# Improving Self-Reported Measures of Innovation Skills in an Interdisciplinary Undergraduate STEM Course



Haider Ali Bhatti
University of California, Berkeley


Perman Gochyyev
University of California, Berkeley


Mark Wilson
University of California, Berkeley

## NARST 2022 Proposal Paper

**Title:** Improving Self-Reported Measures of 21st Century Skills in an Interdisciplinary Undergraduate STEM Course

**Abstract:**
In an effort to push STEM education beyond content-based memorization and towards the development of skills-based practices applicable to the careers of tomorrow, this paper describes how an interdisciplinary STEM course called *Bioinspired Design* improved self-reported measures of 21st Century Skills in undergraduate students. In this course open to all majors and all years, students worked in interdisciplinary teams to translate authentic scientific discoveries from primary literature into societally-impactful bioinspired designs. We hypothesized that students would grow in 21st Century Skills such as Scientific Discovery & Translation Process, Interdisciplinary Thinking, and Interdisciplinary Collaboration as a result of completing the course and engaging in course activities. We assessed this hypothesis through a survey-based methodology analyzed using item response theory. We considered students' growth as aligned with our construct map that included five levels of ascending competence related to 21st Century Skills (Required, Technical, Participant, Active, Leader). A pre/post comparison of students' self-reported 21st Century Skills showed growth in all survey items, equating to 0.89 standard deviations of growth in student ability as a result of completing the *Bioinspired Design* course. This translates to a one "step" increase across all items on the Likert scale (i.e., "Agree" to "Strongly agree").

**Authors:**
Haider A. Bhatti, Perman Gochyyev, Mark Wilson, Robert J. Full

## 1. Subject/problem

The large majority of undergraduate STEM education (USE) is taught through didactic lecture rather than more student-centered approaches (Stains et al., 2018). Student-centered approaches to teaching, particularly the integration of 21st Century Skills (21CS), provide an avenue to make USE more evidence-based and pedagogically robust. For example, a number of 21CS's have been proposed as critical education outcomes (Griffin & Care, 2015; Trilling & Fadel, 2009; NRC, 2012). These include critical thinking and problem solving, creativity and innovation, collaboration, teamwork, and leadership, cross-cultural understanding, communications, information, and media literacy, computing and technology literacy, and career learning of self-reliance (Trilling & Fadel, 2009). By integrating these 21CS's into USE, we can promote a shift towards more student-centered learning.

With this needed shift in mind, we developed an undergraduate STEM course rooted in 21CS called *Bioinspired Design* at a large research university. The course is open to all majors (e.g., STEM, non-STEM) and all years (e.g., freshman through senior) with no prerequisites. Importantly, this course utilizes a 21CS approach to learning by promoting inquiry, teaming, and discovery-based learning, all within the interdisciplinary domain of bioinspired design. This approach empowers learners as they navigate learning environments that are also community-centered and intensely collaborative. Students help one another solve problems, build on each other's knowledge, ask questions to clarify explanations, and suggest avenues that would move their team toward its goal (NRC, 2000). By fostering such communities of learners (Brown & Campione, 1996), teams are further motivated to pursue bioinspired design ideas that are particularly relevant to them and society at-large.

Additionally, our course aligns with the AAAS *Vision and Change* (2011) reports that classify "the ability to tap into the interdisciplinary nature of science" and "the ability to communicate and collaborate with other disciplines" (p. 15) as core competencies within any undergraduate biology course. The concept of interdisciplinarity is further elucidated by Tripp and Shortlidge (2019) their Interdisciplinary Science Framework (IDSF) for USE. In this proposal, we build on these frameworks of interdisciplinarity by considering interdisciplinarity as part of a larger, multidimensional latent construct—21st Century Skills (21CS). Here, we present the development, validation, and results of a pre/post 21CS self-assessment for students enrolled in our *Bioinspired Design* course.

Our course attempts to address many of the aforementioned 21CS's, but here we focus on the specific assessment of the 21CS's that are critical to success in not only our course, but all transformative STEM courses; (a) Scientific Discovery & Translation Process (SDTP); (b) Interdisciplinary Thinking (IT); and (c) Interdisciplinary Collaboration (IC). Our study focused on exploring pre to post changes in self-reported measures of 21CS through a newly developed Likert-type instrument. The rationale motivating this research was to assess how students' self-reported measures of 21CS change as a result of participating in our *Bioinspired Design* course. In this course, students work in interdisciplinary teams to translate authentic scientific discoveries from primary literature into societally-impactful bioinspired designs. As a result of this process, we hypothesized that students engage in the three proposed subdimensions of the 21CS construct as they participate in the course. Once students have completed this course, we hypothesized that students' self-reported measures of 21CS would increase as a result of engaging in the course activities.

## 2. Design or Procedure

This proposal describes a survey-based methodology that utilizes item response theory (IRT) to analyze students' self-reported 21CS in a transformative undergraduate STEM course. We utilized survey-based research methods to administer our 21CS pre/post course survey. The survey consisted of 26 Likert-type items. The items were originally developed as part of a previously studied pre/post course survey developed using the "University" Assessment System (UAS) (Authors et al., 2015; Author & Author, 2000). Students enrolled in the *Bioinspired Design* course completed the survey as a voluntary part of their course activities. Our survey population was a broadly representative sample of undergraduate students at the university because of the inclusive nature of the course (open to all majors, all years, no prerequisites, satisfies a breadth requirement). This is unlike many other discovery-based STEM courses that often have disciplinary barriers to entry (i.e., introductory STEM prerequisites, such as general biology), especially for students from non-STEM backgrounds.

The analysis in this study is based on data collected from 2016 to 2020, resulting in 514 pre- and 432 post-survey responses. Our pre/post survey instrument remained consistent throughout that time, containing the same 26 Likert-type items designed to measure students' self-reported 21CS. Our preliminary analysis of pre/post changes in raw Likert scores showed increases in "agreeability" for all items, including those mapped to the highest levels of the construct map (i.e., items that we considered "hardest" to agree with). Thus, the mean score for each item on the survey increased from pre to post, resulting in a positive delta value for each item. Students showed growth in all skills and subdimensions of our 21CS construct every year after completing the course. However, considering the limitations of analyzing raw scores on self-reported Likert instruments (Chimi & Russell, 2009), we recognized the need to conduct a more in-depth IRT analysis of these data.

An important element of our research design was our measurement framework. This study was based on the UAS four building blocks of assessment proposed by Author (2004). The four building blocks—construct map, items design, outcome space, and measurement model—provide an interconnected and iterative cycle of assessment to ground our quantitative analysis of 21CS. The first building block, the construct map, is a construct definition tool that relies on a developmental perspective to assess student achievement and growth. Construct maps focus on qualitatively different levels of performance on some latent construct. For our proposal, we considered 21CS as an overall latent construct, or dimension of measurement, spread across the three subdimensions of SDTP, IT, and IC. Each of these subdimensions consists of a developmental perspective ranging from Required (lowest level) to Leader (highest level), resulting in three construct maps (Figure 1).

The second building block, the items design, focuses on the match between instruction and the types of assessment, often determined by the match between assessment tasks and levels within the construct maps. In our 21CS survey, each of the 26 items is mapped onto a level within a subdimenison construct map. The items mapped onto the highest levels (Leader) are the most difficult to agree with on the Likert scale whereas the items mapped onto the lowest levels (Required) are the easiest to agree with on the Likert scale. All together, the items within subdimensions form the basis of our multidimensional 21CS construct. The third building block, the outcome space, refers to a procedure for classifying or mapping responses to survey items. Our 21CS survey is a Likert scale instrument in which the outcome space includes the distinct Likert categories ranging from strongly disagree to strongly agree. The final building block, the measurement model, is typically visualized as item-person Wright maps. These Wright maps represent a visualization of the construct map based on student data and responses. Collectively, each of the four building blocks are iterative parts of the assessment cycle, leading to refinements of the instrument, the instruction in the course, and the measurement of the latent construct as a whole.

**Scientific Discovery & Translation Process**

(1) strongly disagree, (2) disagree, (3) neutral, (4) agree, or (5) strongly agree

| Level | *Skill* |
|---|---|
| Leader | Can extract fundamental principles from scientific publications.<br>Can translate fundamental principles from scientific publications to create novel designs. |
| Active | Can critically evaluate scientific publications.<br>Can suggest fundamental principles from scientific publications that could translate into novel designs. |
| Participant | Can understand the logic of scientific publications when discussed by experts.<br>Can understand how experts have translated principles from scientific publications into novel designs. |
| Technical | Reads scientific publications to extract specific facts. |
| Required | Reads scientific publications because required. |

**Interdisciplinary Thinking**

(1) strongly disagree, (2) disagree, (3) neutral, (4) agree, or (5) strongly agree

| Level | *Skill* |
|---|---|
| Leader | Apply concepts and methodologies across disciplines.<br>Uses interdisciplinary concepts and methodologies to advance own discipline.<br>Contributes to advancement in other fields with knowledge from my discipline. |
| Active | Attempt to apply concepts and methodologies across disciplines.<br>Seeks concepts and methodologies from other fields. |
| Participant | Recognizes the benefits from interacting with other disciplines.<br>Aware of concepts and methodologies across disciplines. |
| Technical | Primarily uses technical knowledge in own discipline. |
| Required | Seeks information from own discipline to solve problems.<br>Seeks information from disciplines other than their own when they are required. |

**Interdisciplinary Collaboration**

(1) strongly disagree, (2) disagree, (3) neutral, (4) agree, or (5) strongly agree

| Level | *Skill* |
|---|---|
| Leader | Leadership role in interdisciplinary teams.<br>Contributes to and benefits from interdisciplinary collaborations. |
| Active | Active role in interdisciplinary collaborations.<br>Seeks opportunities outside the discipline. |
| Participant | Participates in diverse teams because recognizes the benefits of interdisciplinary knowledge. |
| Technical | Seek interdisciplinary collaborations to attain technical or factual knowledge outside discipline. |
| Required | Prefer collaboration among members of my discipline.<br>Participates in interdisciplinary teams when required. |

Figure 1: Construct maps for 21st Century Skills subdimensions with each level mapping onto distinct survey items categorized here under "Skill"

Based on this measurement framework, we compiled two separate datasets that represented the pre ($N = 514$) and post-survey ($N = 432$) responses. Each dataset contained pre/post responses to all 26 items. The analysis considered all of the pre responses as one large collection of data to compare with the post responses, a separate collection of data. Therefore, any changes from pre to post were a proxy for changes in students' 21CS as a result of completing the *Bioinspired Design* course. We considered the *Bioinspired Design* course to represent an "intervention" through which we could compare students' 21CS before the intervention and after the intervention (i.e., before and after the course). To conduct our IRT analysis, we utilized the ACER ConQuest software (Adams et al., 2020) to fit the 21CS survey data to a unidimensional partial credit model (PCM). Additionally, item difficulty estimates of the same items from separate administrations were anchored to establish a common metric. Specifically, for the pretest, the item parameters were calibrated by anchoring the items to the estimates obtained from post-test calibration. As a result, both pre and post tests were put on the common scale. This approach to analyze the pre- and post- test provides a strong factorial measurement invariance (Millsap, 2011).

## 3. Analyses and Findings

With our measurement framework and IRT model in mind, we can now answer our main research question—*what is the change in self-reported 21CS from pre to post as a result of completing the Bioinspired Design course?* In other words, how do the results from the anchored pre analysis compare with the results from the anchored post analysis? We begin this comparison by interpreting the pre and post Wright maps in Figure 2. The Wright maps show item responses and persons on a logit scale in which the more positive responses map to higher levels of the construct map. Item parameter thresholds show the average difficulty (in logits) of the item thresholds relative to the other item thresholds. The distribution of the item parameter thresholds on a logit scale is displayed as the item number and threshold level (e.g., 10.4) on the right hand side of the Wright map. The larger the item parameter threshold (or the higher on the Wright Map), the more difficult the item is to agree with and therefore, less likely for a student to strongly agree with the item statement. The left hand side of the Wright map displays the distribution of student abilities as X's. As a student's average ability increases, that student is more likely to agree with the item statements and therefore, the student is exhibiting higher levels of 21CS.

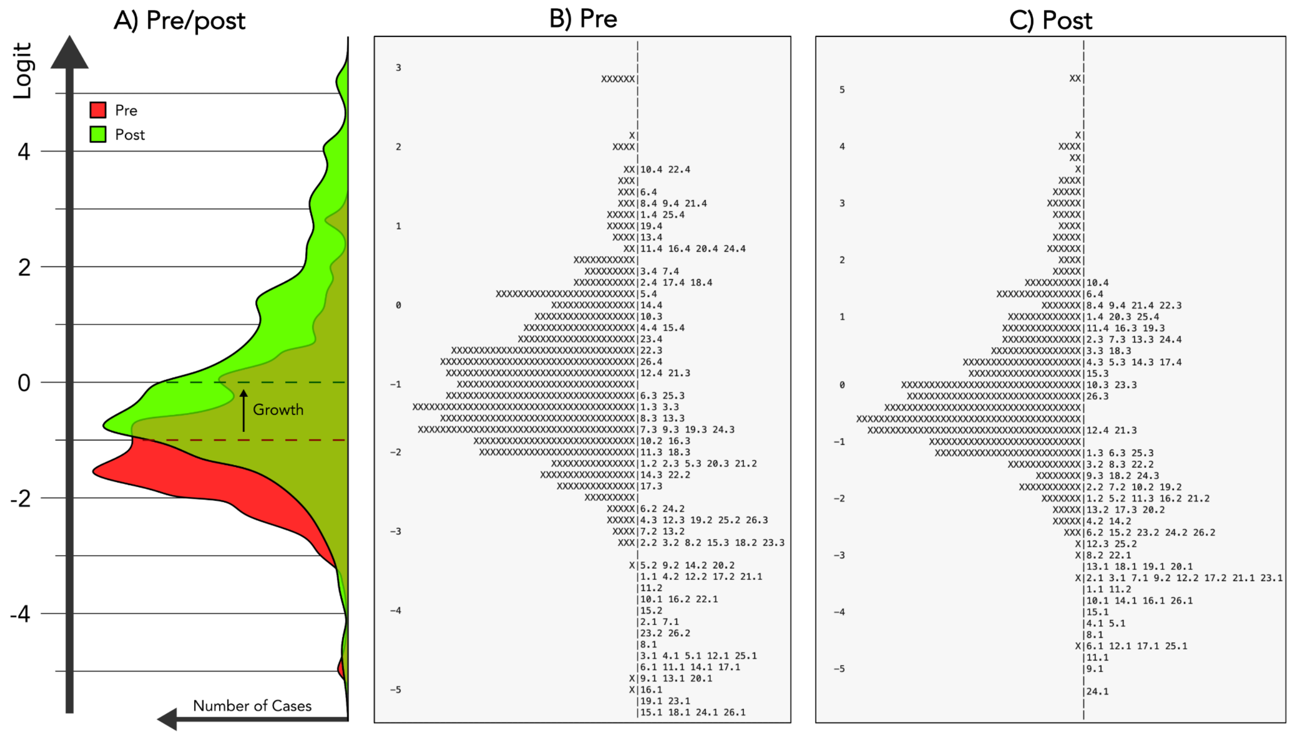


Figure 2: A) Growth from pre to post (+0.97 logits) visualized through overlaid Wright maps; B) Pretest Wright map; C) Post-test Wright map

Both the pre and the post Wright maps show expected banding patterns consistent with leveled thresholds. This shows that the ordered levels of responses (strongly disagree to strongly agree) are consistent with the expected construct map levels. As compared to the pre Wright map, the distribution of student abilities in the post Wright map indicates that more items are needed to assess the higher end of the ability scale. In other words, the pre shows the distribution of student abilities appropriately mapped to more difficult items (Figure 2B), but the post shows a large set of student abilities on the higher end of the logit scale well above any item thresholds (Figure 2C). Overall, more difficult-to-agree-with items are needed to accurately assess the higher end of the student ability distribution in the post. We interpret this to be a result of significant growth in students' 21CS after completing the *Bioinspired Design* course.

The Wright maps also indicate that there is strong quantifiable growth from pre to post in overall student ability. Because of the anchored nature of this analysis, we can summatively quantify this growth based on the outputted constant variable value in the anchored pre. A comparison of the constant in the anchored pre and in the post results in a difference that represents the "anchored" difference between the pre and post on the logit scale. This anchoring links the entire ability distribution between two time points—the post (0.00) and the anchored pre (-0.97). The difference between these two values indicates that the ability distribution of the anchored post is about 1 logit above the pre. Therefore, there is about a 1 logit gain in student ability on this 21CS survey as a result of completing the *Bioinspired Design* course.

Overall, our IRT comparison of pre and post results showed an approximately 1 logit gain (0.97) in student ability as a result of completing the *Bioinspired Design* course. This increase is roughly equivalent to a one "step" increase in the Likert scale across all items (i.e., "Agree" to "Strongly agree"). Put differently, this logit gain translates to approximately 0.89 standard deviations of growth between pre and post. The logit gain can also be converted to a pooled standard deviation between the pre and the post, resulting in a Cohen's *d* effect size value of 0.75. This effect size is nearly a "large" effect (>0.8) based on Cohen's (1988) standards and an especially large effect in the the context of educational interventions (>0.2) (Kraft, 2020).

**4 + 5. Contribution & General Interest**

21CS are a critical part of being a scientifically literate citizen. Our proposal paper supports the NARST 2022 call for promoting educational paths that advance global scientific literacy. The NARST 2022 call for proposals asks, "What is our role toward building trust in science?" We show that engaging in activities related to Scientific Discovery & Translation Process directly improves how the public (i.e., our students) understand the nature of scientific knowledge and the processes by which that knowledge is generated. Additionally, we are also asked, "Who is invited into the scientific community? Who decides? As scholars, what questions are, and should, we pursue that disrupt exclusionary images of science?" We show that engaging in both the Interdisciplinary Thinking and Interdisciplinary Collaboration dismantles artificially created disciplinary boundaries to go far beyond STEM by including artists, designers, social scientists, and entrepreneurs collaborating in diverse teams using scientific discoveries to create inventions that could shape our future. By promoting 21CS, our course fosters a sense of science activism amongst all students, including nonmajors who, along with their STEM major counterparts, will make up a future citizenry that must be able to apply scientific knowledge and practices to make informed decisions as a part of their duties in a democratic society. Thus, we view learning within our course as not just STEM-based, but *STEM-enhanced*, applicable to areas within and outside of STEM.

**Acknowledgements**: We thank Kahraman G. Demir for his help and guidance in constructing Figure 2.

## References


Adams, R.J, Wu, M.L, Cloney, D., and Wilson, M.R. (2020). A*CER ConQuest: Generalised Item Response Modelling Software* [Computer software]. Version 5. Camberwell, Victoria: Australian Council for Educational Research.

American Association for the Advancement of Science. (2011). Vision and change: A call to action, final report. *Washington, DC. Retrieved August*, *1*, 2021.

Author, & Author. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, *13*, 181–208.

Author. (2004). *Constructing Measures: An Item Response Modeling Approach*. Routledge. https://doi.org/10.4324/9781410611697

Authors. (2015). Interdisciplinary Laboratory Course Facilitating Knowledge Integration, Mutualistic Teaming, and Original Discovery. *Integrative and Comparative Biology*, *55*(5), 912–925. https://doi.org/10.1093/icb/icv095

Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. *Innovations in Learning: New Environments for Education.*, 289–325.

Brown, E. R., Smith, J. L., Thoman, D. B., Allen, J. M., & Muragishi, G. (2015). From bench to bedside: A communal utility value intervention to enhance students' biomedical science motivation. *Journal of Educational Psychology*, *107*(4), 1116–1135. https://doi.org/10.1037/edu0000033

Canning, E. A., & Harackiewicz, J. M. (2015). Teach it, don't preach it: The differential effects of directly-communicated and self-generated utility–value information. *Motivation Science*, *1*(1), 47–71. https://doi.org/10.1037/mot0000015

Chamany, K., Allen, D., & Tanner, K. (2008). Making Biology Learning Relevant to Students: Integrating People, History, and Context into College Biology Teaching. *CBE—Life Sciences Education*, *7*(3), 267–278. https://doi.org/10.1187/cbe.08-06-0029

Chimi, C. J., & Russell, D. L. (2009, November). The Likert scale: A proposal for improvement using quasi-continuous variables. In I*nformation Systems Education Conference, Washington, DC* (pp. 1-10).

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Griffin, P., & Care, E. (Eds.). (2015). *Assessment and Teaching of 21st Century Skills*. Springer Netherlands. https://doi.org/10.1007/978-94-017-9395-7

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Millsap, R. E. (2012). *Statistical approaches to measurement invariance*. Routledge.

*National Research Council. (2000). How People Learn: Brain, Mind, Experience, and School: Expanded Edition. The National Academies Press. https://doi.org/10.17226/9853*

National Research Council. Education for life and work: Developing transferable knowledge and skills in the 21st century. National Academies Press; 2012 Dec 18.

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

Siritunga, D., Montero-Rojas, M., Carrero, K., Toro, G., Vélez, A., & Carrero-Martínez, F. A. (2011). Culturally Relevant Inquiry-Based Laboratory Module Implementations in

Upper-Division Genetics and Cell Biology Teaching Laboratories. *CBE—Life Sciences Education*, *10*(3), 287–297. https://doi.org/10.1187/cbe.11-04-0035

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., Eagan, M. K., Esson, J. M., Knight, J. K., Laski, F. A., Levis-Fitzgerald, M., Lee, C. J., Lo, S. M., McDonnell, L. M., McKay, T. A., Michelotti, N., Musgrove, A., Palmer, M. S., Plank, K. M., … Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, *359*(6383), 1468–1470. https://doi.org/10.1126/science.aap8892

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. (pp. xxxi, 206). Jossey-Bass/Wiley.

Tripp, B., & Shortlidge, E. E. (2019). A Framework to Guide Undergraduate Education in Interdisciplinary Science. *CBE—Life Sciences Education*, *18*(2), es3. https://doi.org/10.1187/cbe.18-11-0226