

RESEARCH

Open Access



Measuring university students' interest in biology: evaluation of an instrument targeting Hidi and Renninger's individual interest

Eva Knekta^{1,2}, Ashley A. Rowland³, Lisa A. Corwin³ and Sarah Eddy^{2*}

Abstract

Background: Boosting students' disciplinary interest has long been considered an important mechanism to increase student success and retention in STEM education. Yet, interest is a complex construct and can mean different things to different people, and many of the existing interest questionnaires do not identify a specific theoretical framework underlying their items. To demonstrate that curricular interventions targeting students' interest are effective, educators need a theoretically based instrument to measure interest. The aim of this study was to develop an instrument measuring undergraduate students' interest in the discipline of biology and collect initial validity evidence supporting the proposed use. The instrument structure is based on Hidi and Renninger's (*Educational Psychologist* 41:111–127, 2006) conceptualization of individual interest, and the intended use is to evaluate changes in the biology interests of the US undergraduate students pursuing STEM degrees. To provide evidence of validity, the instrument was completed by 446 biology majors and 489 non-biology majors at two R1 universities. Exploratory and confirmatory factor analyses were applied to evaluate the internal structure of the instrument.

Results: The final three-factor instrument supported by these analyses includes 6 items representing positive feelings towards biology, 5 items representing personal value of biology, and 8 items representing reengagement in biology-related activities. Measurement invariance across biology and non-biology majors was established and subsequent comparisons of these populations demonstrated that biology majors report significantly higher positive feelings, personal value, and reengagement in biology-related activities compared to non-biology majors.

Conclusions: The study findings support the use of the instrument to gain a broad understanding of students' individual interest in biology. With minor adaptations, the instrument could also be evaluated for use in other STEM disciplines and for use by other populations.

Keywords: Interest, Factor analysis, Validation, Higher education, Measurement invariance, Biology education

* Correspondence: seddy@fiu.edu

²Department of Biological Sciences, Florida International University, Miami, FL 33199, USA

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Retention of people in biology career pathways is an essential component of meeting the growing need for STEM and health care professionals in the USA (Dall, West, Chakrabarti, Reynolds, & Iacobucci, 2018; PCAST, 2012). At the undergraduate level, retention efforts often focus on improving the experiences of biology students in college. For example, many colleges and universities are expanding the use of evidence-based teaching practices in biology classrooms (Stains et al., 2018), providing more students with early research experiences (Rodenbusch, Hernandez, Simmons, & Dolan, 2016), and/or building learning communities for their students to participate in (Zhao & Kuh, 2004). These successful programs provide evidence that contexts can influence student behaviors in ways that increases retention.

In addition to context, a focus on understanding students' attitudes and how it relates to context is important to increase retention. Students frequently report leaving a STEM field due to a loss of interest in their discipline (Barr, Gonzalez, & Wanat, 2008; Fouad, Chang, Wan, & Singh, 2017; Seymour & Hunter, 2019). Thus, students' interest, and how it grows and wanes, is one important attitude to understand. Interest may predict persistence because it is tied to motivation, which directly influences behavior (Glynn, Bryan, Brickman, & Armstrong, 2015; Renninger & Hidi, 2016; Ryan & Deci, 2000; Wentzel & Miele, 2016).

In fact, interest is an integral component of several prominent theories of motivation, such as expectancy-value theory, self-determination theory, and social cognitive career theory (Lent, Brown, & Hackett, 2002; Ryan & Deci, 2000; Wigfield & Eccles, 2000). The centrality of interest for academic achievement has also been empirically established. For example, a meta-analysis of more than 150 papers published between 1965 and 1992 found that interest was correlated with measures of achievement across a range of academic contexts such as standardized knowledge tests and course and exam grades (Schiefele, Krapp, & Winteler, 1992). Interest has also been shown to promote attention and recall (Hidi, 1990; McDaniel, Waddill, Finstad, & Bourg, 2000), positive affect and persistence at tasks (Ainley, Hidi, & Berndorff, 2002), self-efficacy (Bong, Lee, & Woo, 2015), goal setting (Durik & Harackiewicz, 2003; Harackiewicz, Barron, Tauer, Carter, & Elliot, 2000), and the use of specific learning strategies (Alexander & Murphy, 1998; for review see Renninger & Hidi, 2016). Finally, Harackiewicz and Hulleman (2010) argue that interest is critically important in its own right and should be considered essential with respect to happiness and life satisfaction (Harackiewicz & Hulleman, 2010).

Thus, interest has clear potential for far-reaching benefits to achievement and persistence within the

undergraduate biology setting. An essential prerequisite for doing research on disciplinary interest, or evaluating the efficacy of interventions targeting biology interest, is to have a tool to measure that interest.

The aim of this study was to develop an instrument measuring individual interest in one STEM field, biology, and collect initial validity evidence supporting the use of the instrument with undergraduate biology majors. The instrument is intended to be used to evaluate changes in individual interest of US undergraduate biology students progressing through their curriculum. The instrument development presented here builds on and further develops two existing questionnaires, the *Study Interest Questionnaire* (SIQ) presented by Schiefele, Krapp, Wild, and Winteler (1993) and items concerning interest included in the student background questionnaire administered in connection to the Program for International Students Assessment 2015 (OECD, 2017). Factor analysis was applied to evaluate the internal structure of the instrument, and measurement invariance was examined to evaluate whether the instrument functions similarly across biology majors and non-biology majors. Finally, differences in individual interest in biology between biology majors and non-biology majors were assessed to evaluate evidence of external validity.

Theoretical framework

When measuring a psychological construct, an essential prerequisite is to present a definition and theoretical basis of the construct in order to clearly convey its intended meaning. This is particularly important for interest because it has varied colloquial uses, definitions, and theoretical frameworks (Krapp, 2002; Renninger & Hidi, 2016; Ryan & Deci, 2000; Schiefele, 1991; Wigfield & Eccles, 2000). This is especially true in biology contexts (Rowland, Knekta, Eddy, & Corwin, 2019). Various theories conceptualize interest in slightly different ways that will result in different measures of interest being used, different outcomes being obtained, and different interpretations being made. In some theories, interest is one component of a larger construct, such as motivation (Ryan & Deci, 2000; Wigfield & Eccles, 2000). In those broader frameworks, interest is usually defined as a unidimensional construct concerning positive feelings (Ryan & Deci, 2000; Wigfield & Eccles, 2000). Other theories consider interest to be a multidimensional construct with affective (e.g., liking), cognitive (e.g., assigning value or storing knowledge), and behavioral (e.g., reengaging with specific content) components (Krapp, 2002; Renninger & Hidi, 2016; Schiefele & Csikszentmihalyi, 1994).

Because we are working with students who are advancing their disciplinary knowledge, we use one of the few interest frameworks that takes a developmental approach to theoretically understand and define interest: the four-phase model of interest development (Hidi & Renninger, 2006; Renninger & Hidi, 2016). The four-phase model of interest development is one of the most cited theories within the biology education community when examining interest (Rowland et al., 2019). This model describes how an initial interest triggered by the environment may develop into a more internal and stable form of interest, termed *individual interest* (Hidi & Renninger, 2006; Renninger & Hidi, 2016).

The first two phases in the model are *triggered situational interest* (phase 1) and *maintained situational interest* (phase 2; Hidi & Renninger, 2006; Renninger & Hidi, 2016). Situational interest arises during an individual's reaction to an experience with specific content or an activity involving that content. In these phases, the interest is elicited by something external to the individual. Situational interest is a psychological state that has an affective component (usually positive feelings but could also involve negative feelings such as fear or disgust; Hidi & Harackiewicz, 2000). It is characterized by focused attention. The two stages of situational interest primarily differ in the longevity of interest beyond the experience. As an example, consider a student who encounters a colorful insect. If they experience triggered situational interest, they may be fleetingly interested in the insect, but they do not pursue the insect when it flies away, nor do they attempt to identify the insect after the encounter. With a maintained situational interest, this student engages longer; they may follow the insect and attempt to identify it.

The later phases of the four-phase model of interest are *emerging individual interest* (phase 3) and *well-developed individual interest* (phase 4). Individual interest is characterized as a "psychological state and a relatively enduring predisposition to reengage with a particular class of content over time" (Renninger & Hidi, 2016, pp. 13). Reengaging with content is a measure that can look like repeatedly choosing to involve oneself with activities relating to a specific class of content. For example, one could repeatedly try to find and identify insects. A person with emerging individual interest is *likely* to independently reengage with content while a person with well-developed individual interest *will* independently reengage with content (Renninger & Hidi, 2016). Hidi, Renninger, and colleagues have described a number of additional characteristics typical for a learner with an individual interest in certain content including positive feelings, stored knowledge, independent reflection, recognition of others' contributions, and personal value associated with the content (Hidi & Renninger, 2006;

Renninger, Ewen, & Lasher, 2002; Renninger & Hidi, 2016). A more developed individual interest is characterized by more positive feelings, higher perceived value for biology, and increased intention to engage in biology-related activities (Renninger & Hidi, 2016). It develops in an iterative manner: increasing knowledge contributes to deepened feelings and increased value for the content, which then spurs continued engagement and additional knowledge acquisition and so on (Renninger & Bachrach, 2015; Renninger & Hidi, 2016). In the example above, a student who has individual interest knows that the insect is a beetle (order Coleoptera), and they value having knowledge about insects. With emerging individual interest, the student may not only seek to identify the insect, but to also understand more about its biology by reading about it after the encounter. The student may acquire a well-developed interest by pursuing a graduate degree in entomology.

It should be noted that positive feelings and value, as described by the four-phase model of interest, are intrinsic in nature (Schiefele, 2009). Thus, the feelings and perceived value are directed towards a certain object or domain and are not based on the relationships of that object to other objects or domains. For example, a person that values biology because competence in biology will help them get a prestigious job, experiences an extrinsic utility value not an intrinsic value of biology. This would not count as value in Hidi and Renninger's conceptualization of interest. Alternatively, valuing biology because having biology knowledge is central to how you see yourself as a person or because studying biology provides a satisfying challenge would count as value. For example, getting to know a new bird species would be valued for a birder.

Existing measures of interest

Interest has been measured in a number of different ways, including observations, neuroscientific techniques, facial expressions, class enrollment data, and questionnaires (Renninger & Hidi, 2016). Each of these methods has its own unique benefits and limitations (Cohen, Manion, & Morrison, 2011). Questionnaires, like the one presented in this study, are suitable for data collection when researchers aim to conduct a longitudinal study of students over time or collect data from a large number of students (Cohen et al., 2011). Since the design of a questionnaire has to be tailored to address the research questions asked, the theory and definitions drawn from, and the context of the study, existing questionnaires concerning interest vary greatly. Some of the first questionnaires of interest were inventories of topics or activities (Renninger & Hidi, 2016). These inventories, for example the ROSE questionnaire described by Schreiner and Sjöberg (2004), measure students' interest in a large

range of biology content or activities (e.g., “How interested are you in learning about the following...”). The inventories consider numerous different topics and or activities and thus differentiate different aspects of one larger domain (e.g., biology), but lack differentiation with respect to the construct of interest (i.e., it is treated as a unidimensional construct). Therefore, they are less suitable to estimate the development of biology interest.

A second method used to measure interest is to use a single item to probe students’ interest (e.g., “To carry out experiments with plants is interesting for me”; Holstermann, Grube, & Bögeholz, 2010). Since interest cannot be directly observed, it is not preferable to use a single item to make inferences about students’ interest (Knekta, Runyon, & Eddy, 2019). Instead, using students’ scores from several items measuring interest in slightly different ways and combining them into a sum or mean score is best practice in measuring unobservable psychological constructs (Knekta et al., 2019).

Wigfield and Eccles (2000), McAuley, Duncan, and Tammen (1987), and Pintrich and de Groot (1990) have published questionnaires with multiple items representing interest. In these questionnaires, interest is a unidimensional construct that contributes to measurement of a multidimensional construct, motivation. In conveying interest as unidimensional, these scales only capture one aspect of interest, such as positive feelings. Yet, to measure how interest develops over time, Renninger and Hidi (2016) recommend measuring positive feelings towards a specific object *and* additional aspects of interest such as students’ engagement with the content over time and stored value. Unidimensional scales of interest cannot capture this level of detail.

We have found only a few questionnaires capturing several dimensions of interest (e.g., Linnenbrink-Garcia et al., 2010; OECD, 2017; Schiefele et al., 1993). The instrument described by Linnenbrink-Garcia et al. (2010) focuses on aspects of situational interest in psychology and mathematics. The *Study Interest Questionnaire* (SIQ) described by Schiefele et al. targets individual interest in one’s university subject and includes items targeting feelings-related valances, value-related valances, and intrinsic character of valance beliefs (Schiefele, 2009; Schiefele et al., 1993). When Renninger & Hidi (2016) refer to Schiefele’s work, they conclude that “Findings from his and his colleagues’ work with the Students Interest Questionnaire suggest that value, feelings, and the choice to engage with content are all associated and are not independent factors” (p. 23). This indicates that although Schiefele et al. (1993) named their dimension differently, Hidi and Renninger consider the SIQ to represent the dimensions value, feelings, and reengagement as described by their four-phase model of interest development. The student background

questionnaire administrated in connection to the Program for International Students Assessment (PISA) in 2015, included five items measuring independent and voluntary engagement in activities related to biology and five items concerning positive emotions towards biology (OECD, 2017).

In this paper, we combine items described by Schiefele et al. (1993), OECD (2017), and newly written items into an instrument measuring positive feelings, personal value, and independent, voluntary reengagement in biology-related content and activities. Starting with existing questionnaires allowed us to more easily target measurement of interest development as described by Hidi and Renninger. Notably, both the SIQ and the instrument described by OECD represent multidimensional measure of individual interest. Further, some validity evidence for these instruments has already been provided. The SIQ instrument includes 18 items representing individual interest in one’s university subject and is adapted for university students. However, the development of and validation studies on the SIQ instrument were made more than two decades ago on a German version of the questionnaire (Schiefele, 2009; Schiefele et al., 1993). The instrument presented by OECD is newly developed but does not include a value component and is adapted for use for high school students. Schiefele (2009) argue that more research on the development of multidimensional interest questionnaires is needed to better understand how individual interest develops.

Methods

Instrument

Our instrument (which we call the Biology Interest Questionnaire, BIQ) aims to measure individual interest—as defined by Renninger & Hidi (2016)—in biology of undergraduates in the USA. The intended use of the instrument is to document changes in biology interest as biology majors progress through their curriculum. The instrument is primarily intended for use with biology majors, but its functionality for other groups of university students was also assessed for comparative purposes. The instrument is a self-report survey and concerns the fairly general domain biology (as opposed to, for example, interest in a specific biology activity or topic area). To evaluate a student’s biology interest the questionnaire includes three different aspects of interest: students’ (1) positive feelings towards biology, (2) perceived personal value of biology, and (3) independent and voluntary reengagement in biology-related content and activities. These three aspects are described as important indicators for understanding individual interest and its development (e.g., Hidi & Renninger, 2006; Linnenbrink-Garcia et al., 2010; Renninger et al., 2002; Renninger & Hidi, 2016). Inspired

by Schiefele (2009) and Linnenbrink-Garcia et al. (2010), we define positive feelings as *individuals' affective experiences while engaging with biology content (e.g., enjoyment, excitement)*. Perceived value refers to the *personal significance of biology (e.g., self-realization, centrality with one's self concept)*. Both positive feelings and value are considered direct measures of the psychological aspects of individual interest. A person that independently reengages in biology-related content and activities is defined as *a person that wants to reengage in biology-related content and does so without needing input from others*. Independent and voluntary reengagement is considered good indicator for development of individual interest and is considered a primary behavioral outcome of a learner with an individual interest in biology (Renninger & Hidi, 2016). Because development and administration of an instrument are a balance between coverage and length, we limited our instrument to measure only the three aspects described since they were suitable to measure individual interest by self-report and to assess in a general content. We did not assess stored knowledge, independent reflection, and recognition of others' contributions, which are three additional characteristics of individual interest described by Renninger and Hidi (2016) that would have made our instrument longer and more difficult to facilitate.

The BIQ was assembled by writing 11 new items, adapting 15 items from the Students Interest Questionnaire (SIQ, Schiefele et al., 1993), and adapting 7 items from the PISA Background Questionnaire (OECD, 2017). The initial instrument included 11 items representing students' positive feelings towards biology, 9 items representing the perceived personal value of biology, and 13 items representing independent and voluntary reengagement in biology-related content and activities (Table 1). Because this instrument is intended for use with a population that likely has an existing interest in biology, we used a 6-point Likert-type scale with a positively packed response scale¹ to rate the feelings- and value-related items (Brown, 2004; Brown, Harris, O'Quin, & Lane, 2017: (1) Strongly disagree, (2) Moderately disagree, (3) Slightly agree, (4) Moderately agree, (5) Mostly agree, and (6) Strongly agree). We predicted that four positive response options rather than the traditional three would help reduce the ceiling effect and result in a higher variation of responses. Overall, we hoped the positively packed scale would make it easier to detect differences in interest between and within students. We did not anticipate that the reengagement items would be equally as easy to endorse as the feeling and value items, and therefore, we chose a balanced 6-

point Likert-type scale for these items ((1) Strongly disagree, (2) Disagree, (3) Slightly disagree, (4) Slightly agree, (5) Agree, and (6) Strongly Agree). We also provided students the option to choose "prefer not to respond" on all items.

Participants and procedures

The questionnaire was distributed in spring 2018 to undergraduates taking introductory biology courses at a large southern US R1 university and a large western US R1 university. Students took the questionnaire 4 weeks before the end of the spring semester. In total, 444 biology majors and 489 non-biology majors completed the questionnaire. Demographic information for the entering cohort of biology majors at the two universities is presented in Table 2.

Data analysis

All data analyses were run R version 3.5.1 (R Core Team, 2018). First, descriptive statistics and correlations between items were examined. In order to investigate how many subscales the instrument represented (collecting validity evidence for the internal structure), factor analysis was applied. Because the instrument was newly developed, we ran an initial exploratory factor analysis (EFA) on half our sample ($n = 222$) to determine the dimensionality of the questionnaire and detect problematic items. This was followed by a confirmatory factor analysis (CFA) ($n = 222$) to confirm the result gained from the EFA. To investigate whether the instrument functions equally for both biology majors and non-majors, measurement invariance was examined (Bashkov & Finney, 2013; Vandenberg & Lance, 2000). EFA was run using the R package psych (Revelle, 2017). CFA and measurement invariance were run using the R package lavaan (Rosseel, 2012).

Exploratory factor analysis

A weighted least square (WLS) estimator was used to extract the variances from the data. Since we hypothesized a correlation between the subscales within the instrument (i.e., feelings, value, and reengagement), an oblique rotation (oblimin rotation) was chosen. Visual inspection of the scree plot, parallel analysis based on eigenvalues from principal components, and factor analysis, as well as theoretical considerations, was used to identify the number of factors² to retain (psych package, Revelle, 2017). Total variance explained, communalities, pattern coefficients, and factor correlations were used to evaluate the fit of the data to the model as well as the fit

¹Positively packed response scale (sometimes also referred to as an unbalanced scale) describes a scale that has more positive response options than negative.

²When performing factor analysis, the factors represent subscales in an instrument.

Table 1 Items included in the original scale. Mean values and standard deviations for the biology major group are given. Original source for the included items is indicated in the last column

# ^a		Mean	Std	Source ^b
These questions ask about your feelings toward biology-related content. Please rate your agreement with the following statements.				
F1	Working with the subject matter and problems of biology is among my favorite activities.	4.23	1.50	Sa
F2	After a long weekend or vacation I look forward to getting back to my biology classes.	3.35	1.52	So
F3	Being involved in biology classes puts me in a good mood.	3.90	1.50	So
F4	I choose to study biology primarily because of the interesting subject matter involved.	4.37	1.52	So
F5	I generally have fun when I am learning biology topics.	4.27	1.44	Po
F6	I like reading about biology.	4.05	1.51	Po
F7	I am happy working on biology topics.	4.32	1.44	Po
F8	I enjoy acquiring new knowledge in biology.	4.76	1.36	Po
F9	I am interested in learning about biology.	4.76	1.39	Po
F10	I think the field of biology is very interesting.	4.86	1.41	D
F11	Biology fascinates me.	4.81	1.40	D
These questions ask about the value that biology contents hold for you. Please rate your agreement with the following statements.				
V1	It is of great personal importance to me to become knowledgeable in biology.	4.69	1.44	D
V2	I really see value in the things that I am learning in biology.	4.59	1.43	D
V3	Studying biology has a lot to do with whom I want to become as a person.	4.48	1.47	D
V4	Compared to other things that are of importance to me (e.g., hobbies, social life), learning about biology is of more importance to me.	3.50	1.56	Sa
V5	Learning about biology is more important to me than leisure and amusement.	3.26	1.60	Sa
V6	Learning about biology has always been important to me.	4.07	1.50	Sa
V7	I am certain that studying biology has a positive influence on my personality.	4.12	1.57	Sa
V8	I am confident that learning about biology directly corresponds to my personal preferences.	4.18	1.50	So
V9	I value the knowledge that I have about biology.	4.76	1.31	D
These questions ask about your involvement in biology-related activities. Please rate your agreement with the following statements. Please note: the response options have changed!				
R1	I often talk outside class about what I am learning in biology classes.	4.16	1.36	Sa
R2	I talk about my hobbies rather than about things that I learned in biology.	2.69	1.20	Sa
R3	In my free time, and unrelated to my coursework, I read magazines, articles, or books related to topics of biology.	3.44	1.46	Pa
R4	If I had enough time, I would work more intensively with certain aspects of biology, even if they had nothing to do with any course requirements.	4.11	1.41	So
R5	Even before coming to college, I voluntarily spent time thinking about topics in biology.	4.02	1.51	Sa
R6	When I am in a library or bookstore, I like to browse through magazines or books related to topics of biology.	3.59	1.50	So
R7	I spend time out in the nature as often as I can.	4.59	1.31	D
R8	If I am watching TV or surfing the internet, I get hooked on biology-related programs or sites.	4.02	1.48	D
R9	I often voluntarily have conversations with others about topics related to biology.	4.03	1.41	Sa
R10	I am engaged in a biology related club.	2.64	1.53	Pa
R11	In my free time, and unrelated to my coursework, I follow news stories related to biology via digital media (e.g., podcasts, online videos, blogs, ebooks, twitter feed).	3.71	1.49	D
R12	I visit biology-related exhibits (e.g., zoos, museums) as often as I can.	3.90	1.56	D
R13	In my free time, and unrelated to my coursework, I attend biology related seminars and presentations.	2.88	1.46	D

a. *F* feelings-related items, *V* value-related items, *R* reengagement-related items, b. *S* item sampled from Schiefele et al. (1993), *P* Item sampled from the PISA survey (OECD, 2017), *D* item developed by the authors, *o* original item used (adaption such as changing "my studies" to "my biology classes" is considered as original item used), *a* items adapted (e.g., R10 was changed from "Attend a biology club" to "I am engaged in a biology related club")

Table 2 Demographic information about the entering cohort biology majors at the two universities

	%	
	University 1	University 2
African American	10	1
American Indian/Alaskan Native	< 1	1
Asian or Pacific islander	5	7
Hispanic	75	12
Non-resident alien	3	-
Unknown	< 1	6
White	7	73
Female	67	56
First generation student	25	12

of individual items to the scales. The size of the pattern coefficient for each item on the theorized factor (the focal factor), on any other subscales, and the difference between these two were considered to determine if the items fit the model. A pattern coefficient > 0.40 on the theorized factor was considered sufficient for retention of the item, and a pattern coefficient > 0.30 on any other subscales was considered problematic. We chose these very inclusive cut off values since the instrument is newly developed, and we did not want to exclude items too early in the process. The total sample size of biology majors for the EFA was 222, which could be considered small but is sufficient for performing factor analysis if the number of items per factor as well as item correlations are high (Gagne & Hancock, 2006; Wolf, Harrington, Clark, & Miller, 2013). Our questionnaire had nine or more items on each subscale, and most items correlations were above 0.5 which both can be considered high values (see correlation matrix in [Supplement material](#); Gagne & Hancock, 2006; Wolf et al., 2013).

Confirmatory factor analysis

To confirm the results from the EFA, CFA was applied on the second half of the biology major sample ($n = 222$). A robust maximum likelihood estimator (MLR) was used to extract the variances from the data. Multiple fit indices (chi-squared value from robust maximum likelihood estimation, MLR χ^2 ; comparative fit index, CFI; the root mean squared error of approximation, RMSEA; and the standardized root-mean-squared residual, SRMR) were consulted to evaluate model fit. The fit indices were chosen to represent an absolute, a parsimony-adjusted, and an incremental fit index (Bandalos & Finney, 2010). Consistent with the recommendations by Hu and Bentler (1999), the following criteria were used to evaluate the adequacy of the models: CFI > 0.95 , SRMR < 0.08 , and RMSEA < 0.06 . Coefficient ω was computed based on the model results

and used to assess reliability (Gignac, 2009). Coefficient ω values > 0.70 were considered acceptable.

Measurement invariance

To investigate whether the interest instrument functioned equally for biology majors and non-biology majors, measurement invariance was examined (Bashkov & Finney, 2013; Vandenberg & Lance, 2000). If the instrument is invariant between the groups of interest, it means that the researcher could confidently use the instrument to compare the two groups with respect to the latent score achieved from the instrument. In a stepwise manner, invariance of different parameters was tested by constraining them to be equal across biology majors and non-biology majors. First, we examined whether the factor structure was invariant across biology majors and non-biology majors (configural invariance). Factor structure refers to the number of factors in the instrument as well as which items are represented on each factor. Secondly, metric invariance, whether the factor loadings of the items were equal across the two groups of students, was tested. Equal factor loading across groups means that each item contributes the same amount to the factors for both groups. Finally, scalar invariance was assessed by constraining the factor loadings and intercepts to be equal across the two groups. Establishing scalar invariance means that students who have the same score for the factor report the same values on the individual items making up that construct. Finally, if at least partial scalar invariance was achieved, differences in latent means for the different subscales between the non-biology and biology majors were tested. This was done by combining the two groups in a multigroup model and constraining the latent means for the non-biology majors to be zero while allowing the latent means for the biology majors to be freely estimated. This test reveals whether there are differences in student responses on interest subscales between biology majors and non-biology majors. In addition to the above described measurement invariance test, covariances between the latent factors were also tested for invariance by constraining the correlations between the latent factors to be equal. Invariance between covariances is not needed in order to use the instrument for comparing the groups. However, we wanted to test whether the relationships between the different latent factors (feelings, value, and reengagement) were equal for biology and non-biology majors in order to a better theoretical understanding of the interest construct. A $\Delta CFI < .01$ between the configural and metric, the metric and the covariance, and between the metric and scalar model, was considered indicative of invariance (Cheung & Rensvold, 2002).

Results

Descriptive statistics

For the total sample ($n = 444$), the 33 items had between 1 and 19 missing values per item (0.002–4% missing values). Multiple imputation using logistic regression was used to estimate missing values (implemented with the MICE package, van Buuren & Groothuis-Oudshoorn, 2011). The first of five sets of imputed datasets was used for the results reported here. The same analysis was later run using the four other imputations, and no substantial differences in results were found. Analysis of Mahalanobis distance³ was run revealing 23 cases with high Mahalanobis distance ($p < .001$) that could be multivariate outliers. Each of the 23 cases was inspected in detail, and we found no justification for removing any of them.

The items had a mean between 2.7 and 4.9, a univariate skewness $< |1.1|$ and kurtosis $< |1.1|$, and the standard deviations ranged between 1.3 and 1.6 (Table 1). Mardia’s multivariate normality test (implemented with the psych package, Revelle, 2017) showed significant multivariate skewness and kurtosis values which indicates multivariate non-normality. The Kaiser’s measure of sampling adequacy value was .97 indicating good factor ability. Multicollinearity was investigated by examining inter-item correlations and tolerance values from multiple regressions (implemented with the olsrr package, Hebbali, 2018). The highest correlation between items was 0.87, and the lowest tolerance value was 0.15. These values indicate rather high correlations between items but are still not multicollinearity (correlations > 0.90 , tolerance < 0.1 indicates multicollinearity; Tabachnick & Fidell, 2013).

The correlations matrix indicated reasonable correlations between the feelings-related items (r between .49 and .87) and between the value-related items (r between .43 and .76; Supplement material Tables 1 and 2). A few low correlation ($r < .3$) between the reengagement-related items were found (Supplement material Table 3), and some items showed high correlations to items on other subscales.

Exploratory factor analysis

Results from the scree plot and the parallel analysis were analyzed to determine the number of factors to retain. Parallel analysis based on all items indicated that between 2 and 4 factors could be relevant to retain. The scree plot leveled out at 4 factors indicating that no more than 4 factors should be retained. Consequently, EFA analyses with 2, 3, and 4 factors including all items were tested. The

total variance explained by the 2, 3, and 4 factor solutions was similar (57%, 60%, and 62%, respectively).

First round of EFAs

Interpretation of the first round of EFAs and the correlation matrix indicated that most of the feelings-related

Table 3 Pattern matrix from the 3-factor solution for first round of EFA with all original items. Bold font indicates items with low pattern coefficients on the focal factor or high pattern coefficients on a second factor

Item	WLS1	WLS2	WLS3	h2	com
F1	0.73			0.75	1.1
F2	0.40		0.36	0.59	2.3
F3	0.58		0.25	0.69	1.5
F4	0.69			0.69	1.1
F5	0.69			0.69	1.1
F6	0.53	0.20		0.60	1.5
F7	0.71			0.78	1.2
F8	0.94			0.79	1.0
F9	0.96			0.86	1.0
F10	0.98			0.81	1.0
F11	0.89			0.73	1.0
V1	0.77			0.57	1.0
V2	0.71			0.65	1.1
V3	0.61			0.56	1.2
V4	0.26		0.70	0.69	1.3
V5		0.20	0.69	0.68	1.2
V6	0.37		0.45	0.68	2.2
V7	0.23	0.25	0.47	0.64	2.0
V8	0.47	0.27	0.22	0.64	2.1
V9	0.70			0.63	1.1
R1	0.27	0.38		0.43	2.0
R2	- 0.26		0.47	0.20	1.6
R3		0.83		0.63	1.0
R4	0.22	0.67		0.57	1.3
R5		0.61		0.51	1.1
R6		0.77		0.62	1.0
R7	0.20	0.37	- 0.33	0.17	2.5
R8		0.73		0.55	1.1
R9		0.72		0.63	1.0
R10	- 0.25	0.49	0.21	0.29	1.9
R11		0.76		0.63	1.0
R12		0.59		0.45	1.0
R13	- 0.20	0.67		0.40	1.2

Pattern coefficients < 0.2 are not showed for clarity
WLS 1/2/3 pattern coefficients for the first/second/third factor retained from the EFA using the WLS estimator, h2 measure of communality, com Hoffman’s index of complexity (the average number of factors necessary to explain the item)

³Mahalanobis distance measures the distance between an individual case and the distribution of all the cases. It can be used to calculate the probability that an individual case within a sample population is an outlier via the combination of multiple variables scores.

items had relatively high correlations to each other and had high pattern coefficients on only one factor for the 2-, 3-, and 4-factor solutions (see pattern matrix from 3-factor solution in Table 3). Thus, they seemed to belong to a single factor that could be considered to represent the feeling aspect of interest. However, a few feelings items (F2, F3, F6, F10) had pattern coefficients above 0.3 on another factor indicating they might not only measure feelings but also something else. Items F8, F9, F10, and F11 all had very high pattern coefficient (> 0.9) which indicates that they might be too similar.

The value-related items did not show a consistent pattern across all three solutions. In the four-factor solution, value items had relatively high pattern coefficients on either the third or the fourth factor. However, several value-related items also had pattern coefficient > 0.3 on the first or second factor indicating cross-loading with the feelings and reengagement scales. In the three-factor solution, some of the value-related items had high pattern coefficients on the same factor as the feelings-related items while other value-related items had high pattern coefficients on the third factor (Table 3). For the two-factor solution, the value-related items split across the first (feelings) and second (reengagement) factors.

Similar to the feelings-related items, most reengagement items had relatively high correlations to each other and had high pattern coefficients on only one factor for the 2-, 3-, and 4-factor solutions. Thus, they seemed to represent the reengagement factor of interest. However, items R1, R2, B7, and R10 had overall low correlation to other reengagement items, relatively low pattern coefficients and or pattern coefficients > 0.3 on a second factor.

Second round of EFAs

In a second round of EFAs, the most problematic items were removed in a stepwise manner. We started with removing R2, R7, and R10 which showed low pattern coefficients, low communalities, and low correlations to other items in the reengagement scale (Table 3 and Supplemental Material Table 3). We also removed R1 and F2 as they had only small differences between the pattern coefficients on the focal vs. other factors. Parallel analysis based on the new set of items indicated that 2 or 3 factors should be retained. When looking at the new 2- and 3-factor solution, the following interpretations were made. Feelings-related items F7, F8, F9, F10, and F11 all had similar wording, correlated very strongly with each other, and had high pattern coefficients (> .89) in all the EFA solutions. Thus, they had a strong influence on the feelings scale. In order to obtain a more parsimonious scale that has good theoretical coverage, rather than being dominated by one element of the feelings-related aspect of interest, only one item from

this set was retained. Items F1, F3, F4, F5, F6, and F7 were deemed as most suitable to keep, and items F8, F9, F10, and F11 were removed.

Across the EFA solutions, the value-related items did not form a discrete single factor. Items V1, V2, V3, V8, and V9 correlated strongly with the feelings-related items (Supplemental Material, Table 2). Items V5, V7, and V8 correlated strongly with the reengagement-related items. In order to obtain EFA solutions with pattern coefficients > 0.4 on the focal factor and low pattern coefficient on the other factors (< 0.3), we could either keep the value-related items V4, V5, and V7 or items V1, V2, V3, V8, and V9. We decided to retain the second set of value-related items as three items for a sub-scale is not optimal. Further, the second set of items better covered our definition of the value aspect of interest.

Third round of EFAs

Consequently, the third round of EFAs was based on items F1, F3, F4, F5, F6, F7, V1, V2, V3, V8, V9, R3, R4, R5, R6, R8, R9, R11, R12, and R13. Parallel analysis

Table 4 Pattern matrix from 2- and 3-factor solution from the third round of EFAs

Item	2-factor solution				3-factor solution				
	WLS1	WLS2	h2	com	WLS2	WLS1	WLS3	h2	com
F1	0.84		0.73	1.0	0.53	0.35	0.73	1.7	
F3	0.74		0.69	1.1	0.50	0.27	0.69	1.7	
F4	0.82		0.70	1.0	0.71		0.73	1.1	
F5	0.79		0.68	1.0	0.95		0.80	1.0	
F6	0.67		0.61	1.1	0.78		0.68	1.0	
F7	0.81		0.77	1.0	0.89		0.86	1.0	
V1	0.88		0.59	1.1			0.86	0.69	1.0
V2	0.88		0.67	1.0			0.79	0.75	1.0
V3	0.76		0.57	1.0			0.79	0.67	1.0
V8	0.60	0.25	0.63	1.3	0.32		0.44	0.64	2.1
V9	0.75		0.60	1.0			0.58	0.63	1.2
R3		0.84	0.64	1.0	0.84			0.64	1.0
R4		0.57	0.55	1.3	0.55	0.26		0.55	1.4
R5		0.65	0.54	1.1	0.67			0.54	1.0
R6		0.79	0.64	1.0	0.77			0.64	1.1
R8		0.64	0.54	1.1	0.66			0.54	1.0
R9		0.76	0.63	1.0	0.76			0.63	1.0
R11		0.77	0.63	1.0	0.80			0.65	1.1
R12		0.61	0.44	1.0	0.61			0.44	1.0
R13		0.70	0.36	1.1	0.68			0.35	1.1

Pattern coefficients < 0.2 are not showed for clarity
WLS 1/2/3 pattern coefficients for the first/second/third factor retained from the EFA, h2 measure of communality, com Hoffman's index of complexity (the average number of factors necessary to explain the item)

indicated that 2 or 3 factors should be kept. A 2-factor solution represented combined feelings and value scale and a separate reengagement scale, with all items having pattern coefficients > 0.4 on the focal factor and pattern coefficients < 0.3 on the other factors (Table 4). Total variance explained was 61%. Thus, although a number of items were removed, this two-factor model explained more variance in the data than the first two-factor model including all original items. In the 3-factor solution, all items had pattern coefficients > 0.4 on the focal factor, and most items had pattern coefficient < 0.3 on the other factors. However, 2 items had pattern coefficients > .3 on the other factors (Table 4). Thus, it continued to be difficult to clearly separate the value-related aspects of interest from the feelings and reengagement aspects. Total variance explained was 64%, slightly more than the 2-factor model.

In summary, the EFAs indicated two models to test with a CFA. A two and a three-factor solution with feelings-related items F1, F3, F4, F5, F6, F7, value-related items V1, V2, V3, V8, V9, and reengagement-related items R3, R4, R5, R6, R8, R9, R11, R12, and R13. At this stage, the two-factor solution showed slightly better psychometric properties; however, the three-factor solution was not precluded.

Confirmatory factor analysis

In order to confirm the factor structure suggested by the EFAs, two CFA models were specified. First, a two-factor CFA model was tested with items F1, F3, F4, F5, F6, F7, V1, V2, V3, V8, and V9 representing a combined feelings and value interest factor and items R3, R4, R5, R6, R8, R9, R11, R12, and R13 representing reengagement factor. The second CFA used a three-factor model where the feelings- and value-related items were separated into distinct factors (Fig. 1). Correlation between the factors was allowed. For identification purposes,

factor loadings for one item on each factor were set to 1. Although items F1 and V8 did not showed optimal psychometric properties for the 3-factor EFA solution, we decided to keep these items so as to retain content coverage until the questionnaire is validated in more samples. If they continue to show poor properties in future analyses, removal or rewording should be considered.

The specified two-factor CFA demonstrated model fit close to our chosen guidelines ($\chi^2 = 367$, $df = 169$, $p < .00$, CFI = 0.94, RMSEA = 0.079, and SRMR = 0.045). Factor loadings were above 0.70 for all items, meaning that for most items around 50% of the variance in the items were explained by the theorized factor. This means that the factors explained the variance in most of the items well. Factor correlations between the two factors were 0.73.

The specified three-factor model also showed good model fit ($\chi^2 = 259$, $df = 167$, $p < .00$, CFI = 0.97, RMSEA = 0.054, and SRMR = 0.042), and all factor loadings were above 0.70 (Fig. 1). The correlation between the feelings factor and value factor was high (0.88). The lowest correlation was found between the value factor and the reengagement factor (0.68). The correlation between the feelings factor and reengagement factor was 0.72.

Because the 3-factor model had slightly higher CFI and lower RMSEA and SRMR and allowed a separation of the feelings and value aspects of interest, we chose to continue with the 3-factor model as the best model for the instrument. The calculated coefficient omegas for the subscales were 0.95, 0.92, and 0.93 for feelings, value, and reengagement, respectively.

Measurements invariance

Results from the stepwise investigation of measurement invariance are presented in Table 5. The configural

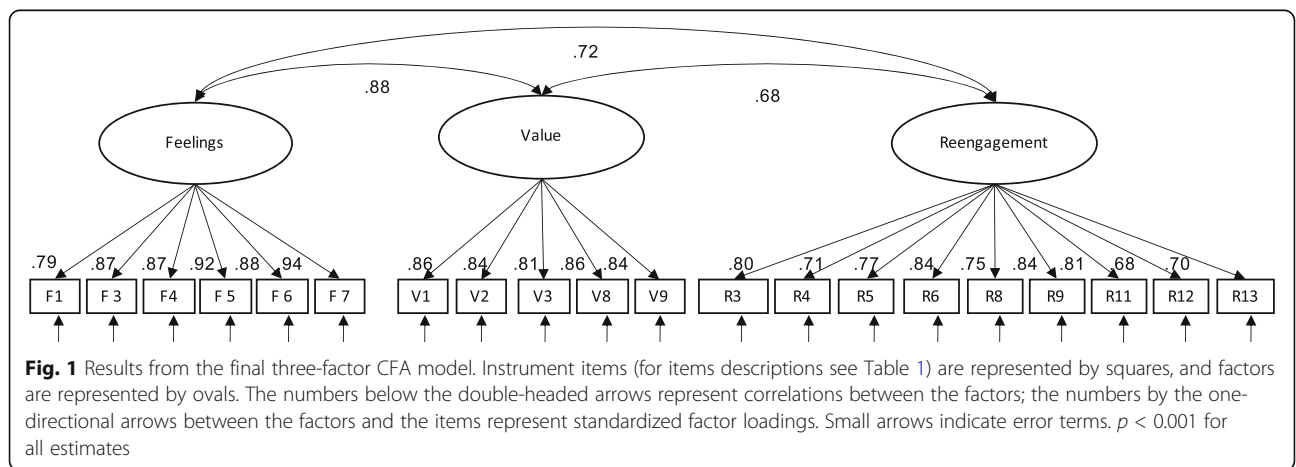


Table 5 Test of measurement invariance of the interest instrument across biology majors ($n = 222$) and non-biology majors ($n = 489$)

	χ^2	df	CFI	Δ CFI	RMSEA [90% CI]	SRMR
Single model biology majors	259	167	0.972		0.054 [0.041–0.066]	0.042
Single model non-biology majors	435	167	0.962		0.065 [0.058–0.073]	0.038
Configural	703	334	0.965		0.062 [0.056–0.068]	0.039
Metric	739	351	0.964	0.001	0.061 [0.055–0.068]	0.050
Factor covariance	750	354	0.963	0.001	0.062 [0.056–0.068]	0.082
Scalar	804	368	0.959	0.005	0.063 [0.057–0.069]	0.050

MLR χ^2 chi-squared value for maximum likelihood estimation with robust standard errors, *df* degrees of freedom, *CFI* comparative fit index, *RMSEA* root mean squared error of approximation, *SRMR* standardized root-mean squared residual, *CI* confidence interval
 $p < .001$ for all models

model had good fit, indicating that the three-factor structure was invariant over biology majors and non-biology majors. Evaluation of the metric model, where the factor loadings were constrained to be equal in the two groups, supported metric invariance (Δ CFI $< .01$). Equal factor loading across groups means that each item contributes equally to the factors for both groups. In the last step, scalar invariance was assessed by constraining intercept and factor loadings to be equal across the two groups. Scalar invariance was supported (Δ CFI $< .01$) meaning that the intercept for each item is equal for both groups. In other words, if biology majors and non-biology majors have equal value on the latent construct (e.g., same feelings related to biology), they also have equal values on the items the construct is based on. Having established scalar invariance means that mean comparison of the different subscales of interest across biology and biology majors are justified.

When testing whether the covariances between the latent variables were equal across groups, Δ CFI was $< .01$, but SRMR (which is sensitive to miss-specified factor correlations; Vandenberg & Lance, 2000) increased from .050 to .082. This indicates that factor correlations between the latent factors may differ between groups. An inspection of the patterns of factor correlations for the metric model (where factor correlations are allowed to differ between groups) showed factor correlations were lower for biology majors than for non-biology majors (standardized factor correlations for non-biology majors were 0.923, 0.828, and 0.835, and for biology majors 0.884, 0.716, and 0.680 between feelings—value, feelings—

reengagement, and value—reengagement). In other words, there seem to be a stronger relation between the different dimensions of interest for non-biology majors compared to biology majors.

Differences between biology majors and non-biology majors

Analysis of the latent means differences between biology and non-biology majors showed that biology majors scored significantly higher on all interest subscales (Table 6).

Density plots based on mean scale scores also showed differences in the each subscale between biology majors and non-biology majors (Fig. 2). Biology majors consistently had a more positive distribution for all subscales. Value showed a rather high positive skewness for biology majors while reengagement was more normally distributed.

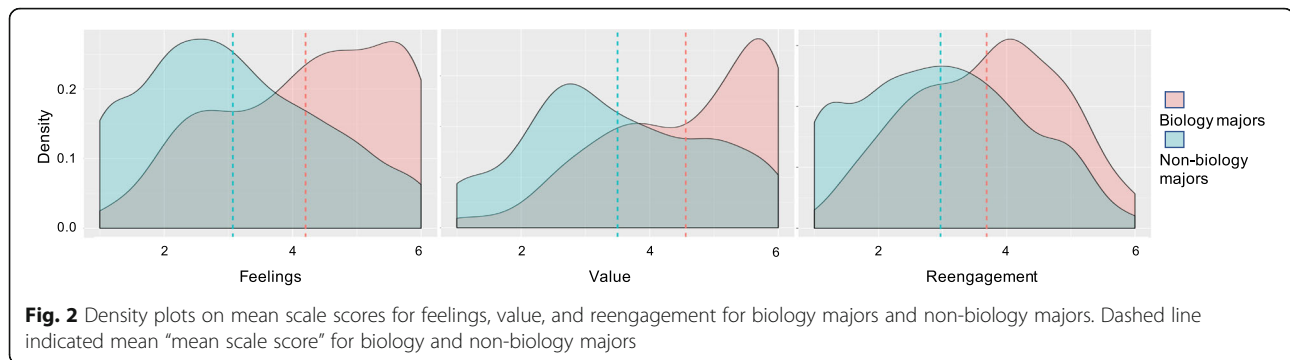
Discussion

Understanding how individual interest develops is an important aspect to understand in order to increase student retention in STEM education. When researching interest or evaluating the efficacy of interventions targeting interest, it is essential to have a tool to measure students' interest. We developed an instrument measuring undergraduate students' individual interest in the discipline of biology and collected initial validity evidence supporting the proposed use of the instrument. Below we first discuss instrument quality in terms of the internal structure, measurement invariance, and reflection about the response scales used for the instrument. We then continue with some theoretical reflection regarding how the different dimensions, feelings, value, and reengagement, relate to each other and what the value and reengagement dimensions really represent as conceptualized in our instrument. We close with implications and future research.

Table 6 Latent mean differences between biology and non-biology majors

	Estimate	Std. Err
Feelings	1.10***	0.110
Value	1.13***	0.111
Reengagement	0.68***	0.094

*** $p < .001$



Instrument quality

The psychometric properties of the final instrument were good: CFA on our data supported the suggested three factor structure, and expected differences between biology and non-biology majors were found. The final instrument includes 6 items representing students’ positive feelings towards biology, 5 items representing perceived value of biology, and 9 items representing students’ independent and voluntary reengagement in biology-related content and activities.

In line with the results from the evaluation of the SIQ instrument (Schiefele et al., 1993), on which our instrument partly builds, the feelings and value scales were closely correlated, and a CFA with only two factors (with feelings and values on the same scale) provided relatively good fit to our data. Unlike the SIQ instrument by Schiefele et al. (1993), we concluded that distinguishing the feelings and value aspects was possible with the BIQ instrument. Thus, our final instrument supports the three distinct scales described above.

Measurement invariance was supported, and thus the instrument can be used to compare the individual interest of biology majors and non-biology majors. Comparisons of latent means showed expected differences between the two groups: biology majors had higher latent means on all three subscales than non-biology majors. This aligns with Renninger and Hidi’s (2016) statement that a more developed individual interest in biology is characterized by more positive feelings, higher perceived value for biology, and increased intention to engage in biology-related activities, and thus the comparison can be considered external validity evidence for the instrument.

Because we expected our population to have rather high individual interest in biology, we used a positively packed response scale for the feelings and value items (Brown, 2004; Brown et al., 2017). Even with this scale, we saw positively skewed responses, especially for the value scale. Mean values for the value-related items included in the final scales ranged between 4.18 and 4.76 (scaled ranged from 1 to 6). Thus, we deemed it justified

to have a positively packed scale since it most likely resulted in larger variation in the responses and avoided ceiling effects compared to having a balanced scale. Further, having the positively packed scale leaves some room for interest to grow in subsequent measures. Still, our data suggest that it is easy for students to agree with the feelings- and value-related items. If the main aim of a study is to understand the feelings and value aspects of interest for populations likely to have a well-developed interest, a rewording of the items to make them even harder to agree on (for example *Learning about biology has always been important to me* could be changed to *Learning about biology has always been very important to me*) in combination with a positively packed scale could be useful. On the other hand, it could be that students beginning a biology major program might very easily agree with *I like reading about biology*. As these students continue to study biology, they might develop a more nuanced picture of what biology is and what their interests are, and consequently, they may develop more interest in some areas of biology and less in others. This might lead to more nuanced responses to the value and feelings scales, which could result in a general decrease of students’ scores on these scales over time. Thus, even though the initial responses started high, we did not deem it necessary to reword the items for our proposed use of the questionnaire.

The reengagement-related items had a balanced scale, and still the variation in responses was larger, and the distribution of the response options was closer to a normal distribution than for the feelings and value scale. In that sense, our results support Hidi and Renninger’s (2006) statement that reengagement is a suitable indicator of individual interest. It might be easy to agree on *I like reading about biology*, but having to consider the behavioral outcome of interest might result in a more objective and less biased estimate of individual interest.

Theoretical reflections regarding the interest construct

Hidi and Renninger (2006) wrote that as individual interest develops, students’ intention to independently

and voluntary reengage in biology-related content and activities would increase to a greater degree than the increase in feelings and value. Furthermore, they stated that as individual interest develops, the correlation between feelings, and value, and reengagement will get stronger (Renninger & Hidi, 2016). Our results did not support these two assumptions. First, we observed a greater difference in feelings and values than in reengagement between the biology majors (which had more interest) and non-biology majors. Second, the invariance analysis of correlation between the three subscales showed that non-biology majors had stronger correlations between all subscales compared to biology majors. Thus, our results indicate that as students' individual interest develops, feelings and values increase more than reengagement, and the correlation between the subconstructs get weaker. Although we cannot know the reasons for this pattern, it might be that biology majors over-report positive feelings and perceived value while their estimate of reengagement is more realistic or even underestimated compared to non-biology majors. This could result in lower correlations between the constructs for biology majors.

As described in the introduction, interest is included in many motivational theories, as well as in theories focusing on interest in particular, and is defined in many different ways (Krapp, 2002; Renninger & Hidi, 2016; Ryan & Deci, 2000; Schiefele, 1991; Wigfield & Eccles, 2000). Thus, the theoretical meaning of interest differs depending on the theoretical framework and definitions used. This is true both on the general level (i.e., what interest is and how many/what dimensions it includes) as well as the specific definitions of the different dimensions of interest. The theory used shapes how an instrument is designed. The instrument used impacts the results, interpretations made, and in the long run the overall understanding of the construct. To accomplish our instrument's aims, we used the four-phase model of interest development to understand interest. We found the value aspect challenging to conceptualize in our instrument. Different papers describing the value component of interest present slightly different conceptualizations of this component (e.g., Linnenbrink-Garcia et al., 2010; Renninger et al., 2002; Schiefele, 2009). We chose to define perceived value as the *personal significance of biology* (e.g., *self-realization, centrality with one's self concept*). This is in line with how Schiefele (2009) and Linnenbrink-Garcia et al. (2010) conceptualize the value component but does differ from the conceptualization given by Renninger et al. (2002). According to Renninger et al. (2002), the stored value component "refers to both a person's developing feelings of competence (...), and the corresponding positive and negative emotions that surface as he or she works to answer curiosity questions (...)" (p. 469). This complex

conceptualization maybe well suited for observational studies, but it is hard to fully capture with the use of a self-report instrument.

Further, according to four-phase model, value is intrinsic in its nature (Hidi & Renninger, 2006). Thus, the perceived value is directly towards a certain object; it is not based on an indirect relationship through other objects or domains. Valuing biology because passing biology courses is necessary to get into medical school does not count as interest in this model. We have strived to word our value items in such way that they reflect the intrinsic nature of interest. However, we cannot rule out the possibility that students' perceptions of the utility of biology have partly affected their responses on the value items.

Similar to considerations of the conceptualization of value, we feel it is important to elucidate some aspects of reengagement. Reengagement items were chosen to reflect independent and voluntary engagement, as opposed to reengagement initiating from an external source or opportunity. However, we must consider that one's engagement is biased by availability and opportunity. Some students might want to independently and voluntarily engage in different biology-related activities but might not have the opportunity to do so. We have strived to include items with as little opportunity bias as possible. We did this by only including items concerning activities that we thought were easily accessible for most students. Despite the potential bias introduced by opportunity, we argue that reengagement is still relevant to include when assessing students' individual interest. Unfortunately, students who want to engage but do not have the opportunity are not completing the positive feedback loop that further strengthens interest. Renninger and Hidi's (2016) framework suggests that engagement increases knowledge, and knowledge contributes to the deepening of feelings about and valuing for a certain object. As value and feelings develops, they lead to continued engagement and so on, resulting in further interest development. Thus, whatever the cause of the lack of reengagement, understanding the level of reengagement will be important for characterizing interest, especially when following students longitudinally.

Limitations and future research

The current instrument shows promising properties is brief and easily administered, yet covers several relevant aspects of interest. However, validation is a continuous and iterative process that involves accumulation of evidence to support proposed score interpretation (AERA, APA, & NCME, 2014), and therefore we recommend more studies collecting validity evidence to further strengthen the validity of the interpretation of the scores from the BIQ for researchers' proposed use. Repeated

factor analysis on new samples should be conducted to confirm the internal structure of the BIQ found here. Future studies could also apply Rasch techniques to yield additional insight into the internal structure of the instrument and the item response processes (Boone, 2016).

The use of the BIQ with students in introductory biology classes was tested here. Further validity evidence is still needed to support the use of this instrument for longitudinal studies and with upper division biology majors.

Finally, additional types of validity evidence could add nuance to our understanding of the BIQ and interest itself. For example, interest is a complex construct that is closely related to other constructs such as relevance, utility, and curiosity (Priniski, Hecht, & Harackiewicz, 2018; Silvia, 2006). Empirical studies comparing results from the BIQ with results from utility, curiosity and or relevance instruments could provide additional validity evidence for the BIQ based on external relationships and contribute to our theoretical understanding of the interest construct. Inclusion of more aspects of individual interest described by Renninger and Hidi (2016) in future iterations of this instrument could also further the understanding of how the different aspects of individual interest relate to each other and how they change as interest develops.

Conclusions

The study findings support the use of the BIQ to gain a broad understanding of students' individual interest in biology. The BIQ was developed within the four-phase model of interest development framework. In addition, the feelings subscale could be relevant for other theoretical frameworks of interest, e.g., expectancy value theory (Wigfield & Eccles, 2000) or self-determination theory (Ryan & Deci, 2000). The reengagement subscale could be an indication of interest independent of the theory used. The BIQ provides researchers, departments, and teachers an instrument they can use when evaluating interventions targeting students' interest. We encourage STEM education researchers in other disciplines to adapt and collect validity evidence for the instrument in their disciplines in order to gain a broad understanding of students' interest across disciplines and begin building a compendium of knowledge supporting students' persistence in STEM education.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s40594-020-00217-4>.

Additional file 1.

Abbreviations

BIQ: Biology Interest Questionnaire; CFA: Confirmatory factor analysis; CFI: Comparative fit index; EFA: Exploratory factor analysis; MLR: Robust maximum likelihood estimator; MLR χ^2 : Chi-squared value from robust maximum likelihood estimation; PISA: Program for International Students Assessment; RMSEA: The root mean squared error of approximation; SIQ: Study Interest Questionnaire; SRMR: The standardized root-mean-squared residual; WLS: Weighted least square

Acknowledgements

We would like to thank Ann Renninger for helpful feedback regarding the conceptualization of interest.

Authors' contributions

EK and AR reviewed the literature and designed the survey instrument with input from SE and LC. All authors discussed the theoretical framework choices. EK cleaned the data, analyzed, and interpreted the data with input from SE. All authors discussed the instrument revision. EK had the main responsibility for writing the paper, but all authors contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Funding

Research in this paper was funded by NSF #1712160 and #1711082.

Availability of data and materials

The datasets generated and analyzed during the current study are available anonymized from the corresponding author by request.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Science and Mathematics Education, Umeå University, 901 87 Umeå, Sweden. ²Department of Biological Sciences, Florida International University, Miami, FL 33199, USA. ³Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, Boulder, CO 80309, USA.

Received: 21 November 2019 Accepted: 13 April 2020

Published online: 19 May 2020

References

- Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology, 94*(3), 545–561. <https://doi.org/10.1037/0022-0663.94.3.545>.
- Alexander, P. A., & Murphy, P. K. (1998). Profiling the differences in students' knowledge, interest, and strategic processing. *Journal of Educational Psychology, 90*(3), 435–447.
- American Educational Research Association [AERA], American Psychological Association [APA], & National Council for Measurement in Education [NCME]. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Bandalos & Finney. (2010). Factor analysis. Exploratory and confirmatory. In Hancock & Mueller (Eds.), *The reviewer's guide to quantitative methods in the social science* (pp. 93–114). New York: Routledge.
- Barr, D. A., Gonzalez, M. E., & Wanat, S. F. (2008). The leaky pipeline: Factors associated with early decline in interest in premedical studies among underrepresented minority undergraduate students. *Academic Medicine, 83*(5), 503–511.
- Bashkov, B. M., & Finney, S. J. (2013). Applying longitudinal mean and covariance structures (LMACS) analysis to assess construct stability over two time points: An example using psychological entitlement. *Measurement & Evaluation in Counseling & Development, 46*(4), 289–314. <https://doi.org/10.1177/0748175613497038>.
- Bong, M., Lee, S. K., & Woo, Y.-K. (2015). The role of interest and self-efficacy in pursuing mathematics and science. In K. A. Renninger, M. Nieswandt, & S. Hidi (Eds.), *Interest in Mathematics and Science Learning* (pp. 33–48). Washington, DC: American Educational Research Association.
- Boone, W. J. (2016). Rasch analysis for instrument development: Why, when, and how? *CBE—Life Sciences Education, 15*(4), rm4.

- Brown, G. T. L. (2004). Measuring attitude with positively packed self-report ratings: Comparison of agreement and frequency scales. *Psychological Reports*, 94(3), 1015–1024. <https://doi.org/10.2466/pr0.94.3.1015-1024>.
- Brown, G. T. L., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: Identifying and understanding non-invariance. *International Journal of Research & Method in Education*, 40(1), 66–90. <https://doi.org/10.1080/1743727X.2015.1070823>.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5.
- Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education*. (7) Milton Park, Abingdon, Oxon, [England]: \
- Dall, T., West, T., Chakrabarti, R., Reynolds, R., & Iacobucci, W. (2018). 2018 Update the complexities of physician supply and demand: Projections from 2016 to 2030. *Final Report Association of American Medical Colleges*. <https://doi.org/10.13140/RG.2.2.25694.48963>.
- Durik, A. M., & Harackiewicz, J. M. (2003). Achievement goals and intrinsic motivation: Coherence, concordance, and achievement orientation. *Journal of Experimental Social Psychology*, 39(4), 378–385.
- Fouad, N. A., Chang, W. H., Wan, M., & Singh, R. (2017). Women's reasons for leaving the engineering field. *Frontiers in Psychology*, 8, 875.
- Gagne, P., & Hancock, G. R. (2006). Measurement model quality, sample size, and solution propriety in confirmatory factor models. *Multivariate Behavioral Research*, 41(1), 65–83.
- Gignac, G. E. (2009). Psychometrics and the measurement of emotional intelligence. In C. Stough, D. H. Saklofske, & J. D. A. Parker (Eds.), *Assessing emotional intelligence: Theory, research, and applications* (pp. 9–40). New York: Springer US. https://doi.org/10.1007/978-0-387-88370-0_2.
- Glynn, S. M., Bryan, R. R., Brickman, P., & Armstrong, N. (2015). Intrinsic motivation, self-efficacy, and interest in science. In K. A. Renninger, M. Nieswandt, & S. Hidi (Eds.), *Interest in mathematics and science learning* (pp. 189–202). Washington, DC: American Educational Research Association.
- Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology*, 92(2), 316–330.
- Harackiewicz, J. M., & Hulleman, C. S. (2010). The importance of interest: The role of achievement goals and task values in promoting the development of interest. *Social and personality psychology compass*, 4(1), 42–52. <https://doi.org/10.1111/j.1751-9004.2009.00207.x>.
- Hebbali, A. (2018). Tools for building OLS Regression Models, version 0.5.1. Available from <https://cran.r-project.org/web/packages/olsr>
- Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, 60(4), 549–571. <https://doi.org/10.3102/00346543060004549>.
- Hidi, S., & Harackiewicz, J. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research*, 70, 151–179.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4.
- Holtermann, N., Grube, D., & Bögeholz, S. (2010). Hands-on activities and their influence on students' interest. *Research in Science Education*, 40(5), 743–757. <https://doi.org/10.1007/s11165-009-9142-0>.
- Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Knekta, E., Runyon, C., & Eddy, S. (2019). One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research. *CBE—Life Sciences Education*, 18(1), rm1. <https://doi.org/10.1187/cbe.18-04-0064>.
- Krapp, A. (2002). An educational-psychological theory of interest and its relation to SDT interest: A rediscovered motivational concept. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 405–426). Rochester: University of Rochester Press.
- Lent, R. W., Brown, S. D., & Hackett, G. (2002). Social cognitive career theory. In D. Brown (Ed.), *Career choice and development* (pp. 255–311). San Francisco: Jossey-Bass.
- Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement*, 70(4), 647–671. <https://doi.org/10.1177/0013164409355699>.
- McAuley, E., Duncan, T., & Tammen, V. V. (1987). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, 60, 48–58.
- McDaniel, M. A., Waddill, P. J., Finstad, K., & Bourg, T. (2000). The effects of text-based interest on attention and recall. *Journal of Educational Psychology*, 92(3), 492–502.
- OECD. (2017). *PISA 2015 Assessment and analytical framework: Science, reading, mathematics, financial literacy and collaborative problem solving*. Paris: PISA, OECD Publishing <https://doi.org/10.1787/9789264281820-en>.
- PCAST. (2012). *Engage to Excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: PCAST.
- Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33–40. <https://doi.org/10.1037/0022-0663.82.1.33>.
- Priniski, S. J., Hecht, C. A., & Harackiewicz, J. M. (2018). Making learning personally meaningful: A new framework for relevance research. *The Journal of Experimental Education*, 86(1), 11–29. <https://doi.org/10.1080/00220973.2017.1380589>.
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing <https://www.R-project.org/>.
- Renninger, K. A., & Bachrach, J. E. (2015). Studying triggers for interest and engagement using observational methods. *Educational Psychologist*, 50(1), 58–69. <https://doi.org/10.1080/00461520.2014.999920>.
- Renninger, K. A., Ewen, L., & Lasher, A. K. (2002). Individual interest as context in expository text and mathematical word problems. *Learning and Instruction*, 12(4), 467–490. [https://doi.org/10.1016/S0959-4752\(01\)00012-3](https://doi.org/10.1016/S0959-4752(01)00012-3).
- Renninger, K. A., & Hidi, S. E. (2016). *The power of interest for motivation and engagement*. Abingdon: Routledge. <https://doi.org/10.4324/9781315771045>.
- Revelle, W. (2017). psych: Procedures for personality and psychological research, Northwestern University, Evanston, Illinois, USA. Retrieved February 24, 2019, from <https://CRAN.R-project.org/package=psychVersion=1.7.8>
- Rodenbusch, S. E., Hernandez, P. R., Simmons, S. L., & Dolan, E. L. (2016). Early engagement in course-based research increases graduation rates and completion of science, engineering, and mathematics degrees. *CBE—Life Sciences Education*, 15(2), ar20.
- Rosseel, Y. (2012). lavaan: An R Package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Rowland, A. A., Knekta, E., Eddy, S., & Corwin, L. A. (2019). Defining and measuring students' interest in biology: An analysis of the biology education literature. *CBE—Life Sciences Education*, 18(3), ar34. <https://doi.org/10.1187/cbe.19-02-0037>.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>.
- Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, 26(3–4), 299–323. <https://doi.org/10.1080/00461520.1991.9653136>.
- Schiefele, U. (2009). Situational and individual interest. In K. R. Wenzel & A. Wigfield (Eds.), *Educational psychology handbook series. Handbook of motivation at school* (pp. 197–222). Routledge/Taylor & Francis Group: New York.
- Schiefele, U., & Csikszentmihalyi, M. (1994). Interest and the quality of experience in classrooms. *European Journal of Psychology of Education*, 9(3), 251–270.
- Schiefele, U., Krapp, A., Wild, K.-P., & Winteler, A. (1993). Der "Fragebogen zum Studieninteresse" (FSI). / The Study Interest Questionnaire (SIQ). *Diagnostica*, 39, 335–351.
- Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 183–212). Hillsdale: Erlbaum Retrieved January 22, 2019, from http://opus.kobv.de/ubp/volltexte/2009/3352/pdf/schiefele1992_8.pdf.
- Schreiner, C., & Sjøberg, S. (2004). *Sowing the seeds of ROSE. Background, rationale, questionnaire development and data collection for the ROSE (The relevance of science education): A comparative study of students' views of science and science education*. Oslo: Acta Didactica, Department of Teacher Education and School Development, Faculty of Education, University of Oslo.
- Seymour, E., & Hunter, A.-B. (2019). *Talking about leaving revisited: Persistence, relocation, and loss in undergraduate STEM education*. Cham: Springer Nature Switzerland AG.
- Silvia, P. J. (2006). *Exploring the psychology of interest*. Oxford: Oxford University Press.

- Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., De Chenne-Peters, S. E., & Levis-Fitzgerald, M. (2018). Anatomy of STEM teaching in North American universities. *Science*, *359*(6383), 1468–1470.
- Tabachnick & Fidell. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson.
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1–67 URL <https://www.jstatsoft.org/v45/i03/>.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70. <https://doi.org/10.1177/109442810031002>.
- Wentzel, K. R., & Miele, D. B. (2016). *Handbook of motivation at school*. Abingdon: Routledge.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>.
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, *73*(6), 913–934.
- Zhao, C. M., & Kuh, G. D. (2004). Adding value: Learning communities and student engagement. *Research in Higher Education*, *45*(2), 115–138.

Publisher's Note

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

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
