020; Smith et al. 2016). So, these recommendations may not be representative of the evolution education community. Thus, it is an important time to establish current consensus recommendations from a wide range of experts and research teams about evolution acceptance measurement to guide the field forward.

It is also important to begin establishing a potential consensus about measurement of evolution acceptance across research fields. There are distinct lines of peer reviewed evolution education research in cognitive and developmental sciences (Evans 2001; Evans and Lane 2011; Gervais 2015; Shtulman 2006; Shtulman and Calabi 2012; Shtulman and Schulz 2008; Shtulman and Valcarcel 2012), sociology and social psychology (Baker 2013; Baker et al. 2018; Elsdon-Baker 2015; Hill 2014; Leicht et al. 2022; McPhetres et al. 2021; McPhetres and Zuckerman 2018; Unsworth and Voas 2018, 2021), educational psychology (Dole and Sinatra 1998; Sinatra et al. 2014; Southerland et al. 2001), theology (Austriaco 2019;

Loke 2016; McGrath 2021), and discipline based education research (Asghar 2013; Barnes et al. 2020; Eddy et al. 2013; Glaze et al. 2014; Graves 2019; Holt et al. 2018; Jensen et al. 2019; Nehm and Schonfeld 2007; Sbeglia and Nehm 2020; Verhey 2005; Wiles 2014) that seek to understand low acceptance of evolution. Unfortunately, researchers in these fields rarely speak across their disciplinary boundaries about how to measure evolution acceptance. This siloed approach to understanding evolution acceptance means that studies may only consider the disciplinary knowledge of those who are conducting the studies and the literature base surrounding evolution acceptance from that discipline. Even within disciplines, researchers define and measure evolution acceptance differently, making studies difficult to compare (Lloyd-Strovas and Bernal 2012; Smith 2010). To mitigate this issue, we gathered experts from relevant disciplines to meet and contribute to an expert review of instruments being used in evolution acceptance research.

An additional focus of the meeting was to evaluate instruments based on their suitability for religious populations. While many factors have been shown to predict students' evolution acceptance [for example, their understanding of evolution (Barnes et al. 2019; Carter and Wiles 2014; Glaze et al. 2014; Nadelson and Southerland 2010; Rutledge and Warden 2000), understanding of the nature of science (Dunk et al. 2017; Dunk and Wiles 2018), religiosity (Dunk and Wiles 2018; Gutowski et al. 2023; Heddy and Nadelson 2012; Jensen et al. 2019), religious affiliation, and demographic factors such as education level (Heddy and Nadelson 2013), race/ethnicity (Sbeglia and Nehm 2018), and gender (Sbeglia and Nehm 2018)], one's perceived conflict between their religion and evolution has been shown to be the most predictive factor for religious students' acceptance of evolution in the United States (Barnes et al. 2021b). Despite this importance of perceived conflict with religion, little attention has been paid to validity concerns of evolution acceptance instruments for individuals with different religious identities (Barnes et al. 2022a; Beniermann et al. 2023; Misheva et al. 2023).

Specifically, few prior studies have examined content validity evidence based on religious identity (Beniermann et al. 2023). Content validity refers to the extent that the items of the instrument align with the construct being measured (see, AERA et al. 2014). While many developed instruments included expert review in their initial creation, it is not clear that the instruments were reviewed to consider different religious groups (Glaze et al. 2020; Nadelson and Southerland 2012; Rutledge and Warden 1999) except in the case of one single instrument recently published (Beniermann et al. 2021a). Thus, items may have construct irrelevant aspects that experts with

diverse experience with religious populations may be able to identify. For evolution acceptance, this means evaluating the extent to which students are answering survey questions based on their evolution acceptance and not a separate construct such as their religious identity or their understanding of evolution (i.e., their ability to answer questions on an evolution test). Is there sufficient content validity evidence for atheists compared to Muslim and Christian students? For example, several existing instruments ask the extent to which a student agrees that religious texts like the Bible conflict with evolution (Beniermann et al. 2023; Rutledge and Warden 1999; Smith 2010; Smith et al. 2016) but many students may subscribe to a religious text that is not the Bible. A Christian student would likely answer this based on their own evolution acceptance and an “agree” would indicate rejection of evolution but this same response from an atheist student would likely have little to do with their evolution acceptance (Barnes et al. 2022a). Other items within instruments refer to God as a creator, making them difficult to navigate for respondents who are non-religious or those who belong to polytheistic religions. Content validity based on religious identification has been evaluated recently using statistical methods (Beniermann et al. 2023) and the authors identified several items on evolution acceptance instruments that differentially functioned for highly religious populations in Germany. Thus, we posit that we need to further evaluate content validity of instrumentation with expert reviewers and through a variety of religious lenses if we want to improve evolution acceptance measurement.

Why is evaluating evolution acceptance instruments important?

There is a growing concern that “evidence” used to make evidence-based decisions in education is inaccurate because of a lack of rigor in the development of the instruments (Dunk et al. 2019; Lloyd-Strovas and Bernal 2012; Mead et al. 2019; Nehm and Mead 2019). While we acknowledge that a tremendous amount of work and, in many cases, rigorous efforts have gone into developing evolution acceptance instruments, we argue that these instruments can still be improved, and this would better establish trust in the field of evolution education. The concerns about current evolution acceptance instruments have been so extensive that *Evolution: Education and Outreach* dedicated an entire recent issue to measurement in evolution education studies (Nehm and Mead 2019). Further, a recent perspective piece coauthored by evolution education researchers, including members of the network, in *Nature: Ecology and Evolution*, highlighted the need for more consistent evolution acceptance measurement (Dunk et al. 2019). These recent

publications indicate that it is an important time for evolution education researchers to come together to address the problem of evolution acceptance measurement.

Although many instruments to measure evolution acceptance have been developed and published, the range of existing instruments means that evolution education researchers must decide the best way to measure evolution acceptance and that researchers are often trying to compare findings from studies that use different instruments. Recent studies have addressed this by administering multiple instruments together to compare them (Barnes et al. 2019; Metzger et al. 2018; Romine et al. 2018) and showed many similarities between instruments. However, when researchers administered multiple evolution acceptance surveys to undergraduate biology students they found that different instruments used to measure evolution acceptance given to the same students led to different research results and conclusions (Barnes et al. 2019). This indicated that inconsistencies of results across studies that have been noted by researchers (Beniermann et al. 2023; Lloyd-Strovas and Bernal 2012; Smith 2010) could have been due to differences in the instrument used to measure evolution acceptance.

Convening a group to provide expert evaluation of current evolution acceptance instruments

We convened a group of 16 experts in evolution education research to address these issues through a National Science Foundation Research Coordination in Undergraduate Biology Education (RCN UBE) incubator grant (all of whom are authors on this report). The idea originated as a group discussion with PI team S.E.B., M.E.B., J.R.W., and J.J. and one network member (R.D.P.D.) at the 2019 summer meeting for the Society for the Advancement in Biology Education Research (SABER) about how best to measure evolution acceptance to allow comparisons between studies. To move forward and make meaningful strides in improving evolution acceptance, we posited that evolution education researchers across disciplines and representing different religious beliefs need to come together to identify areas of consensus.

With any invited meeting with limited spots, it is impossible to include all individuals who have contributed to the evolution acceptance literature. We were able to recruit individuals from the following disciplinary perspectives: discipline-based education research, science education, cognitive psychology, theology, psychometrics, and evolutionary biology, most of whom have published at least one novel measure of evolution acceptance.

We also gathered experts with experience measuring evolution acceptance among populations of varying religiosity and religious affiliations. We had participants who have taught evolution at a religious institution (J.R.K.,

J.J.), have taught evolution at a secular institution with a significant religious population (E.A.H., A.L.T., M.E.B.), and who have taught evolution at a secular institution with an average religious population (J.P.C., J.R.W., M.E.B.) Our experts themselves came from a diversity of religious and non-religious backgrounds (Table 1). Notably, over half of our participants have been religious at one time and a third of the participants are currently religious. Our shared positionality that guided our work as a network is that (1) Evolution acceptance is an important construct that can be measured, (2) We do not advocate the teaching of alternatives in opposition to evolution (e.g., special creationism, intelligent design), and (3) One can be religious and still accept the fundamental tenets of evolution. Our network came together with these as shared core values.

To balance the current disparate reviews and recommendations that exist in the evolution acceptance literature, we focused on areas of consensus reached by the group to clarify areas of agreement among a diversity of experts. This meeting report includes experts' consensus views on strengths, weaknesses, and uses of current evolution acceptance instruments, including potential limitations of each instrument for religious students. Since

varying definitions of evolution acceptance likely underlie inconsistencies between current instruments, we also developed a potential consensus definition of evolution acceptance to guide future instrument development.

Method for determining consensus recommendations

The network spent significant time evaluating nine total evolution acceptance instruments: six multi-item instruments that have multiple sources of peer-reviewed validation evidence and three single-item instruments that have been widely used in the evolution education literature. These included the Measure of Acceptance of the Theory of Evolution (MATE) (Rutledge and Warden 1999) and its revised version the MATE 2.0 (Barnes et al. 2022a), the Generalized Acceptance of Evolution Evaluation (GAENE) (Smith et al. 2016) and its revised version the GAENE 3.0 (Glaze et al. 2020), the Inventory of Student Evolution Acceptance (I-SEA) (Nadelson and Southerland 2012), and the Evolution Education Questionnaire, Attitudes Towards Evolution Subscale (ATEVO-EEQ) (Beniermann et al. 2021a). All instruments in a ready to use format can be found in the Additional file 1. Prior to

Table 1 Description of network members' childhood religious background, experience with conflict between religion and evolution, experience rejecting evolution, current religious identity, and self-reported experiences/expertise with specific religious populations

Network member	Childhood religious background	Past conflict with religion and evolution?	Rejected evolution in the past?	Current religious or philosophical identification:	Self-reported prior experience/expertise with religious populations:
1	Christian (Methodist)	Yes	Yes	Atheist	Christian
2	None	No	No	Agnostic	Judeo-Christian
3	Christian (CJCLDS ¹)	No	No	Christian	Judeo-Christian (CJCLDS)
4	Christian (church of Christ)	Yes	Yes	Humanist ²	Evangelical Christian, Muslim, Humanist
5	Christian (evangelical), secular	Yes	Yes	Atheist	Judeo-Christian
6	Agnostic	No	No	Buddhist	Christian, Muslim, Hindu, Buddhist
7	Atheist	No	No	Agnostic	None
8	Christian (Lutheran)	No	No	Christian	Christian (mainline and evangelical)
9	Atheist	No	No	Christian	Christian, Jewish, Muslim
10	Agnostic	No	No	Atheist	Judeo-Christian
11	Jewish	No	No	Atheist	Judeo-Christian
12	Christian (evangelical Southern Baptist)	Yes	Yes	Unitarian Christian	Christian (mainline and evangelical), Unitarian
13	Methodist, Unitarian Universalist, Secular	No	Unsure	None	Judeo-Christian
14	Christian (Roman Catholic)	Yes	No	Agnostic	Christian (Roman Catholic)
15	Muslim (Sunni Islam)	Unsure	Yes	Muslim (Sunni Islam)	Muslim, Atheist, Catholic, Christian
16	Agnostic	No	No	Agnostic	Judeo-Christian

¹ The official name of Latter-day Saints is the Church of Jesus Christ of Latter-day Saints (CJCLDS), formerly known as LDS or Mormon, but both terms are no longer recommended

² Humanism is a progressive philosophy of life that affirms our ability and responsibility to lead ethical lives of personal fulfillment that aspire to the greater good

the meeting each network member was sent a bibliography of the published papers of the instruments as well as the most up to date research at that time on evolution acceptance measurement (the bibliography included: Barnes et al. 2019; Barnes et al. 2022a; Beniermann et al. 2021a; Glaze et al. 2020; Mead et al. 2019; Nadelson and Southerland 2012; Romine et al. 2018; Rutledge and Sadler 2007; Rutledge and Warden 1999; Sbeglia and Nehm 2018, 2019; Smith 2009; Smith et al. 2016). So, if some network members were not familiar with some of these papers, they were able to familiarize themselves with them before coming to the meeting.

To determine consensus opinions among experts on evolution acceptance measurement, we evaluated the extent of content validity for each instrument based on expert review (see American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014, pg. 13–15 for a description of content validity). To evaluate validity evidence based on expert review for each instrument, the Co-PI team asked the network members to individually review each of these instruments based on their strengths and weaknesses before they came to the meeting, including item-level feedback, and to specifically consider how the instruments would perform with different student populations, including different religious populations. All responses were collected digitally through Qualtrics®. The Co-PI team summarized member responses and presented them to the group, and members were able to go through each summary and either agree or disagree, as well as add additional comments. Subsequently at the meeting, network members were assigned to one of four groups to discuss a group of instruments and report back to the larger network the strengths and weaknesses of each instrument. Then, each group presented their evaluation and discussed each instrument in detail with the larger group. Participants were given anonymous surveys at various points in our discussions to pinpoint any disagreements that participants did not want to report in front of the group. The Co-PI team subsequently identified all points of consensus among the group and co-wrote the meeting report, which was then reviewed by all members of the network for approval. Of note, several of our network members have authored one of these instruments (MATE 2.0: M.E.B., T.M., S.E.B.; I-SEA: L.S.N.; GAENE 3.0: A.L.T.), which represents both a source of information and a conflict of interest. To try to minimize this concern, no author of an instrument was part of the group to review that instrument, although everyone participated in individual feedback and the group discussion.

Below, we provide a brief background and description of each instrument, previous critiques from the published literature about that instrument, and what network

members identified as strengths and weaknesses of that instrument.

What are the strengths and weaknesses of current instruments that exist to measure evolution acceptance?

Multi-item instruments

Measure of Acceptance of the Theory of Evolution (MATE) (Rutledge and Warden 1999)

History, description, and prior critiques The original MATE (Rutledge and Warden 1999) has been by far the most used instrument in evolution education research to measure evolution acceptance (Barnes et al. 2022a; Mead et al. 2019), despite persistent criticisms of this measure (Beniermann et al. 2023; Lloyd-Strovas and Bernal 2012; Smith 2009). The MATE was originally created to measure evolution acceptance among high school biology teachers in Indiana, but has since been used to measure evolution acceptance of a wide variety of populations both within and outside the United States (Athanasidou and Papadopoulou 2012; Deniz et al. 2008; Moore and Cotner 2009; Trani 2004). The MATE consists of 20 items rated on a 5-point Likert scale from strongly disagree to strongly agree, with “undecided” as the midpoint. Content validity was established via expert review from evolutionary biologists, science educators, and a philosopher of science. While the authors reported that a factor analysis of the MATE revealed only one factor (Rutledge and Warden 1999), subsequent work has shown that the instrument may be multidimensional, with negatively worded items and positively worded items falling on separate dimensions (Metzger et al. 2018; Romine et al. 2018; Sbeglia and Nehm 2019).

The authors of the MATE did not provide an explicit definition of evolution acceptance, but they did describe what aspects of evolution they sought to address: (1) evolutionary processes, (2) available evidence of evolutionary change, (3) evolutionary theory’s ability to explain phenomena, (4) human evolution, (5) the age of the earth, (6) validity of science as a way of knowing, and (7) the current status of evolutionary theory within the scientific community. Subsequent authors have provided a post-hoc definition of evolution acceptance based on the survey items (Romine et al. 2017). The MATE has been criticized for including the constructs of understanding of evolution, understanding of the nature of science, and scientists’ views of evolution, which is likely in part because of its ambiguous definition of evolution acceptance (Smith 2010). Importantly, the MATE could lead to inflated correlations with these factors because researchers often seek to determine the relationship between constructs such as acceptance and understanding (Barnes et al. 2019, 2022a; Smith 2010). This instrument also

repeatedly uses the word “evolution” without specifying the scale (microevolution or macroevolution) or context (humans or non-humans). Researchers often justify using the MATE because of its prior prevalent use (Metzger et al. 2018; Nadelson and Southerland 2012; Romine et al. 2017; Sbeglia and Nehm 2019; Smith et al. 2016).

Network meeting identified strengths and weaknesses After careful evaluation and discussion, network members identified both strengths and weaknesses of the MATE and offered recommendations for its use going forward. Generally, network members agreed that a key feature of the MATE is that it has been highly used, so there is a broad literature base of it being used with high school students, high school teachers, K-12 teachers in training, undergraduates (majors and non-majors), in several international settings and in different religious populations. The questions on the MATE are relatively easy to understand. Additional strengths of the instrument, from a researcher’s perspective, is that it is relatively short, and therefore can be completed by respondents rather quickly.

However, several serious weaknesses were identified. Though there was overall agreement that the language was appropriate for undergraduate biology students, some network members noted some confusing wording in certain items, such as the word “meaning” in “The theory of evolution brings meaning to the diverse characteristics and behaviors observed in living forms.” Another weakness is that some items assess an understanding of evolution potentially giving students the impression that the instrument is looking for the “right” answer instead of eliciting students’ personal opinions or acceptance of the information. Additionally, due to the conflation of understanding of evolution and the nature of science with acceptance, students who score at the same level of overall acceptance (i.e., the composite score) may vary widely in their acceptance, rejection, or knowledge of different aspects of evolution or the conflated constructs. The network members noted the several published critiques of this instrument for these reasons and others, including that the MATE operates multi-dimensionally (measures multiple constructs). In addition, due to the presence of negatively and positively coded items that are similar (e.g., “The age of the earth is less than 20,000 years” and “The age of the earth is at least 4 billion years”), the MATE can seem redundant; while psychometrically desirable, this redundancy may make respondents feel as though they are answering the same question repeatedly, which can contribute to survey fatigue.

Considering the relationship between evolution and religion, the network members’ assessment was that a significant weakness of the MATE is that it is Christian-centric (or at least Judeo-Christian-centric) since it

directly references the Bible. Given this, there are validity concerns with the MATE for Muslim students and for agnostic or atheist students who do not consider biblical texts as valid or who may not even be familiar with the Bible. There is additional risk, given its direct reference to religion, that the MATE may be viewed as presenting conflict between science/evolution and religion. For example, the item “The theory of evolution cannot be correct since it disagrees with the Biblical account of creation” might be read as meaning that these ideas must be mutually exclusive.

Recommendation for the MATE: Because of the listed weaknesses, we as a network recommended that the MATE is valuable only for use in comparison studies with prior research in which it has been employed, given that it has been used for decades and has been used in many contexts. The content validity issues of the MATE (Smith et al. 2016), and subsequent response process validity issues that have been uncovered with the MATE (Barnes et al. 2022a) make it not suitable for most evolution education researcher interests.

MATE 2.0 (Barnes et al. 2022a)

History, description, and prior critiques The MATE 2.0 (Barnes et al. 2022a) is a revised version of the MATE meant to address prior criticisms. The MATE 2.0 consists of 9 items rated on a 5-point Likert scale from strongly disagree to strongly agree, with “neutral” as the midpoint. To revise the original MATE, the authors of the MATE 2.0 first evaluated the response process validity of the original MATE by conducting cognitive interviews with 62 undergraduate students. The results empirically documented what researchers had previously voiced as potential limitations of the original MATE. The authors also newly found that students were unsure of what was meant by “evolution” in many items; these students answered based on an interpretation of “evolution” that includes concepts they agree with (e.g., microevolution) and excludes concepts they disagree with (e.g. human macroevolution). Based on these findings and criticisms from the prior literature, the authors revised some items, deleted items that could not be revised, and then conducted 29 additional cognitive interviews on the revised instrument, documenting a reduction in the response process errors from the previous version. The authors also provided a definition of evolution acceptance for their new instrument: “The agreement that it is scientifically valid that all species have evolved from prior species” and added a prompt that instructs the survey taker to answer based on their own opinion. The authors administered the MATE 2.0 to students in 22 undergraduate biology classes across the United States and gathered structural validity evidence through a Rasch dimensionality analysis, reliabil-

ity evidence by calculating a Cronbach's alpha coefficient, and concurrent validity evidence through correlations with other measures of evolution acceptance. An important change from the MATE to the MATE 2.0 is that the focus of each item on the MATE 2.0 is on macroevolution and species change, whereas some items on the original MATE could have been interpreted as asking about microevolution alone.

Network meeting identified strengths and weaknesses After careful evaluation and discussion of the revised MATE 2.0, network members identified both strengths and weaknesses and offered a recommendation for its use going forward. This revised instrument is shorter and simpler than the original MATE. In response to criticisms of the original MATE, the MATE 2.0 includes a user prompt explicitly explaining that student responses should reflect their own opinions. Thus, this revised version minimizes conflation of understanding and acceptance of evolution.

Although the MATE 2.0 addresses many of the notable problems with the original instrument, network members still identified some weaknesses. First, it is relatively new and consequently has not yet been tested in as many contexts as the original instrument. Similarly, because of the significant differences between it and the original MATE, it cannot be used for comparison with prior research that used the original MATE. Network members also pointed out the potential for students to have different conceptions of what instrument items mean when asking about “previous” or “earlier” “species,” such as in “Current scientific evidence suggests that new species can evolve from earlier species.”

Considering its functioning with religiously diverse participants, the MATE 2.0 has the advantage of not referencing religion and thus not being specific to Christianity. Thus, due to its more religiously inclusive nature, it is likely to be more applicable across different settings and within more religiously diverse populations, including Muslim students as well as agnostic and atheist individuals. However, because the MATE 2.0 only assesses macroevolution acceptance, it may not accurately capture the nuance of individuals who accept some evolutionary principles but do not accept an old Earth or, alternatively, it may suffer from a floor effect in highly resistant populations who only accept microevolution. Interestingly, it refers heavily to speciation or change specifically at the species level to assess macroevolution, which some creationists—who reject large-scale evolutionary change—may construe as acceptable “microevolution” or small-scale change within “created kinds” (generally believed to be at the taxonomic level of “family” or

“order”), causing an overestimation of acceptance among those populations.

Recommendation for the MATE 2.0: It was the conclusion of the network that the MATE 2.0 is preferred in almost all situations to the original MATE with the exception of comparison to previous studies employing the original instrument. In addition, the MATE 2.0 is recommended for populations with heterogeneous worldviews where a more Christian-centric instrument would be less appropriate.

Inventory of Student Evolution Acceptance (I-SEA) (Nadelson and Southerland 2012)

History, description, and prior critiques The I-SEA was created to improve upon the original MATE by subdividing the measurement of evolution acceptance into three distinct subscales for microevolution, macroevolution, and human evolution (Nadelson and Southerland 2012). The authors made these distinctions based on prior research showing that students commonly perceive differences between microevolution and macroevolution, and between human and non-human evolution, even though these concepts are not biologically different (Nehm and Ha 2011; Reznick and Ricklefs 2009).

The I-SEA consists of 24 items with 8 items in each subscale; it is scored on a 5-point Likert scale ranging from strongly agree to strongly disagree, with “undecided” as the midpoint. During instrument development, an initial set of items was drafted based on student interviews, and think-aloud interviews were performed with four high school and college students to assess item clarity (Nadelson and Southerland 2012). Content validity was established using expert reviews by nine college biology faculty. The authors performed an exploratory factor analysis to reduce the initial item set to the 24 items present in the final draft, and then re-tested this final draft using a confirmatory factor analysis. Results of the confirmatory factor analysis showed that grouping the items into the three subscales fit the data well. Validation was conducted using high school and college student populations (Nadelson and Southerland 2012). However, a later study using Rasch analysis based on data from 2130 undergraduates found that items in the human evolution subscale may form two clusters, which approximately correspond to human microevolution and human macroevolution, but these findings were preliminary (Sbeglia and Nehm 2019). Furthermore, additional cognitive interviews with 22 college students revealed that students with little to no prior exposure to college-level biology struggled to interpret many of the items because they lacked relevant scientific knowledge (Misheva et al. 2023).

Network meeting identified strengths and weaknesses After careful evaluation and discussion, network members identified both strengths and weaknesses of the I-SEA and offered a recommendation for its use going forward. One of its main advantages is that it divides items into three distinct subscales: microevolution, macroevolution, and human evolution, which can be administered as one large survey or individual subscales. This is useful because prior research suggests that students can have very different levels of acceptance of evolution depending on whether they are thinking about small changes in populations over time (microevolution), large changes like speciation (macroevolution), or human evolution (Barnes et al. 2019, 2021a, 2022b; Evans 2008; Pew 2013; Rachmatullah et al. 2018). So, this instrument allows researchers to disaggregate evolution acceptance by these constructs and use these results to assess the efficacy of interventions that target specific forms of evolution rejection (e.g., rejection of human evolution). Additionally, these three subscales clearly align with common instructional goals and allow researchers (and instructors) to obtain a more detailed profile of a student's views on evolution than is possible with a single-construct measure. It also contains negatively coded items as attention checks, which some could consider more psychometrically sound. However, Romine et al. (2018) shows that these negatively coded items on the I-SEA show multi-dimensional behavior and Sbeglia and Nehm (2019) recommend reversing these negatively worded items. At the meeting, we discussed the idea that the negatively coded items entail a greater amount of cognitive load, which may be responsible for the multi-dimensional behavior. This is supported by Romine et al. (2017), who found that the negatively worded dimension has two times more variance shared with knowledge of macroevolution than the positively worded dimension.

A main weakness identified by network members was that the I-SEA contains a number of items that use scientific jargon and refer to biological concepts that may be unfamiliar to students with less biology background (e.g., "Humans were derived from ancestral primates"), which can lead to conflation between knowledge about and acceptance of evolution. As such, the I-SEA may be more appropriate for student populations that have had some prior exposure to college-level biology, such as students in an upper-level biology course. The phrasing of some items may also be grammatically challenging and may pose difficulties for students who speak English as a second language (e.g., "I think there is an abundance of observable evidence to support the theory describing how variations within a species can happen"). Some items seem to measure two statements leading to potential complications for respondents who agree with one part but not another (e.g., "Species were created to

be perfectly suited to their environment, so they do not change"). An additional weakness is that the I-SEA does not explicitly define evolution acceptance leaving some items open for interpretation by the survey taker as to whether they are referring to microevolution or macroevolution (e.g., "I think that humans evolve"). Furthermore, the I-SEA is a relatively long instrument (24 items), which can lead to survey fatigue if paired with other measures and may present logistical constraints for administration (e.g., may be too time-consuming to administer during class). While the subscales can be hailed as a strength, they also present a weakness in that the I-SEA cannot be used as a unidimensional measure (Romine et al. 2018; Sbeglia and Nehm 2019). Further, several network members expressed the opinion in line with Sbeglia and Nehm's (2019) preliminary findings that suggest the human evolution scale may need to be further parsed into human micro- and human macroevolution.

Considering its functioning with religiously diverse populations, items within the I-SEA do not specify any one religion's account of creation or reference a God or higher power, making the instrument suitable for deployment among religiously diverse populations or those without religious belief.

Recommendation for I-SEA: The network members agreed that the I-SEA is an appropriate measure for heterogeneous religious and non-religious populations, but may be somewhat challenging for respondents with little or no prior biology exposure. It is a particularly desirable scale for researchers interested in delineating microevolution, macroevolution, and human evolution constructs.

Generalized Acceptance of Evolution Evaluation (GAENE) (Smith et al. 2016)

History, description, and prior critiques The original GAENE was published in 2016 to address prior critiques of other evolution acceptance instruments (Smith et al. 2016). It is intended for high school and college students. The original GAENE consists of 13 items rated on a 5-point Likert scale, from strongly disagree to strongly agree, with "no opinion" as the midpoint. Item development was guided by an explicit definition of acceptance that was meant to disentangle acceptance from understanding and belief: "Evolution acceptance is 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" (Smith et al. 2016, pg. 8).

Items on the GAENE were iteratively developed, with two rounds of pilot testing that included a focus group with five high school students, written feedback from 26 high school students, and interviews with seven university students. Further validation evidence was

gathered by expert review, as well as reliability, factor, and Rasch analyses, which showed acceptable reliability and validity of the GAENE for the intended population. A seemingly particularly strong aspect of the development of this survey was its attention to using a Rasch modeling framework from the outset of designing the survey. The items, for instance, were created to elicit a wide range of evolution acceptance levels. However, subsequent research revealed that items on the GAENE that were designed to elicit strong responses had the greatest misfit compared to items from other measures such as the MATE and I-SEA; these include, “*everyone should understand evolution*” and “*nothing in biology makes sense without evolution*” (Romine et al. 2018).

Network meeting identified strengths and weaknesses After evaluation and discussion, network members identified both strengths and weaknesses of the original GAENE and offered a recommendation for its use going forward. Although there is a more recent version of this instrument (i.e., the GAENE 3.0), the authors of the revision recommend that this original version is still suitable for classroom and research purposes; given that it is ten items shorter, this version may be more suitable if a short and less emotionally charged survey is desired. It is an easily administrable survey, and it contains no reverse-coded items, which can decrease response processing errors but may not be as psychometrically ideal. In addition, the instrument provides researchers with a definition of evolution, although some network members suggested it was unclear due to the ambiguity of the wording: “*The mental act or policy of deeming, positing, or postulating...*” Network members noted that this was a broader definition than other measures in which acceptance is defined based primarily on “agreement” rather than this broader characterization of acceptance, which would include behaviors in addition to agreement. It is perhaps for this reason that items on the GAENE depart from other measures in that they include behavior-based inquiries such as “*I would be willing to argue in favor of evolution in a public forum such as a school club, church group, or meeting of public school parents.*” Notably, the intent of the measure according to its authors was not to include an advocacy element to acceptance, but rather to elicit whether a person differentiated between their willingness to engage in that conversation based on the inherent pressures they feel from a variety of sources (e.g., peers as opposed to strangers or family members). While the instrument was not designed to measure advocacy, but rather more extreme viewpoints of evolution acceptance that could manifest as behavioral changes, people who have engaged with the instrument, including network

members, have expressed concern that the instrument measures an advocacy construct in addition to evolution acceptance.

Network members identified several specific weaknesses in this instrument. Items were made to introduce a wide variation in responses, but one of the main issues is that items on this instrument conflate what network members referred to as “perceived importance of evolution” (e.g., “Everyone should understand evolution”) and “advocacy for evolution” (e.g., “I would be willing to argue in favor of evolution in a small group of friends.”). Thus, while researchers can assess acceptance with this measure, researchers will also be capturing perceived importance and advocacy for evolution. While the instrument does provide a definition of acceptance for researchers, evolution is not clearly defined for respondents and may lead to some ambiguity in certain items, such as “*People who plan to become biologists need to understand evolution.*” Past research shows that if evolution is not defined, students will interpret it to be concepts they agree with such as microevolution, which would lead to inflated measures of evolution acceptance (Barnes et al. 2022a). Additionally, some items contain terminology that is potentially ambiguous from a nature of science perspective, such as “*evolution is a scientific fact;*” prior Rasch analysis has also shown this item to have poor fit (Sbeglia and Nehm 2018).

While the GAENE does not specifically address religious concepts or Biblical creation stories, some items could elicit strong emotional responses (e.g., “I would be willing to argue in favor of evolution in a public forum such as a school club, church group, or meeting of public school parents.”). Thus, while this survey has the benefit of not being specific to Judeo-Christian religions like some other measures, the wording of some items may prime a defensive response among more religious students, producing low levels of agreement and suspicion of the survey administrator’s motives. Along those lines, if these items are administered in an undergraduate classroom, it may give students a mistaken impression of the goals or agenda of the instructor, making students resistant to evolution instruction. These issues are likely less important for non-religious students.

Recommendation for GAENE: Because the scope of the GAENE includes advocacy, the network agreed that researchers should use caution when using this measure for evolution acceptance. Our collective consensus definition of evolution acceptance did not include advocacy (Fig. 1), so there is not alignment between our consensus definition and this instrument. However, if one is interested in a broader range of possible evolution acceptance, then the GAENE may be appropriate, particularly if researchers are interested in defining

evolution acceptance as the importance of evolution to respondents or respondents' willingness to advocate for evolution. It could be acceptable in the contexts of populations that are already highly accepting of evolution.

GAENE 3.0 (Glaze et al. 2020)

History, description, and prior critiques The GAENE 3.0 is an extension of the original GAENE, and contains 10 additional items. In line with the Rasch philosophy on which the instrument was built, the authors of the original GAENE intended to provide a new version that would elicit a wider distribution of scores for better discrimination at the ends of the distribution (those with very low and very high acceptance). The new GAENE operates from the same explicit definition of evolution acceptance as the previous GAENE. It is called the GAENE 3.0 because the original version went through multiple iterations. To expand upon the original GAENE, the authors used their own expertise and interviews with five acquaintances known to have strong views on evolution to generate additional items they believed would elicit more variation in the scores. Using the Rasch measurement framework, they showed that the new items did elicit the most extreme responses. The authors also provided evidence of structural, convergent, and divergent validity for the extended measure. The authors suggest that the GAENE 3.0 be used in place of the original GAENE among populations that have high acceptance of evolution (Glaze et al. 2020).

Network meeting identified strengths and weaknesses Network members carefully evaluated and discussed the GAENE 3.0 containing the additional ten items and identified both strengths and weaknesses. As the authors intended, the GAENE 3.0 includes additional items absent from the original that capture a wider range of variation, giving researchers a more normal distribution and more discrimination at the extreme ends of the scale. Thus, the lengthened GAENE 3.0 may be a good measure for high accepting populations for which there will be less emotional reactance to extreme items and the researchers will be more likely to avoid a positively skewed distribution of responses.

However, network members identified that a number of weaknesses present in the original GAENE remain, or are potentially exacerbated, in the GAENE 3.0. In addition to the original importance and advocacy questions, the GAENE 3.0 has particularly provocative language, such as "I would *bet my life* on the claim that evolution is true." Additionally, added items still have terminology that is potentially ambiguous from a nature of science perspective, such as "evolution is true", and items in which "evolution" is not defined so it could be interpreted

as microevolution or macroevolution (e.g., "All evidence supports the claim that evolution is true"). Lastly, with the addition of items, the GAENE 3.0 is longer than the original instrument and thus may pose logistical challenges for administration or survey fatigue.

Similar to the original GAENE, the same issues continue to persist with strong language being even more prevalent (e.g., "Evolution is the most important theory devised by man"), which could prompt extreme reactions from religious students and activate a fear of instructor agendas. Thus, it may not be appropriate for highly religious audiences. Further, some items may conflate students' particular academic interests with their acceptance; a student with aspirations to become an expert in infectious diseases might strongly agree that all of life shares common ancestry but also think Germ Theory is the most important theory devised by man and not evolution.

Recommendation for GAENE 3.0: The network members' recommendations for the GAENE 3.0 are similar to the original instrument, perhaps with added caution for highly religious audiences for whom evolution rejection is prevalent, because many of additional items were meant to elicit even more extreme responses than the original GAENE. The extremes were added because the original instrument had the issue psychometrically that it did not elicit very high or very low responses, so it is important to note that any instrument that gets at the full range may have the risk of eliciting strong responses from some groups.

Evolution Education Questionnaire, Attitudes Towards

Evolution Subscale (ATEVO-EEQ) (Beniermann et al. 2021b)

History, description, and prior critiques The Attitudes Towards Evolution instrument (ATEVO) was initially published in 2019 as part of a German-language book on evolution acceptance (Beniermann 2019). The ATEVO was subsequently incorporated into the Evolution Education Questionnaire on Acceptance and Knowledge (EEQ) in 2021, which includes additional items measuring knowledge. The ATEVO was then translated into 23 languages including English. The ATEVO subscale of the EEQ (henceforth ATEVO-EEQ) consists of eight items scored on a 5-point Likert scale ranging from "agree" to "disagree" with "undecided" as the midpoint. This instrument was designed to measure students' views on evolution in general and on the evolutionary origins of the human mind, with four items on each topic. The ATEVO-EEQ received expert validation from experts in biology, evolution, and philosophy, including people with a range of views on evolution and creation. Further quantitative validation was carried out using a sample of 9,311 participants consisting of high school students, college students,

science teachers, and the general public in Germany; principal component analysis revealed that the ATEVO-EEQ is two-dimensional, with the two clusters corresponding to items about evolution in general and items about the human mind (Beniermann et al. 2021b).

Network meeting identified strengths and weaknesses After careful evaluation and discussion, network members identified both strengths and weaknesses of the ATEVO-EEQ and offered a recommendation for its use going forward. This instrument is unique from all other instruments evaluated in that it measures a second construct: evolutionary origins of the human mind. Thus, it gets at a potentially new dimension of evolution acceptance. In addition, some network members commented positively on the strength of the phrase, “In my personal opinion...”, as this may be an effective way to indicate to respondents that they should answer based on their own views, reducing potential conflation between evolution understanding and evolution acceptance.

That being said, many of these same strengths were also listed as potential weaknesses of the survey. First, measuring the evolutionary origins of the human mind may present particular challenges for students with little to no knowledge of psychology and human evolution. Thus, the ATEVO-EEQ may be less useful for general biology courses, where some students may not have prior exposure to the item topics, and thus lack pre-existing views. In addition, experts raised concerns about two items (“In my personal opinion, our consciousness is a product of natural evolutionary processes” and “In my opinion, our sense of morality is partly the result of natural evolution”); the words “consciousness” and “morality” are ambiguous and may be interpreted differently by different students. These items may be provocative to those who see consciousness and morality as religious/spiritual constructs. Even though the EEQ makes no explicit mention of religion, it may still force the user to “take a side” between religious and scientific explanations of things like consciousness and morality. Lastly, network members noted that no specific definition for evolution acceptance is offered to researchers or respondents.

Recommendation for ATEVO-EEQ: It was the network’s recommendation that the ATEVO-EEQ may not be appropriate for general biology courses or for highly religious populations. It is best suited for respondents in courses focused on the evolution of human behavior and cognition, such as evolutionary psychology.

Single-item instruments

Gallup poll (Gallup Inc 2019)

History, description, and prior critiques Since 1982, Gallup has been polling the United States general public on

their views of evolution by asking respondents to identify which of three statements “come closest to your views on the origin and development of human beings”, with choices of “God created human beings pretty much in their present form at one time within the last 10,000 years or so”, “Humans developed from less advanced forms of life, but God guided this process,” and “Humans developed from less advanced forms of life, but God had no part in this process.” This question has not changed since it was first released in 1982 (Gallup Inc 2019). Education researchers have used the Gallup poll to measure acceptance of evolution among both teachers and students, and both within and outside the United States (Berkman et al. 2008; Berkman and Plutzer 2011; Hanley et al. 2014).

Network meeting identified strengths and weaknesses After careful evaluation and discussion, network members identified both strengths and weaknesses of the Gallup poll. The network members identified that the major strength of the Gallup poll is the possibility of comparison to over 40 years of data. It is also quick to administer (including being able to administer as a clicker question in class) and easy to score since it is a single item.

However, this poll question also has major weaknesses, which greatly limit its utility. The narrowness of the question in asking only about human macroevolution was seen as a major limitation by some members. A second major weakness of this question is that it seems to imply that humans are the pinnacle of evolution, having evolved from “less advanced forms,” which is an inaccurate representation of evolution. Finally, a single question lacks structural validity and there was concern that a single question could not differentiate student levels of acceptance of evolution because it restricts it to a binary of accepting or not.

Considering the relationship between evolution and religion, this item is problematic in that all answer choices are in direct relation to God’s role in the process. This is an issue for individuals who identify as atheists and do not believe in the existence of God because all answer choices assume the existence of God. Further, the specific terminology and capitalization of “God” is representative of only Abrahamic religions (e.g., Christianity and Judaism) and exclusive of any religion that is polytheistic (e.g., Hinduism). Further, most Muslims use the Arabic term for God, which is Allah. In addition, for some religious audiences, there are not enough answer choices. For example, some religious individuals would agree with “God created human beings pretty much in their present form...” but would not agree with the young Earth implications of the second half, “...at one time within the last 10,000 years or so.”

Recommendation for Gallup: Due to the significant weaknesses in this question, we do not recommend its use in evolution education research in biology classes unless one is interested in comparison studies with historical Gallup data. Even then, we encourage caution since 40 years of data does not strengthen the result from a poll with inherent weaknesses.

Pew Research Center poll (Pew Research Center 2019)

History, description, and prior critiques Since 2005, the Pew Research Center has been asking respondents in the United States to select which of three statements comes closest to their own views, with responses of “Humans have evolved over time due to processes such as natural selection; God or a higher power had no role in this process,” “Humans have evolved over time due to processes that were guided or allowed by God or a higher power,” and “Humans have existed in their present form since the beginning of time.” Early versions of the poll involved a two-step process with an initial question asking whether respondents agreed that “Humans have evolved over time” or if “Humans have existed in their present form since the beginning of time”, with distinction between the theistic and non-theistic evolutionary positions solicited only after evolution had been chosen in the first step. In 2019, Pew reported on an experiment in which the prior two-step polling process was compared with a single item allowing respondents to select from the three main categories simultaneously with a single-question approach (Pew Research Center 2019). Differences in responses between the two methodologies were stark, calling reported results from previous polling efforts into question, as far more (31% vs. 18%) respondents selected the creationist view when they were not first offered a theistic evolution option. The Pew Research Center has said that all subsequent surveys will use the single-step question with the three choices.

Network meeting identified strengths and weaknesses After careful evaluation and discussion, network members identified both strengths and weaknesses of the Pew Research Center poll. The strengths of the items are similar to those of the Gallup poll in that nationwide longitudinal data are available for comparisons and being a single item, it is easy to administer.

However, it suffers from similar weaknesses. The network members identified that in addition to the limitations of a single item survey to measure a construct as complicated as evolution acceptance, the major weakness of the Pew is that it is not explicit about what the term “evolved” means, which could imply either macroevolution or microevolution. The reference to natural selection in the question implies a focus on microevolution, which

is often not thought to be sufficient to say something about evolution acceptance. Additional wording issues are present in the item, including ambiguity regarding what is meant by humans having “existed in their present form since the beginning of time,” which different kinds of special creationists might not entirely agree with whether they conceive of a divine creation of humanity only 6 days after the origin of the universe or millions of years after the beginning of life on Earth.

Considering the relationship between evolution and religion, the question implies the existence of God in either of the choices that would be considered scientifically accurate, which is exclusive of atheists, agnostics, and non-Christian religions. Additionally, there are not enough answer options to capture the range of views. For example, it lacks a God-guided view that is less interventionist (e.g., God created the natural laws and set them in motion but did not guide the process).

Recommendation for Pew: Due to the significant weaknesses in this question, we do not recommend its use in evolution education research in biology classes unless one is interested in comparison studies with historical Pew data. Similar to the Gallup poll, some members noted this could be useful for demographic description purposes.

Miller 2006 poll

History, description, and prior critiques Through a series of national and international public opinion surveys with various collaborators, Jon D. Miller from the Institute for Social Research at the University of Michigan has been studying evolution acceptance among respondents since 1985. Miller and colleagues have been asking members of the general public whether the statement “Human beings, as we know them today, developed from earlier species of animals,” is true, false, or if they are not sure. Among the more widely cited studies using this item, a 2006 report from Miller, Scott, and Okamoto compared evolution acceptance among respondents from 32 European countries, Japan, and the United States (Miller et al. 2006). Prior critique of this measure has highlighted the reliability shortcomings of using a single-item measure, and the potential validity issues of framing the answers choices as true or false, which may cause confusion about whether the item is asking for personal opinion or scientific consensus (Smith and Siegel 2016).

Network meeting identified strengths and weaknesses After careful evaluation and discussion, network members identified both strengths and weaknesses with this poll item. The network members thought that this question was better than either of the other single-item questions because it did not explicitly mention God and it

focused only on the macroevolution of humans. However, weaknesses included concern about the wording of this question, including what “developed” means and that it was presented as a true or false question, which implies a “right” answer as opposed to an opinion.

Considering the relationship between evolution and religion, this item is noticeably without any mention of God and thus could be more appropriate for heterogeneous religious populations. However, similar to the issue with the original Pew two-question format, by not offering any theistic evolutionary options, this item might underestimate evolution acceptance among highly religious populations.

Recommendation for Miller: If evolution education researchers require one question to ask about evolution acceptance due to survey fatigue, then we recommend this question, except perhaps in highly religious populations in which it may underestimate evolution acceptance. However, we do not advocate the use of this single-item instrument generally to measure evolution acceptance because of the combined limitations of the lack of construct validity due to one question, concerns that it cannot differentiate among levels of evolution acceptance, and problems in its wording.

Aligning the decision of which instrument to use to measure evolution acceptance with your goals for research

The network had broader discussion about instrument limitations. Overall, we noted the overwhelming number of instruments present in the literature makes it difficult to discern which instrument to use and the different foci of each instrument makes it difficult to compare results across studies (Barnes et al. 2019; Beniermann et al. 2023; Lloyd-Strovas and Bernal 2012). Further, the large number of instruments could create challenges to understand which intervention is most effective; it is hard to know if any difference in outcomes is due to differences in the intervention, differences in the populations, or differences in instrumentation. Additionally, definitions of evolution acceptance vary across instruments and items, and this limits the generalizability and comparability of studies (Barnes et al. 2019; Beniermann et al. 2023). While having multiple instruments that are slightly different can strengthen the evidence for a result, if the instruments are too different from one another, this could lead to disparate results. Some studies showed different measures of evolution acceptance lead to similar results (Metzger et al. 2018) but in other studies multiple instruments have led to different study results and conclusions (Barnes et al. 2019) indicating a need for greater standardization of measurement in evolution education studies. Although having some variation

in measurement tools can provide greater fit to the study context and increase validity of the data, having too many instruments can lead to varying results and conclusions, especially when they operate on different definitions of evolution acceptance.

Further, all network members noted that there are limitations to each instrument, with many members noting specific problems with questions in the existing set of instruments. Some of these critiques were about specific wording, but other critiques were more extensive and related to the lack of clarity for survey takers about a definition of evolution in the survey. Specifically, network members agreed that students taking evolution surveys often cannot tell whether the word “evolution” is referring to micro- or macroevolution or if it includes humans, so different survey respondents may be answering based on a different conception of evolution (Barnes et al. 2022a; Nadelson and Southerland 2012). Since students tend to accept microevolution more than macroevolution or human evolution (Barnes et al. 2019, 2022b), items that do not specify the context may lead one student to answer based on microevolution and another student to answer based on human evolution. Two students answering the same item based on a different context has been shown to lead to different evolution acceptance scores among students who have a similar acceptance of evolution (Barnes et al. 2022a). Further, network members indicated that several evolution acceptance instrument items conflate acceptance and knowledge/understanding or conflate one’s understanding of the nature of science and evolution acceptance (Smith et al. 2016). Since some instruments have knowledge/understanding and nature of science constructs embedded within their items, studies that use these instruments may show inflated relationships among these constructs and evolution acceptance. When researchers report “evolution acceptance” scores as part of their research project, few people reading the paper will critically examine the instrument; the research community tends to assume the instrument is valid and reliable, but the extent of validity and reliability is likely different between instruments and populations. Thus, network members collectively agreed that instrumentation reform is desired and necessary to drive this work forward in a meaningful and productive way.

Given that instrumentation reform is lengthy and new instrumentation will not be readily available to researchers in the short term, network members made an effort to offer the evolution education research community guidance on how best to use the current instruments available to measure evolution acceptance. As a group activity, we created a guide for researchers to use to make decisions on which instruments might be most appropriate to a given study population, research

Table 2 A table to guide researchers to which evolution acceptance instrument could be best suited for their population and data collection needs

	Original MATE	MATE 2.0	I-SEA	Original GAENE	GAENE 3.0	EEQ	Gallup	Pew	Miller
For highly religious populations		X	X						
For non-religious populations		X	X	X	X				
For populations high in evolution acceptance		X	X	X	X				
For populations low in evolution acceptance		X	X						
For students with limited biology knowledge		X		X	X				
For students with prior biology knowledge		X	X	X	X				
For measuring macroevolution acceptance alone		X	X						
For differentiating between micro, macro, and human evolution acceptance			X						
For measuring acceptance of evolutionary psychology						X			
For comparisons with other studies using the same measure	X	X	X	X	X	X	X	X	X

An "X" indicates which instruments would be appropriate for each category, based on the consensus of the meeting participants

Table 3 Features of each evolution acceptance instrument that researchers may consider when deciding which evolution acceptance instrument could be best suited for their needs

	Original MATE	MATE 2.0	ISEA	Original GAENE	GAENE 3.0	EEQ	Gallup	Pew	Miller
# of items	20	9	24	13	22	8	1	1	1
Contains reverse coded items	Yes	Yes	Yes	No	No	Yes	No	No	No
Scale	5 pt Likert, SA-SD*	5 pt Likert, SA-SD	5 pt Likert, SA-SD	5 pt Likert, SA-SD	5 pt Likert, SA-SD	5 pt Likert, D-A	None: Choose closest view	None: Choose closest view	None: True, False, Not sure
Scale mid-point	Undecided	Neutral	Undecided	No opinion	No opinion	Undecided	N/A	N/A	N/A
Conflates religion	Yes	No	No	No	No	No	Yes	Yes	No
Conflates understanding of evolution	Yes	No	Yes	No	No	Yes	No	Yes	No
Conflates nature of science	Yes	Yes	No	Yes	Yes	No	No	No	No
Conflates advocacy	No	No	No	Yes	Yes	No	No	No	No
Contains an unspecified use of the word "evolution"	Yes	No	Yes	Yes	Yes	Yes	No	Yes	No
Includes a definition of evolution acceptance	No	Yes	No	Yes	Yes	No	No	No	No

SA: Strongly Agree; SD: Strongly Disagree; A: Agree; D: Disagree

design, and stated research outcomes (Table 2). Additionally, Table 3 offers an overview of instrument characteristics that can be considered as one determines suitability for their own needs. Each instrument in its entirety can be found in the Additional file 1.

Network consensus definition of evolution acceptance

One reason for inconsistency in measurement of evolution acceptance is likely due to the different definitions that researchers use to create evolution acceptance measures (Ingram and Nelson 2006; National Academy of Sciences 1998, 2008; Sinatra et al. 2003; Smith 1994; Smith and Scharmann 1999; Southerland et al. 2001; Southerland and Sinatra 2003; Wiles 2014) and multiple researchers have suggested that if we are to measure evolution acceptance in a way that provides valid inferences and reliable comparisons between studies, the field needs to agree on a consensus definition of evolution acceptance (Barnes et al. 2019; Beniermann et al. 2023). As of now, the published definitions of evolution acceptance that underly current measures (Table 4) have not been reviewed by a large network of evolution acceptance experts. Given we had 16 evolution acceptance researchers, the network members set out to meet this need and identify a network consensus definition of evolution acceptance that the field may be able to use going forward for instrument development.

Why does a consensus definition matter?

A foundational step of creating an instrument to measure a construct is to first define that construct (AERA et al. 2014, pp. 75–78) and if definitions among instruments vary substantially, this could lead to inconsistencies in the measurement of evolution acceptance between instruments. Researchers have used different definitions of “evolution acceptance,” many of which have been distinguished from “believing in evolution” (Ingram and Nelson 2006; National Academy of Sciences 1998, 2008; Sinatra et al. 2003; Smith 1994; Smith and Scharmann 1999; Southerland et al. 2001; Southerland & Sinatra 2003; Wiles 2014). Believing in evolution implies a subjective judgement based on faith similar to believing in religion. Thus, evolution education researchers have pushed to distinguish believing in evolution from accepting evolution by defining acceptance as based on “a systematic evaluation of the evidence” leading to “a learner’s personal assessment of the validity of [evolution]” (Sinatra et al. 2003, p. 512). However, requiring a systematic evaluation of the evidence for acceptance of evolution creates a substantial overlap between the constructs of understanding and accepting evolution. Understanding evolution is the degree to which someone has a good conceptual grasp of evolutionary theory and has substantial knowledge of facts about evolution. Students can have a good understanding of evolution, score very well on evolution exams, and yet still reject the veracity of evolutionary theory (Hermann 2012). Thus, evolution education researchers usually distinguish understanding

Table 4 Definitions of evolution acceptance from authors of instruments evaluated by the network

Instrument	Author’s definitions of evolution acceptance
Original MATE	Authors do not provide an explicit definition but says acceptance includes: 1) the processes of evolution, (2) the available evidence of evolutionary change, (3) the ability of evolutionary theory to explain phenomena, (4) the evolution of humans, (5) the age of the earth, (6) the independent validity of science as a way of knowing, and (7) the current status of evolutionary theory within the scientific community
MATE 2.0	Authors provide an explicit definition of acceptance of evolution in the article as “ <i>The agreement that it is scientifically valid that all species have evolved from prior species.</i> ”
I-SEA	Authors do not provide a single definition of evolution acceptance, but they define acceptance broadly: “ <i>Acceptance of a construct is based on an examination of the validity of the knowledge supporting the construct, the plausibility of the construct for explaining phenomenon, persuasiveness of the construct, and fruitfulness or productivity of the empirical support for the construct</i> ”
Original GAENE	Authors provide an explicit definition of acceptance of evolution in the article as “ <i>Evolution acceptance is 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.</i> ”
GAENE 3.0	Authors provide an explicit definition of acceptance of evolution in the article as “ <i>Evolution acceptance is 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.</i> ”
EEQ-ATEVO	Authors do not provide an explicit definition of evolution acceptance, but relate it to “ <i>a positive attitude toward evolution</i> ”
Gallup	None given
Pew	None given
Miller	Authors define acceptance of evolution as agreement with the statement “ <i>Human beings, as we know them today, developed from earlier species of animals.</i> ”

and acceptance of evolution as different constructs that may be related to one another, depending on the population of students. For instance, there is some indication that among more highly religious populations, there are more people who understand evolution but do not personally think evolution is plausible (Weisberg et al. 2018). Further, the relationship between understanding and acceptance can be different depending on acceptance of microevolution, macroevolution, and human evolution. When students are considering whether they accept microevolution, their understanding of evolution seems to matter more than when they evaluate macroevolution or human evolution (Barnes et al. 2019). Finally, some items on current instruments seem to imply a definition of evolution acceptance that includes or excludes religious belief, but evolution acceptance is arguably separate from religious belief and should not include any reference to religion. Thus, when researchers include concepts and measures that conflate understanding, religious beliefs and acceptance, the conclusions they make from their data may be misinformed.

Network derived consensus definition of evolution acceptance

With grounding in the prior literature, we wanted to document network member views on the definition of acceptance of evolution to work toward a consensus definition.

Methods for determining consensus definition of evolution acceptance

As mentioned previously, each network member was sent a bibliography of peer reviewed published papers of the instruments reviewed in this study as well as the most up to date research at that time on evolution acceptance measurement (the bibliography included: Barnes et al. 2019; Barnes et al. 2022a; Beniermann et al. 2021b; Glaze et al. 2020; Mead et al. 2019; Nadelson and Southerland 2012; Romine et al. 2018; Rutledge and Sadler 2007; Rutledge and Warden 1999; Sbeglia and Nehm 2018, 2019; Smith 2009; Smith et al. 2016). Network members then completed an individual survey that asked an open-ended question about their definition of evolution acceptance. The Co-PI team constructed an initial draft of a definition based on the similarities in network participants' responses. We then had a group discussion

at the meeting about the definition to see if any additional points would be raised. Surprisingly, there was a high degree of similarity in network member individual responses and little disagreement about what should and should not be included in a definition of evolution acceptance. Network members were in consensus that evolution acceptance can be defined as a person's agreement that evolution is valid, and the best explanation from science for the unity and diversity of life on Earth. Network members also agreed on aspects of evolution that needed to be accepted, including speciation, the common ancestry of life, and that humans share this common ancestry. In Fig. 1 we present our final definition of evolution acceptance that all network members agreed to.

Future research and recommendations

Through our network, we identified directions for future research specifically related to evolution acceptance measurement that could be uniquely addressed by a collaborative network. One direction that could improve evolution acceptance measurement would be to revise the instruments (Barnes et al. 2022a; Beniermann et al. 2023; Dunk et al. 2019; Kuschmierz et al. 2020; Sbeglia and Nehm 2018, 2019; Smith 2010). Much of what the network identified as problematic within instruments was the wording of individual items that was unclear or seemed to measure more than evolution acceptance alone. Some of these issues likely exist because of a limited number of perspectives that were incorporated into the initial draft, including a small number of initial cognitive interviews to determine whether students were interpreting the items the way they were intended. These issues will not be resolved using statistical techniques that prior studies have tried to use to establish validity evidence of evolution acceptance measures (Barnes et al. 2019; Metzger et al. 2018; Romine et al. 2018; Sbeglia and Nehm 2018, 2019) because statistical techniques will not identify response process errors. Rasch analysis can potentially identify response process errors by looking at person misfit (Wright and Stone 1979), but to our knowledge few researchers have published these analyses with evolution acceptance instruments (Beniermann et al. 2023). Our network began to address these issues by including expert reviews from within the network composed of individuals from different disciplines, religious

Evolution acceptance: agreeing 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 non-human ancestors.

Fig. 1 Network members' consensus definition of evolution acceptance

backgrounds, religious expertise, and evolutionary biology knowledge. An additional way that these issues should be systematically identified and resolved is to identify response process errors through cognitive interviews with students from diverse populations by having network members work together to recruit and interview students from different demographics, including different religious backgrounds. A similar approach was taken when revising the original MATE instrument to create the MATE 2.0, which largely reduced the number of response process errors from students (Barnes et al. 2022a), but the instrument would have benefited from even broader perspectives of experts and more student voices incorporated.

Further, the field could normalize the refinement and revision process for instruments as well as discourage the publication of data from instruments that have been heavily critiqued. The progress in standardized measurement of evolution acceptance has been significant, but extremely slow over the last 30 years. For two decades, most researchers were using their own unique instruments to measure evolution acceptance and in the last decade we have seen the majority of studies conducted using the original MATE, despite continuous criticisms of this instrument. Although, the MATE instrument was heavily critiqued it was used for more than 20 years before it was revised. The fact that the original GAENE was revised only 4 years after its initial publication indicates that some researchers are recognizing the need for refinement of these instruments, but even the GAENE was revised without considering criticisms of the first publication (Sbeglia and Nehm 2018). So, it is important that we not only normalize the refinement and revision process, but that our refinement and revision is based upon identified areas of consensus regarding the definition of evolution acceptance within the field of evolution education research.

A tradeoff to revising instruments more often is that studies using the revised instruments will be slightly less comparable to older studies. However, we argue this is more desirable than continuing to use instruments with known significant flaws. Based on our network findings, we encourage researchers to continue to think about how and when to improve existing instruments and argue that a larger network collaboration would be uniquely positioned to achieve this effectively because network members could work together to pilot the revised instrument, achieving more generalizable validity evidence.

Another direction that the field could take would be to create an entirely new instrument by collating existing items to make an instrument with multiple sub-scales and/or by creating new items under a common definition and a broad team to devise the new additional items.

While either approach would be time-intensive, it could represent a collective effort to clarify further what the community's consensus might be on how to define and measure evolution acceptance.

Conclusions

Through the initial establishment of this network of interdisciplinary scholars with intentionally different viewpoints, experiences, backgrounds, and expertise relevant to studying evolution acceptance, we were able to offer expert-based consensus recommendations on current instruments and a definition of evolution acceptance. Additionally, by capitalizing on the diversity of network member experiences with and viewpoints on religion, we were able to identify specific limitations of evolution acceptance instruments for students with different religious identities. These collective perspectives offer a cohesive viewpoint, which hopefully can be both a resource and a guide for the next steps in evolution acceptance instrument design and revision. We hope to build from the initial momentum of this network to grow the network and usher in a new era of evolution education research that uses collaborative methods among a diverse team to create more cohesive efforts to measure and improve evolution acceptance. We encourage others who are interested to reach out to the corresponding author and join this growing community of scholars.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12052-024-00194-0>.

Additional file 1: Instruments evaluated.

Acknowledgements

We would like to thank ASU's RISE Center and specifically Lynn Highton for her help with the logistics for the meeting.

Author contributions

The leadership team for this network consisted of SEB, MEB, JJ, JRW and TM; they organized the RCN-UBE meeting and wrote the meeting report. RDPD, EAH, JRK, LSN, WLR, AS, ALT, and YZ participated in the in-person meeting along with the leadership team. All authors participated in pre-meeting activities and reviewed the manuscript.

Funding

This RCN-UBE meeting was funded by the National Science Foundation RCN-UBE grant #2018558. Additional funding for this publication was provided by an award from the Society for the Study of Evolution.

Availability of data and materials

All surveys reviewed in this meeting, including prompts, scales, and items, are available in the Additional file 1.

Declarations

Ethics approval and consent to participate

Data was not collected from research participants, therefore ethical approval and consent for publication of individual data were not applicable.

Competing interests

Several network members have authored one of these instruments (MATE 2.0: M.E.B., T.M., S.E.B.; I-SEA: L.S.N.; GAENE 3.0: A.L.T.), which represents both a source of information and a conflict of interest. To try to minimize this concern, no author of an instrument was part of the group to review that instrument, although everyone participated in individual feedback and the group discussion.

Received: 3 January 2024 Accepted: 4 January 2024

Published online: 22 January 2024

References

- AERA, APA, NCME. The standards for educational and psychological testing. American Educational Research Association. 2014.
- Asghar A. Canadian and Pakistani Muslim teachers' perceptions of evolutionary science and evolution education. *Evol Educ Outreach*. 2013;6(1):10. <https://doi.org/10.1186/1936-6434-6-10>.
- Athanasiou K, Papadopoulou P. Conceptual ecology of the evolution acceptance among Greek education students: knowledge, religious practices and social influences. *Int J Sci Educ*. 2012;34(6):903–24. <https://doi.org/10.1080/09500693.2011.586072>.
- Austriaco N. A theological fittingness argument for the evolution of *Homo sapiens*. *Theol Sci*. 2019;17(4):539–50.
- Baker JO. Acceptance of evolution and support for teaching creationism in public schools: the conditional impact of educational attainment. *J Sci Study Relig*. 2013;52(1):216–28. <https://doi.org/10.1111/jsr.12007>.
- Baker JO, Rogers D, Moser T. Acceptance of evolution among American Mormons. *J Contemp Relig*. 2018;33(1):123–34. <https://doi.org/10.1080/13537903.2018.1408295>.
- Barnes ME, Dunlop HM, Holt EA, Zheng Y, Brownell SE. Different evolution acceptance instruments lead to different research findings. *Evol Educ Outreach*. 2019;12(1):4. <https://doi.org/10.1186/s12052-019-0096-z>.
- Barnes ME, Dunlop HM, Sinatra GM, Hendrix TM, Zheng Y, Brownell SE. "Accepting Evolution Means You Can't Believe in God": Atheistic perceptions of evolution among college biology students. *CBE Life Sci Educ*. 2020;19(2): ar21. <https://doi.org/10.1187/cbe.19-05-0106>.
- Barnes ME, Roberts JA, Maas SA, Brownell SE. Muslim undergraduate biology students' evolution acceptance in the United States. *PLoS ONE*. 2021a;16(8): e0255588. <https://doi.org/10.1371/journal.pone.0255588>.
- Barnes ME, Supriya K, Zheng Y, Roberts JA, Brownell SE. A new measure of students' perceived conflict between evolution and religion (PCoRE) is a stronger predictor of evolution acceptance than understanding or religiosity. *CBE Life Sci Educ*. 2021b;20(3): ar42. <https://doi.org/10.1187/cbe.21-02-0024>.
- Barnes ME, Misheva T, Supriya K, Rutledge M, Brownell SE. A revised measure of acceptance of the theory of evolution: introducing the MATE 2.0. *CBE Life Sci Educ*. 2022a;21(1): ar10. <https://doi.org/10.1187/cbe.21-05-0127>.
- Barnes ME, Riley R, Bowen C, Cala J, Brownell SE. Community college student understanding and perceptions of evolution. *CBE Life Sci Educ*. 2022b;21(3): ar46. <https://doi.org/10.1187/cbe.21-09-0229>.
- Beniermann A. *Evolution – von Akzeptanz und Zweifeln*. Wiesbaden: Springer; 2019. <https://doi.org/10.1007/978-3-658-24105-6>.
- Beniermann A, Kuschmier P, Pinxten R, Aivelo T, Bajrami A, Bohlin G, Brennecke JS, Cebesoy UB, Cvetković D, Đorđević M, Dvořáková RM, Futo M, Nicoleta G, Korfiatis K, Lendvai A, Mogias A, Paolucci S, Petersson M, Pietrzak B, et al. Evolution Education Questionnaire on Acceptance and Knowledge (EEQ)—Standardised and ready-to-use protocols to measure acceptance of evolution and knowledge about evolution in an international context. 2021a; Zenodo. <https://doi.org/10.5281/zenodo.7779165>
- Beniermann A, Kuschmier P, Pinxten R, Aivelo T, Bohlin G, Brennecke JS, Cebesoy UB, Cvetković D, Đorđević M, Dvořáková RM, Futo M, Geamana N, Korfiatis K, Lendvai A, Mogias A, Paolucci S, Petersson M, Pietrzak B, Porozovs J, et al. Evolution Education Questionnaire on Acceptance and Knowledge (EEQ)—Standardised and ready-to-use protocols to measure acceptance of evolution and knowledge about evolution in an international context. 2021b. <https://helda.helsinki.fi/handle/10138/333584>
- Beniermann A, Moormann A, Fiedler D. Validity aspects in measuring evolution acceptance: evidence from surveys of preservice biology teachers and creationists. *J Res Sci Teach*. 2023;60(6):1223–65. <https://doi.org/10.1002/tea.21830>.
- Berkman MB, Plutzer E. Defeating creationism in the courtroom, but not in the classroom. *Science*. 2011;331(6016):404–5. <https://doi.org/10.1126/science.1198902>.
- Berkman MB, Pacheco JS, Plutzer E. Evolution and creationism in America's classrooms: a national portrait. *PLoS Biol*. 2008;6(5): e124. <https://doi.org/10.1371/journal.pbio.0060124>.
- Bowen CD, Summersill AR, Jensen JL, Brownell SE, Barnes ME. A comparison of online and in-person evolution instruction that includes religious cultural competence. *J Microbiol Biol Educ*. 2022. <https://doi.org/10.1128/jmbe.00067-22>.
- Carter BE, Wiles JR. Scientific consensus and social controversy: exploring relationships between students' conceptions of the nature of science, biological evolution, and global climate change. *Evol Educ Outreach*. 2014;7(1):6. <https://doi.org/10.1186/s12052-014-0006-3>.
- Deniz H, Donnelly LA, Yilmaz I. Exploring the factors related to acceptance of evolutionary theory among Turkish preservice biology teachers: toward a more informative conceptual ecology for biological evolution. *J Res Sci Teach*. 2008;45(4):420–43. <https://doi.org/10.1002/tea.20223>.
- Dole JA, Sinatra GM. Reconceptualizing change in the cognitive construction of knowledge. *Educ Psychol*. 1998;33(2–3):109–28.
- Dunk RDP, Wiles JR. 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. *bioRxiv*. 2018. <https://doi.org/10.1101/280479>.
- Dunk RDP, Petto AJ, Wiles JR, Campbell BC. A multifactorial analysis of acceptance of evolution. *Evol Educ Outreach*. 2017;10(1):4. <https://doi.org/10.1186/s12052-017-0068-0>.
- Dunk RDP, Barnes ME, Reiss MJ, Alters B, Asghar A, Carter BE, Cotner S, Glaze AL, Hawley PH, Jensen JL, Mead LS, Nadelson LS, Nelson CE, Pobiner B, Scott EC, Shtulman A, Sinatra GM, Southerland SA, Walter EM, Brownell SE, Wiles JR. Evolution education is a complex landscape. *Nat Ecol Evol*. 2019;3(3):327–9. <https://doi.org/10.1038/s41559-019-0802-9>.
- Eddy SL, Crowe AJ, Wenderoth MP, Freeman S. How should we teach tree-thinking? An experimental test of two hypotheses. *Evol Educ Outreach*. 2013;6(1):13. <https://doi.org/10.1186/1936-6434-6-13>.
- Elsdon-Baker F. Creating creationists: the influence of 'issues framing' on our understanding of public perceptions of clash narratives between evolutionary science and belief. *Public Underst Sci*. 2015;24(4):422–39. <https://doi.org/10.1177/0963662514563015>.
- Evans EM. Cognitive and contextual factors in the emergence of diverse belief systems: creation versus evolution. *Cogn Psychol*. 2001;42(3):217–66. <https://doi.org/10.1006/cogp.2001.0749>.
- Evans EM. Conceptual change and evolutionary biology: a developmental analysis. *International Handbook of Research on Conceptual Change*. 2008, 263–294.
- Evans EM, Lane JD. Contradictory or complementary? Creationist and evolutionist explanations of the origin(s) of species. *Hum Dev*. 2011;54(3):144–59. <https://doi.org/10.1159/000329130>.
- Fiedler D, Sbeglia GC, Nehm RH, Harms U. How strongly does statistical reasoning influence knowledge and acceptance of evolution? *J Res Sci Teach*. 2019. <https://doi.org/10.1002/tea.21547>.
- Gervais WM. Override the controversy: Analytic thinking predicts endorsement of evolution. *Cognition*. 2015;142:312–21. <https://doi.org/10.1016/j.cognition.2015.05.011>.
- Glaze AL, Goldston MJ, Dantzer J. Evolution in the Southeastern USA: factors influencing acceptance and rejection in pre-service science teachers. *Int J Sci Math Educ*. 2014;13(6):1189–209. <https://doi.org/10.1007/s10763-014-9541-1>.

- Glaze AL, Snyder S, Devereaux RA, Smith MU. The Generalized acceptance of Evolution Evaluation (GAENE) 3.0: enhancement and validation of a new measure of acceptance. *J Higher Educ Theory Pract*. 2020. <https://doi.org/10.33423/jhetp.v20i7.3157>.
- Graves JL. African Americans in evolutionary science: where we have been, and what's next. *Evol Educ Outreach*. 2019;12(1):18. <https://doi.org/10.1186/s12052-019-0110-5>.
- Green K, Delgado C. Crossing cultural borders: results of an intervention on community college biology students' understanding and acceptance of evolution. *Int J Sci Educ*. 2021;43(4):469–96. <https://doi.org/10.1080/09500693.2020.1869854>.
- Gutowski R, Aptyka H, Großschedl J. An exploratory study on students' denominations, personal religious faith, knowledge about, and acceptance of evolution. *Evol Educ Outreach*. 2023;16(1):9. <https://doi.org/10.1186/s12052-023-00187-5>.
- Hanley P, Bennett J, Ratcliffe M. The inter-relationship of science and religion: a typology of engagement. *Int J Sci Educ*. 2014;36(7):1210–29. <https://doi.org/10.1080/09500693.2013.853897>.
- Heddy BC, Nadelson LS. A global perspective of the variables associated with acceptance of evolution. *Evol Educ Outreach*. 2012;5(3):412–8. <https://doi.org/10.1007/s12052-012-0423-0>.
- Heddy BC, Nadelson LS. The variables related to public acceptance of evolution in the United States. *Evol Educ Outreach*. 2013;6(1):3. <https://doi.org/10.1186/1936-6434-6-3>.
- Hermann RS. Cognitive apartheid: on the manner in which high school students understand evolution without believing in evolution. *Evol Educ Outreach*. 2012;5(4):619. <https://doi.org/10.1007/s12052-012-0446-6>.
- Hill JP. Rejecting evolution: the role of religion, education, and social networks. *J Sci Study Relig*. 2014;53(3):575–94. <https://doi.org/10.1111/jssr.12127>.
- Holt EA, Ogden TH, Durham SL. The positive effect of role models in evolution instruction. *Evol Educ Outreach*. 2018;11(1):11. <https://doi.org/10.1186/s12052-018-0086-6>.
- Gallup Inc. Evolution, creationism, intelligent design. Gallup.Com. 2019. <https://news.gallup.com/poll/21814/Evolution-Creationism-Intelligent-Design.aspx>
- Ingram EL, Nelson CE. Relationship between achievement and students' acceptance of evolution or creation in an upper-level evolution course. *J Res Sci Teach*. 2006;43(1):7–24. <https://doi.org/10.1002/tea.20093>.
- Jensen JL, Manwaring KF, Gill RA, Sudweeks RS, Davies RS, Olsen JA, Phillips AJ, Bybee SM. Religious affiliation and religiosity and their impact on scientific beliefs in the United States. *Bioscience*. 2019;69(4):292–304. <https://doi.org/10.1093/biosci/biz014>.
- Kuschmierz P, Meneganzin A, Pinxten R, Pievani T, Cvetković D, Mavrikaki E, Graf D, Beniermann A. Towards common ground in measuring acceptance of evolution and knowledge about evolution across Europe: a systematic review of the state of research. *Evol Educ Outreach*. 2020;13(1):18. <https://doi.org/10.1186/s12052-020-00132-w>.
- Leicht C, Sharp CA, LaBouff JP, Zarzeczna N, Elsdon-Baker F. Content matters: perceptions of the science-religion relationship. *Int J Psychol Relig*. 2022;32(3):232–55. <https://doi.org/10.1080/10508619.2021.2003111>.
- Lindsay J, Arok A, Bybee SM, Cho W, Cordero AM, Ferguson DG, Galante LL, Gill R, Mann M, Peck SL, Shively CL, Stark MR, Stowers JA, Tenneson M, Tolman ER, Wayment T, Jensen JL. Using a reconciliation module leads to large gains in evolution acceptance. *CBE Life Sci Educ*. 2019;18(4):ar58. <https://doi.org/10.1187/cbe.19-04-0080>.
- Lloyd-Strovas JD, Bernal XE. A review of undergraduate evolution education in U.S. Universities: building a unifying framework. *Evol Educ Outreach*. 2012;5(3):435. <https://doi.org/10.1007/s12052-012-0435-9>.
- Loke ATE. Reconciling evolution and biblical literalism: A proposed research program. *Theol Sci*. 2016;14(2):160–74.
- McGrath AE. A consilience of equal regard: Stephen Jay Gould on the relation of science and religion. *Zygon*. 2021;56(3):547–65.
- McPhetres J, Zuckerman M. Religiosity predicts negative attitudes towards science and lower levels of science literacy. *PLoS ONE*. 2018;13(11):e0207125. <https://doi.org/10.1371/journal.pone.0207125>.
- McPhetres J, Jong J, Zuckerman M. Religious Americans have less positive attitudes toward science, but this does not extend to other cultures. *Soc Psychol Personal Sci*. 2021;12(4):528–36. <https://doi.org/10.1177/1948550620923239>.
- Mead LS, Kohn C, Warwick A, Schwartz K. Applying measurement standards to evolution education assessment instruments. *Evol Educ Outreach*. 2019;12(1):5. <https://doi.org/10.1186/s12052-019-0097-y>.
- Metzger KJ, Montplaisir D, Haines D, Nickodem K. Investigating undergraduate health sciences students' acceptance of evolution using MATE and GAENE. *Evol Educ Outreach*. 2018;11(1):10. <https://doi.org/10.1186/s12052-018-0084-8>.
- Miller JD, Scott EC, Okamoto S. Public acceptance of evolution. *Science*. 2006;313(5788):765–6. <https://doi.org/10.1126/science.1126746>.
- Misheva T, Brownell SE, Barnes ME. "It's More Of A Me-Thing Than An Evolution Thing": exploring the validity of evolution acceptance measures using student interviews. *CBE Life Sci Educ*. 2023;22(4):1–13. <https://doi.org/10.1187/cbe.23-01-0022>.
- Moore R, Cotner S. The creationist down the hall: does it matter when teachers teach creationism? *Bioscience*. 2009;59(5):429–35. <https://doi.org/10.1525/bio.2009.59.5.10>.
- Nadelson LS, Southerland SA. Examining the interaction of acceptance and understanding: how does the relationship change with a focus on macroevolution? *Evol Educ Outreach*. 2010;3(1):82–8. <https://doi.org/10.1007/s12052-009-0194-4>.
- Nadelson LS, Southerland SA. A more fine-grained measure of students' acceptance of evolution: development of the inventory of student evolution acceptance—I-SEA. *Int J Sci Educ*. 2012;34(11):1637–66. <https://doi.org/10.1080/09500693.2012.702235>.
- National Academy of Sciences. Teaching about evolution and the nature of science. Washington: The National Academies Press; 1998. <https://doi.org/10.17226/5787>.
- National Academy of Sciences (2008) Science, Evolution, and Creationism. <https://doi.org/10.17226/11876>
- Nehm RH, Ha M. Item feature effects in evolution assessment. *J Res Sci Teach*. 2011;48(3):237–56. <https://doi.org/10.1002/tea.20400>.
- Nehm RH, Mead LS. Evolution assessment: Introduction to the special issue. *Evol Educ Outreach*. 2019;12(1):7. <https://doi.org/10.1186/s12052-019-0098-x>.
- Nehm RH, Schonfeld IS. Does increasing biology teacher knowledge of evolution and the nature of science lead to greater preference for the teaching of evolution in schools? *J Sci Teacher Educ*. 2007;18(5):699–723. <https://doi.org/10.1007/s10972-007-9062-7>.
- Pew Research Center (2019) The Evolution of Pew Research center's survey questions about the origins and development of life on earth. <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/>
- Pew. Public's views on human evolution. Pew Research Center's Religion & Public Life Project. 2013. <http://www.pewforum.org/2013/12/30/public-views-on-human-evolution/>
- Pobiner B. Accepting, understanding, teaching, and learning (human) evolution: obstacles and opportunities. *Am J Phys Anthropol*. 2016;159(S61):232–74. <https://doi.org/10.1002/ajpa.22910>.
- Rachmatullah A, Nehm RH, Roshayanti F, Ha M. Evolution education in Indonesia: Pre-service biology teachers' knowledge, reasoning models, and acceptance of evolution. In: *Evolution education around the globe*. Springer; 2018, 335–355.
- Reznick DN, Ricklefs RE. Darwin's bridge between microevolution and macroevolution. *Nature*. 2009;457(7231):837. <https://doi.org/10.1038/nature07894>.
- Rice JW, Olson JK, Colbert JT. University evolution education: the effect of evolution instruction on biology majors' content knowledge, attitude toward evolution, and theistic position. *Evol Educ Outreach*. 2011;4(1):137. <https://doi.org/10.1007/s12052-010-0289-y>.
- Romine WL, Walter EM, Bosse E, Todd AN. Understanding patterns of evolution acceptance—a new implementation of the measure of acceptance of the theory of evolution (MATE) with Midwestern university students. *J Res Sci Teach*. 2017;54(5):642–71. <https://doi.org/10.1002/tea.21380>.
- Romine WL, Todd AN, Walter EM. 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. *Evol Educ Outreach*. 2018;11(1):17. <https://doi.org/10.1186/s12052-018-0093-7>.
- Rutledge ML, Sadler KC. Reliability of the measure of acceptance of the theory of evolution (MATE) instrument with university students. *Am Biol Teach*. 2007;69(6):332–5.

- Rutledge ML, Sadler KC. University students' acceptance of biological theories— is evolution really different? *J Coll Sci Teach*. 2011;41(2):38.
- Rutledge ML, Warden MA. The development and validation of the measure of acceptance of the theory of evolution instrument. *Sch Sci Math*. 1999;99(1):13–8. <https://doi.org/10.1111/j.1949-8594.1999.tb17441.x>.
- Rutledge ML, Warden MA. Evolutionary theory, the nature of science & high school biology teachers: critical relationships. *Am Biol Teach*. 2000;62(1):23–31. <https://doi.org/10.2307/4450822>.
- Sbeglia GC, Nehm RH. Measuring evolution acceptance using the GAENE: Influences of gender, race, degree-plan, and instruction. *Evol Educ Outreach*. 2018;11(1):18. <https://doi.org/10.1186/s12052-018-0091-9>.
- Sbeglia GC, Nehm RH. Do you see what I-SEA? A Rasch analysis of the psychometric properties of the inventory of student evolution acceptance. *Sci Educ*. 2019;103(2):287–316. <https://doi.org/10.1002/sce.21494>.
- Sbeglia GC, Nehm RH. Illuminating the complexities of conflict with evolution: validation of the scales of evolutionary conflict measure (SECM). *Evol Educ Outreach*. 2020;13(1):23. <https://doi.org/10.1186/s12052-020-00137-5>.
- Shtulman A. Qualitative differences between naïve and scientific theories of evolution. *Cogn Psychol*. 2006;52(2):170–94. <https://doi.org/10.1016/j.cogpsych.2005.10.001>.
- Shtulman A, Schulz L. The relation between essentialist beliefs and evolutionary reasoning. *Cogn Sci*. 2008;32(6):1049–62.
- Shtulman A, Valcarcel J. Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*. 2012;124(2):209–15. <https://doi.org/10.1016/j.cognition.2012.04.005>.
- Shtulman A, Calabi P. Cognitive constraints on the understanding and acceptance of evolution. *Evolution Challenges: Integrating Research and Practice in Teaching and Learning about Evolution*, 2012, 47–65.
- Sinatra GM, Southerland SA, McConaughy F, Demastes JW. Intentions and beliefs in students' understanding and acceptance of biological evolution. *J Res Sci Teach*. 2003;40(5):510–28. <https://doi.org/10.1002/tea.10087>.
- Sinatra GM, Kienhues D, Hofer BK. Addressing challenges to public understanding of science: epistemic cognition, motivated reasoning, and conceptual change. *Educ Psychol*. 2014;49(2):123–38.
- Smith MU. Counterpoint: belief, understanding, and the teaching of evolution. *J Res Sci Teach*. 1994;31(5):591–7. <https://doi.org/10.1002/tea.3660310512>.
- Smith MU. Current status of research in teaching and learning evolution: I. Philosophical/epistemological issues. *Sci Educ*. 2009;19(6–8):523–38. <https://doi.org/10.1007/s11191-009-9215-5>.
- Smith MU. Current status of research in teaching and learning evolution: I. Philosophical/epistemological issues. *Sci Educ*. 2010;19(6):523–38. <https://doi.org/10.1007/s11191-009-9215-5>.
- Smith MU, Scharmann LC. Defining versus describing the nature of science: a pragmatic analysis for classroom teachers and science educators. *Sci Educ*. 1999;83(4):493–509. [https://doi.org/10.1002/\(SICI\)1098-237X\(199907\)83:4%3c493::AID-SCE6%3e3.0.CO;2-U](https://doi.org/10.1002/(SICI)1098-237X(199907)83:4%3c493::AID-SCE6%3e3.0.CO;2-U).
- Smith MU, Siegel H. On the relationship between belief and acceptance of evolution as goals of evolution education. *Sci Educ*. 2016;25(5):473–96. <https://doi.org/10.1007/s11191-016-9836-4>.
- Smith MU, Snyder SW, Devereaux RS. The GAENE—Generalized Acceptance of Evolution Evaluation: development of a new measure of evolution acceptance. *J Res Sci Teach*. 2016;53(9):1289–315. <https://doi.org/10.1002/tea.21328>.
- Southerland SA, Sinatra GM. Learning about biological evolution: a special case of intentional conceptual change. Sherry A. Southerland and Gale M. Sinatra. In *Intentional Conceptual Change*. Routledge; 2003.
- Southerland SA, Sinatra GM, Matthews MR. Belief, knowledge, and science education. *Educ Psychol Rev*. 2001;13(4):325–51. <https://doi.org/10.1023/A:1011913813847>.
- Trani R. I won't teach evolution; It's against my religion. And now for the rest of the story. *Am Biol Teacher*. 2004;66(6):419–27. <https://doi.org/10.2307/4451708>.
- Unsworth A, Voas D. Attitudes to evolution among Christians, Muslims and the Non-Religious in Britain: differential effects of religious and educational factors. *Public Underst Sci*. 2018;27(1):76–93. <https://doi.org/10.1177/0963662517735430>.
- Unsworth A, Voas D. The Dawkins effect? Celebrity scientists, (non)religious publics and changed attitudes to evolution. *Public Underst Sci*. 2021;30(4):434–54. <https://doi.org/10.1177/0963662521989513>.
- Verhey SD. The effect of engaging prior learning on student attitudes toward creationism and evolution. *Bioscience*. 2005;55(11):996–1003. [https://doi.org/10.1641/0006-3568\(2005\)055\[0996:TEOEP\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2005)055[0996:TEOEP]2.0.CO;2).
- Weisberg DS, Landrum AR, Metz SE, Weisberg M. No missing link: knowledge predicts acceptance of evolution in the United States. *Bioscience*. 2018;68(3):212–22. <https://doi.org/10.1093/biosci/bix161>.
- Wiles JR. Gifted students' perceptions of their acceptance of evolution, changes in acceptance, and factors involved therein. *Evol Educ Outreach*. 2014;7(1):4. <https://doi.org/10.1186/s12052-014-0004-5>.
- Wiles JR, Alters B. Effects of an educational experience incorporating an inventory of factors potentially influencing student acceptance of biological evolution. *Int J Sci Educ*. 2011;33(18):2559–85. <https://doi.org/10.1080/09500693.2011.565522>.
- Wright BD, Stone MH. *Best test design*. San Diego: Mesa Press; 1979.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.