Physiology is hard: a replication study of students’ perceived learning difficulties

Tara Slominski, Sarah Grindberg, and Jennifer Momsen
Department of Biological Science, North Dakota State University, Fargo, North Dakota

Submitted 2 March 2018; accepted in final form 4 February 2019

INTRODUCTION

Human Anatomy and Physiology (HAP) is widely recognized as a difficult course, often characterized by high drop, withdrawal, and failure rates (10, 23). Unfortunately, the HAP education community has little research that explicitly investigates why undergraduate students struggle to succeed in HAP. Research suggests students find physiology difficult (Sturges D, Maurer T. Internet J Allied Health Sci Pract 11: 1–10, 2013). However, without replication, these claims are limited in their generalizability. There is a need in physiology education research to replicate studies like these across different institutions to support generalizations. We, therefore, replicated both of these studies by collecting survey responses from 466 students at 4 different institutions and 17 instructors at 15 different institutions. We found that students in our study identified similar factors as the students surveyed in the original study. Students most strongly agreed with items that attributed the difficulty of HAP to the nature of the discipline, as opposed to the way physiology is taught or the way students approach learning it. Faculty in our sample, like those in the original study by Michael, agreed most strongly with items that attributed physiology’s difficulty to discipline specific factors. Our data reinforce the results of Sturges and Maurer and Michael. We can more confidently claim that HAP students and faculty believe the difficulty in learning physiology is the result of inherent features of the discipline itself and not factors related to instruction or the students themselves.

While these results are valuable, they do not reflect students’ perceptions of the difficulties in learning human physiology. Sturges and Maurer (22) adapted the Michael (14) survey for use with students and reported the responses of 276 students (183 from HAP I, 93 from HAP II) from a single time point at a university in Georgia. These results revealed how students perceived the difficulties associated with learning physiology, as opposed to how faculty think students perceive these difficulties. Students in this population did not attribute their learning difficulties to the manner in which physiology was taught; rather, they attributed their learning difficulties to the inherent difficulty of the discipline. These results indicate some student agreement with faculty, as per the Michael study. Furthermore, students in this study showed high agreement with statements attributing learning difficulties to teleological tendencies, dynamic process, and thinking in terms of cause and effect.

The findings reported by Sturges and Maurer (22), while providing some insight into student perceptions of the learning difficulties associated with physiology, are limited in their generalizability. In particular, the data were gathered from a single student population at one institution, a 4-yr university in southeast Georgia. These results may not reflect the perceptions of students from across the nation and at other institution types. Much like in psychology, replication studies are critical in education research (13). Direct replications use the same methods as the original study, seeking to corroborate the findings of the initial study. Through direct replication studies, we can generalize results to different populations (19).

To our knowledge, there has been no formal replication of either the Sturges and Maurer (22) or Michael (14) studies. There is a need to replicate these studies with new populations to verify the generalizability of these results. We, therefore, used a direct replication approach, administering the same surveys from Sturges and Maurer and Michael to new populations at different institutions. Such a replication allows us to make more generalizable claims.

HOW WE TEACH | Generalizable Education Research

Physiology is hard: a replication study of students’ perceived learning difficulties

Tara Slominski, Sarah Grindberg, and Jennifer Momsen
Department of Biological Science, North Dakota State University, Fargo, North Dakota

Submitted 2 March 2018; accepted in final form 4 February 2019

Slominski T, Grindberg S, Momsen J. Physiology is hard: a replication study of students’ perceived learning difficulties. Adv Physiol Educ 43: 121–127, 2019; doi:10.1152/advan.00040.2018.—

Human Anatomy and Physiology (HAP) has long been recognized as a difficult course. A 2007 study (Michael J. Adv Physiol Educ 31: 34–40, 2007) sought to better understand this difficulty by asking faculty for their perceptions of why students struggle to learn in HAP. Later research built on these findings by investigating why students find physiology difficult (Sturges D, Maurer T. Internet J Allied Health Sci Pract 11: 1–10, 2013). However, without replication, these findings are limited in their generalizability. There is a need in physiology education research to replicate studies like these across different institutions to support generalizations. We, therefore, replicated both of these studies by collecting survey responses from 466 students at 4 different institutions and 17 instructors at 15 different institutions. We found that students in our study identified similar factors as the students surveyed in the original study. Students most strongly agreed with items that attributed the difficulty of HAP to the nature of the discipline, as opposed to the way physiology is taught or the way students approach learning it. Faculty in our sample, like those in the original study by Michael, agreed most strongly with items that attributed physiology’s difficulty to discipline specific factors. Our data reinforce the results of Sturges and Maurer and Michael. We can more confidently claim that HAP students and faculty believe the difficulty in learning physiology is the result of inherent features of the discipline itself and not factors related to instruction or the students themselves.

The findings reported by Sturges and Maurer (22), while providing some insight into student perceptions of the learning difficulties associated with physiology, are limited in their generalizability. In particular, the data were gathered from a single student population at one institution, a 4-yr university in southeast Georgia. These results may not reflect the perceptions of students from across the nation and at other institution types. Much like in psychology, replication studies are critical in education research (13). Direct replications use the same methods as the original study, seeking to corroborate the findings of the initial study. Through direct replication studies, we can generalize results to different populations (19).

To our knowledge, there has been no formal replication of either the Sturges and Maurer (22) or Michael (14) studies. There is a need to replicate these studies with new populations to verify the generalizability of these results. We, therefore, used a direct replication approach, administering the same surveys from Sturges and Maurer and Michael to new populations at different institutions. Such a replication allows us to make more generalizable claims.
METHODS

To gather student perceptions of the difficulties of learning human physiology, we used an adaptation of the survey generated by Michael (14) (Table 1). Michael’s survey is organized to reflect a three-factor model for student difficulty: physiology as a discipline, how students learn physiology, and how physiology is taught. The survey consists of 18 items representing these 3 factors. These 18 items are presented in a Likert-type fashion, asking students to indicate their level of agreement with the given statement (strongly disagree, disagree, neutral, agree, strongly agree, etc.).

Sturges and Maurer (22) made minor changes to the Michael (14) survey to make it more appropriate for use with undergraduate students. For example, the Michael version used words and phrases like “simultaneously,” “reason causally,” and “teleological thinking.” Sturges and Maurer changed these items to use phrases like “at the same time,” “in terms of cause and effect,” and “thinking in terms of their purpose,” respectively. We also used these modifications in our survey. In the Sturges and Maurer version of the survey, the term “AP” was used, whereas Michael used the term “physiology.” In our survey, we used the term “physiology.” We also asked a series of demographic questions to better describe and compare survey participants (major, ethnicity, sex, ACT/SAT score, etc.). We used the same version of the survey for the instructor participants (not including student demographic questions).

Study sites for both the student survey and the faculty survey were solicited in Fall 2016 via personalized e-mails and three professional listservs (American Physiological Society, Human Anatomy and Physiology Society, and Society for the Advancement of Biology Education). These e-mails asked for faculty participants who would be willing to 1) administer the student survey in their HAP class and 2) complete the faculty survey.

Twenty-five instructors agreed to participate in our study. Participation included completing the faculty survey and administering the student survey in their HAP or physiology courses. We sought institutional review board (IRB) approval from each institution to administer the student survey, and we received IRB approval from our home institution to administrate the faculty survey. We received IRB approval to administrate the student survey in 15 of the 25 courses. Of the 15 courses, the instructors of 10 courses successfully administered the student survey in the last 2 wk of the Fall 2016 semester. We excluded data from courses that 1) were not considered an introductory HAP course, or 2) did not achieve a student survey response rate >40%. Student surveys were administered electronically through Qualtrics survey software (16).

Of those 25 instructors who agreed to participate, 17 completed the faculty survey between late October and early December of 2016. Faculty surveys were administered electronically through Qualtrics survey software. All participants in our study (both students and faculty) were 18 yr of age or older, in compliance with approved IRB protocol SM17069.

Analysis. Data for this study are composed of responses to Likert-scale items. There is a longstanding debate as to how Likert-scale data should be analyzed. Many researchers in numerous disciplines argue that Likert data can be mapped to an interval scale, and thus parametric measures are appropriate (3, 4). Conversely, a substantial body of literature contends that Likert data are inherently ordinal and should be analyzed using nonparametric measures to generate sound conclusions (2, 6, 12). In addition, previous studies have found that students’ interpretations of “agree” vs. “strongly agree” are not consistent; these categories represent different magnitudes of response across and within populations (1, 20).

While the debate on Likert-scale analysis is likely to continue for the foreseeable future, individuals on both sides of this argument agree on one important facet: individual Likert items should not be considered a “Likert scale” and, in most cases, should not be analyzed individually (3, 4, 9). Both Sturges and Maurer (22) and Michael (14) assumed an interval scale and compared means of individual Likert items. We opted not to do this, because of the existing body of literature on Likert-scale analysis. Instead, we chose to peruse a nonparametric approach and compare the rank order of the 18 individual items.

To compare responses from Sturges and Maurer (22) and Michael (14) with responses from our populations, we compared the rank order of the items. To facilitate this ranking, we collapsed the five-point Likert scale into a three-point scale consisting of “agree,” “neutral,” and “disagree.” For example, the “agree” category in the three-point scale encompasses all student responses that indicated either “strongly agree” or “agree.” As a result, all agreements are accounted for, but less emphasis is placed on the intensity of that agreement. In terms of our findings, students responding with “agree” can draw on similar experiences, arguments, or emotions as students who respond with “strongly agree,” so the implications of our data are not compromised. We then ranked each item in terms of percent agreement, that is, how much of the population agreed to each of the 18 items. Those items with the highest percent agreement were ranked as most important.

Table 1. Summary of survey items

| Item                                                                 | Likert Scale |
|----------------------------------------------------------------------|--------------|
| 1) What is it about the subject matter of physiology that makes it hard to learn? (Discipline) | S D N A SA   |
| 2) What is it about the way physiology is taught that makes it hard to learn? (Teaching) | S D N A SA   |
| 3) What is it about the way students attempt to learn physiology that makes it hard? (Students) | S D N A SA   |

See Supplemental Table S1 for list of survey items. SD, strongly disagree; D, disagree; N, neutral; A, agree; SA, strongly agree.

1 Supplemental Tables S1 (full difficulty survey) and S2 (full list of demographic survey responses) for this article are available online at the Advances in Physiology Education website.
Students in our sample reported a wide array of declared majors, with responses spanning from criminal justice and music performance to nursing and pharmacy. Although there was a great deal of variation across and within each institution, a large number of students at each institution reported being either a nursing or pre-nursing student (Fig. 1). Similarly, three institutions (B, C, and D) had a substantial number of exercise science majors take the survey.

Because different institutions offer different programs (e.g., Audiology and Speech Language Sciences is offered at institution C but not at institution B), we were cautious in how we compared students’ majors across institutions. For example, we could not compare perceptions of medical laboratory sciences students to the perceptions of nursing students, as there are only 7 medical laboratory sciences in our sample compared with more than 100 nursing or pre-nursing students. Also, medical laboratory sciences students only reside at one of the four institutions, so we could not be confident that any observed differences could be attributed to declared major alone.

Do student opinions align with the opinions gathered by Sturges and Maurer? There was a statistically significant lack of independence between the rankings provided by students in our pooled population and the rankings provided by Sturges and Maurer (22) (Table 4, \( t = 0.725, P < 0.001 \)). Students in both populations exhibited similar levels of agreement to the 18 items of the difficulty survey, and thus the resulting rankings of the items are similar. This finding was consistent when rankings were compared at the institutional level: rankings from each institution had a statistically significant lack of independence among themselves when compared with the pooled population ranking and the Sturges and Maurer population (Table 5).

Students in our population reported the most agreement to item 5 (“Physiology, like other life sciences, seems to encourage thinking about things in terms of their purpose”) and item 3 (“Understanding physiology requires the ability to think in terms of cause and effect”) (Table 4). The Sturges and Maurer population also reported the strongest agreement to these items (Table 4).

Do student opinions of difficulty vary across intuitions? Students at all four institutions reported similar levels of agreement on all survey items pertaining to the discipline (Table 4). A Kruskal-Wallis test with a Bonferroni correction (\( \alpha \text{adjusted} = 0.0028 \)) showed a statistically significant difference in reported agreement for the following seven items: items 10, 11, 12, 13, 15, 16, and 18 (Table 4). However, students collectively reported little agreement to these items, with the highest of these items ranked as the seventh most impactful factor.
contributing to physiology’s difficulty. Students at all institutions reported similar levels of agreement for the 11 other items on the survey. Students identified items 3 and 5 as the top two sources of difficulty at institutions A, B, and C. At institution D, students reported the most agreement with items 2, 3, and 5.

Do faculty opinions align with the opinions gathered by Michael? There was a statistically significant lack of independence between the rankings provided by faculty in our population and the rankings provided by faculty in the Michael (14) study (Table 6, $\tau = 0.588, P < 0.01$). Faculty in both populations reported similar levels of agreement to the 18 items of the difficulty survey, and thus the resulting rankings of the difficulty items are similar.

The faculty in our survey reported the strongest agreement to item 3 (“Understanding physiology requires the ability to think in terms of cause and effect”), item 4 (“Understanding physiology requires at least some limited ability to think about dynamic systems”), and item 14 (“Students believe that ‘learning’ is the same thing as ‘memorizing’”). All 17 faculty participants in our study indicated they either “agreed” or “strongly agreed” with these items, resulting in a mean agreement score of 1.0 for all three items. The faculty in Michael’s population also reported the strongest level of agreement to these three items (Table 6).

**DISCUSSION**

We conducted a direct replication study to establish the generalizability of the findings of Sturges and Maurer (22) and Michael (14). Our results, coupled with their findings, docu-

---

**Table 4. Student ranking of survey items**

| Item | Factor | Rank |
|------|--------|------|
| 5    | Discipline     | Pooled | Institution A | Institution B | Institution C | Institution D | S&M |
| 3    | Discipline     | 1      | 2             | 2             | 1             | 1             | 1   |
| 2    | Discipline     | 2      | 1             | 1             | 2             | 3             | 2   |
| 4    | Discipline     | 3      | 6             | 3             | 2             | 2             | 5   |
| 7    | Discipline     | 4      | 4             | 4             | 4             | 4             | 3   |
| 1    | Discipline     | 5      | 5             | 5             | 5             | 6             | 6   |
| 16*  | Student        | 6      | 13            | 6             | 11            | 5             | 11  |
| 14   | Student        | 7      | 7             | 7             | 6             | 7             | 8   |
| 8    | Teaching       | 8      | 8             | 8             | 9             | 8             | 4   |
| 6    | Discipline     | 9      | 7             | 10            | 13            | 9             | 12  |
| 15*  | Student        | 10     | 10            | 9             | 17            | 12            | 16  |
| 18*  | Student        | 11     | 9             | 11            | 10            | 11            | 7   |
| 11*  | Teaching       | 12     | 13            | 12            | 12            | 9             | 9   |
| 13*  | Teaching       | 13     | 12            | 13            | 8             | 16            | 13  |
| 17   | Student        | 14     | 10            | 16            | 7             | 15            | 10  |
| 12*  | Teaching       | 15     | 13            | 14            | 16            | 13            | 14  |
| 10*  | Teaching       | 16     | 16            | 17            | 14            | 14            | 15  |
| 9    | Teaching       | 18     | 16            | 18            | 18            | 18            | 18  |

See Supplemental Table S1 for list of survey items. S&M, Sturges and Maurer (22). *Results are from Kruskal-Wallis rank sum test by four institutions (with Bonferroni correction for 18 groups = 0.0028).
ment a general trend in physiology education, namely that it is the nature of the discipline itself, and not instruction or student learning approaches that make physiology so challenging to learn. Our replication, using the original surveys, gathered responses from students and instructors from across the nation, asking them what makes physiology difficult to learn. Students and faculty agreed that the discipline itself is challenging, and this may be due to the complexity of causal reasoning and the approachability of teleological reasoning. We also find evidence to support the claim from Michael that faculty also attribute students' difficulty to learning physiology to their tendency to equate learning with memorizing.

Replicating earlier studies. The studies conducted by Sturges and Maurer (22) and Michael (14) provide valuable insight into student difficulties in HAP, but are limited to a single population. Direct replication studies like the present study help to establish generalizations across the broader HAP community and are essential in education research. For example, the demographics of HAP populations can vary greatly across the United States (as demonstrated by Fig. 1). These demographic differences could result in different skills, motivations, and approaches to learning, and thus result in different sources of difficulty in HAP classrooms. Furthermore, no study has sampled students and instructors simultaneously, such that instructors and students responded to the survey with the same course in mind. Thus our study goals included replicating both the Sturges and Maurer and Michael studies across institutions in the U.S. to support the generalizability of their findings.

Our sampling efforts used listservs and direct solicitation of faculty, resulting in the recruitment of four HAP populations that had a number of distinct differences from the original population surveyed by Sturges and Maurer (22). First, freshmen only made up 1% of the population of Sturges and Maurer, but often comprised more than one-third of our HAP population (Table 3). This is an important consideration, because freshman- and sophomore-level students often have a more novice-like approach to learning new material, relying on extensive memorization and using less effective and/or alternative reasoning pathways in response to contextual surface features (5, 7). These novice-like learning tendencies, in conjunction with limited experience learning in undergraduate science classrooms, could lead a student to attribute physiology’s difficulty to teaching practices or studying approaches, as opposed to the discipline itself.

Second, the four institutions sampled in our study also showed variation in the majors represented and the proportions of those majors at each institution (Fig. 1). For example, Exercise Science students were present at all institutions, but made up 13% of the population at institution B and 38% of the population at institution D. Institution B also contained a significant number of Pharmacy/Prepharmacy students (e.g., 21%), whereas other institutions (i.e., A and C) contained no Pharmacy/Prepharmacy students. Declared major could impact student prior knowledge, motivations, and performance goals and, by extension, their learning approaches in a HAP course. Students with different majors could experience different challenges in the HAP learning environment.

Despite such demographic variability, we found significant agreement about the difficulties of physiology across our study populations and institutions. Students identify the discipline as the primary driver of difficulty when learning physiology, and our findings thus reinforce the results of Sturges and Maurer (22). From this collective body of work, we can now state with some confidence that HAP students struggle with the discipline, particularly using causal reasoning effectively, often resorting to a less productive, teleological approach. These challenges persist across student populations, instructional designs, and institutional settings. HAP educators and education researchers should consider the unique features of the discipline and the ways students think about the discipline, in addition to more traditional educational variables (e.g., instructional strategies or student engagement) when striving to reduce student difficulties in learning physiology.

**Table 5. Student ranking correlations across institutions**

| Institution | Pooled | Institution A | Institution B | Institution C | Institution D | S&M |
|-------------|--------|---------------|---------------|---------------|---------------|-----|
| Pooled      | 0.763  | 0.935         | 0.669         | 0.852         | 0.725         |     |
| Institution A | 0.723  | 0.678         | 0.691         | 0.723         |               |     |
| Institution B | 0.603  | 0.813         | 0.660         |               |               |     |
| Institution C | 0.625  | 0.761         |               |               |               |     |
| Institution D | 0.695  |               |               |               |               |     |

Kendall rank correlation coefficient is shown. S&M, Sturges and Maurer (22). All P values < 0.001.

**Table 6. Faculty perception ranking compared with Michael**

| Item   | Factor | Slominski Percent | Agreement | Slominski Rank | Michael Rank |
|--------|--------|------------------|-----------|----------------|--------------|
| 3      | Discipline | 1.000            | 1         | 1              |              |
| 4      | Discipline | 1.000            | 1         | 3              |              |
| 14     | Students   | 1.000            | 1         | 2              |              |
| 1      | Discipline | 0.882            | 4         | 8              |              |
| 2      | Discipline | 0.882            | 4         | 5              |              |
| 5      | Discipline | 0.765            | 6         | 15             |              |
| 15     | Students   | 0.765            | 6         | 6              |              |
| 16     | Students   | 0.765            | 6         | 6              |              |
| 6      | Discipline | 0.647            | 9         | 4              |              |
| 13     | Teaching   | 0.588            | 10        | 14             |              |
| 7      | Discipline | 0.529            | 11        | 11             |              |
| 18     | Students   | 0.529            | 11        | 9              |              |
| 17     | Students   | 0.471            | 13        | 18             |              |
| 10     | Teaching   | 0.412            | 14        | 16             |              |
| 9      | Teaching   | 0.353            | 15        | 11             |              |
| 8      | Discipline | 0.294            | 16        | 10             |              |
| 11     | Teaching   | 0.294            | 16        | 13             |              |
| 12     | Teaching   | 0.059            | 18        | 17             |              |

See Supplemental Table S1 for list of survey items. Slominski values are from this study. Michael values are from Ref. 14.
occurred as opposed to focusing on how the actual mechanisms drive the physiological events (24). In contrast, causal reasoning requires students to reason more dynamically about the structures, behaviors (or mechanisms), and functions of a complex system. This type of reasoning requires thinking across different biological levels or scales within the organism. Students need to be able to understand how the dynamic interactions that occur within these levels give rise to functions at higher levels (11). The results of this research demonstrate that students are aware of both of these reasoning approaches, but they struggle to move beyond teleological reasoning strategies when faced with physiological phenomena.

Faculty perceptions. The faculty responses in our study mirrored the responses gathered by Michael in 2007 (14) and again contribute to the generalizability of our findings. Every faculty member surveyed agreed that students struggle to learn physiology because of 1) students’ causal reasoning abilities (item 3); 2) students’ systems thinking abilities (item 4); and 3) students’ belief that memorizing is the same as learning (item 14). While faculty and students surveyed in this research agreed on the role of causal reasoning and, to some extent, systems thinking skills in learning HAP, faculty also identified a learning strategy (memorization) as a source of difficulty in HAP classrooms.

A student in an HAP course has 12 or more years of learning experience. Memorization is a common learning strategy used in middle and high school, and it is likely they learned to rely on memorization as a means to study. From physics education research, we know that students’ tendency to memorize often persists into their undergraduate career: struggling undergraduate students report trying to approach learning physics by compartmentalizing and memorizing individual pieces of information (4). Extending these findings to the HAP classroom and our study, we hypothesize that struggling HAP students may fail to recognize that learning is not the same as memorizing. Students may conflate the ability to recall a number of basic facts and functions of a given organ or tissue with confidence that they have learned and understand the material. However, the student has not practiced reasoning causally; thus they may underperform their expectations on physiology-related questions and may resort to teleological reasoning strategies. Students may not report their learning strategies (e.g., memorization) as a contributor to learning HAP because they have limited experience with other study or learning strategies.

Implications for teaching. Starting in elementary school, students’ develop an inclination to answer “why” instead of “how.” A study by Southerland et al. (21) found primary and secondary students offered “why” responses, despite being asked “how” questions. Even when explicitly asked for a mechanistic understanding of biological phenomenon, students responded with a teleological rational for the occurrence of the phenomenon.

We found that students at multiple institutions agree that physiology encourages teleological reasoning. Because students have identified this inclination as a challenge to their own learning, we would urge instructors to take note of how students use teleological reasoning in their own classroom and explicitly and repeatedly address this issue with their students. We encourage instructors to provide examples to students that demonstrate how a teleological reasoning approach could result in a faulty or incomplete understanding of a physiological event. By training students to recognize how and when they use teleological reasoning in HAP classrooms, students may become careful and deliberate in the language they use when explaining the mechanisms that drive physiological phenomena. Furthermore, if instruction explicitly targets mechanisms spanning multiple organizational levels and/or multiple body systems, students may be more equipped to move beyond teleological explanations and reason about physiology with a more causal, mechanistic approach.

It is interesting to note that students reported that most teaching factors were not a major source of difficulty. For example, students did not report high agreement to the items pertaining to high memorization demands (item 11) and the high degree of jargon (item 12). However, there was disagreement across institutions on four of the six teaching factors: item 10, “Teachers do a poor job of defining and communicating our learning objectives”; item 11, “Teachers expect too many memorized facts and too little understanding at the same time”; item 12, “Teachers and authors use language imprecisely, use too much jargon, and use too many acronyms, all to the detriment of learning”; and item 13, “Teachers talk too much and students talk too little.” This variation across institution responses likely results from variation in instructional practices. For example, perhaps some instructors make good use of clear and measurable learning objectives, and thus students in those courses would disagree with item 10, whereas students at other institutions (where learning objectives are not utilized to the same degree) respond with more agreement. Whereas there was variation across these items (and items 15, 16, and 18), students did not find these factors as impactful as the inherent features of the discipline (items 1–5 and 7), of which there were no observed differences across institutions.

Implications for research. In our study, faculty and students agreed on the most influential sources of difficulty when learning physiology: “Understanding physiology requires the ability to think in terms of cause and effect, and physiology, like other life sciences, seems to encourage thinking about things in terms of their purpose.” We advocate for more qualitative investigations into this tendency to use teleological reasoning in physiology (as opposed to causal reasoning). Developing a better understanding of when and how students employ teleological reasoning will better support instruction that mitigates this detrimental tendency.

The findings of our replication study, while not new or flashy, establish the generalizability of these claims about student difficulties when learning physiology and provide valuable insight into the difficulties students experience. Students most strongly agreed with items that attributed the difficulty to the nature of the discipline, and not the way HAP is taught or how students approach learning. Faculty in our sample also confirmed the original study by Michael (14), agreeing most strongly to items that attribute difficulty to the nature of HAP as a discipline. Faculty also highly attributed student difficulty to memorization strategies, whereas students did not report this to be a considerable source of their own difficulty.

However, it is important to note that our data reflect student and faculty perceptions of difficulty and not the difficulties realized during instruction. It is possible that students are not consciously aware of (or able to put into words) all of the sources of difficulties they experience when learning in the
HAP classroom. Future research should focus on unpacking student difficulties in real time through more qualitative means.

Conclusion. Our replication research establishes the generalizability of the claims made by Michael (14) and Sturges and Maurer (22) by surveying students and instructors from different institutions across the United States. Our study population represents students enrolled in both large and small courses, with a wide range of career goals. The results of this research reinforce Michael and Sturges and Maurer in identifying the discipline of physiology as inherently difficult for students. Students and instructors both identify challenges associated with casual and teleological reasoning in physiology. This body of research underscores a need for qualitative and quantitative research that explores how students reason in HAP.

ACKNOWLEDGMENTS

We thank the students and instructors who participated in our study. Without their participation, this study could not have been conducted.

GRANTS

This study was supported by the North Dakota State University Graduate School, and the North Dakota Experimental Program to Stimulate Competitive Research Infrastructure Improvement Program-Doctoral Dissertation Assistance.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

T.S. performed experiments; T.S. and S.G. analyzed data; T.S., S.G., and J.M. interpreted results of experiments; T.S. prepared figures; T.S. drafted manuscript; T.S., S.G., and J.M. edited and revised manuscript; T.S., S.G., and J.M. approved final version of manuscript.

REFERENCES

1. Adams WK, Perkins KK, Podolefsky NS, Dubson M, Finkelstein ND, Wieman CE. New instrument for measuring student beliefs about physics and learning physics: the Colorado Learning Attitudes about Science Survey. Phys Rev Spec Top Phys Educ Res 2: 010101, 2006. doi:10.1103/PhysRevSTPER.2.010101.
2. Bishop FA, Herron RL. Use and misuse of the Likert item responses and other ordinal measures. Int J Exerc Sci 8: 297–302, 2015.
3. Caritô J, Perla R. Resolving the 50-year debate around using and misusing Likert scales. Med Educ: 1150–1152, 2008. doi:10.1111/j.1365-2923.2008.03172.x.
4. Caritô J, Perla RJ. Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antitodes. J Soc Sci 3: 106–116, 2007.
5. Chi MT, Feltoich PJ, Glaser R. Categorization and representation of physics problems by experts and novices. Cogn Sci 5: 121–152, 1981. doi:10.1207/s15516709cog0502_2.
6. Clason DL, Dormody TJ. Analyzing data measured by individual Likert-type items. J Agric Educ: 35: 31–35, 1994. doi:10.5032/jae.1994.04031.
7. diSessa AA. Knowledge in pieces. In: Constructivism in the Computer Age, edited by Forman G, Pufall PB. Hillsdale, NJ: Lawrence Erlbaum, 1988, p. 49–70.
8. Hammer D, Elby A. Tapping epistemological resources for learning physics. J Learn Sci 12: 53–90, 2003. doi:10.1207/s15327809JLS1201_3.
9. Harpe SE. How to analyze Likert and other rating scale data. Curr Pharm Teach Learn 7: 836–850, 2015. doi:10.1016/j.cplt.2015.08.001.
10. Harris DE, Hannum L, Gupta S. Contributing factors to student success in anatomy & physiology: lower outside workload & better preparation. Am Biol Teach 66: 168–175, 2004. doi:10.2307/4451650.
11. Hmeleo-Silver CE, Marathe S, Liu L. Fish swim, rocks sit, and lungs breathe: expert-novice understanding of complex systems. J Learn Sci 16: 307–331, 2007. doi:10.1080/10508400701413401.
12. Jamieson S. Likert scales: how to (ab)use them. Med Educ 38: 1217–1218, 2004. doi:10.1111/j.1365-2929.2004.02012.x.
13. Makel MC, Plucker JA. Facts are more important than novelty: replication in the education sciences. Educ Res 43: 304–316, 2014. doi:10.3102/0013189X1455513.
14. Michael J. What makes physiology hard for students to learn? Results of a faculty survey. Adv Physiol Educ 31: 34–40, 2007. doi:10.1152/advan.00057.2006.
15. Michael JA. Students’ misconceptions about perceived physiological responses. Am J Physiol 274: S90–S98, 1998. doi:10.1152/advances.1998.274.6.S90.
16. Qualtrics. Qualtrics (Online). https://www.qualtrics.com [31 Oct 2018].
17. R Core Team. The R Project for Statistical Computing (Online). http://www.R-project.org/ [18 Feb 2018].
18. Richardson DR. A survey of students’ notions of body function as teleologic or mechanistic. Am J Physiol 258: S8–S10, 1990. doi:10.1152/advances.1990.258.6.S8.
19. Schmidt S. Shall we really do it again? The powerful concept of replication is neglected in the social sciences. Rev Gen Psychol 13: 90–100, 2009. doi:10.1037/a0015108.
20. Semsar K, Knight JK, Birol G, Smith MK. New instrument for measuring student beliefs about physics. J Soc Sci 38: 528–348, 2001. doi:10.1002/jssc.1013.
21. Southerland SA, Abrams E, Cummins CL, Anzelmo J. Understanding students’ explanations of biological phenomena: conceptual frameworks or p-prims? Sci Educ 85: 268–278, 2011. doi:10.1177/0036669510371838.
22. Sturges D, Maurer T. Allied health students’ perceptions of class difficulty: the case of undergraduate human anatomy and physiology. Internet J Allied Health Sci Pract 11: 1–10, 2013.
23. Sturges D, Maurer TW, Allen D, Gatch DB, Shankar P. Academic performance in human anatomy and physiology classes: a 2-yr study of academic motivation and grade expectation. Adv Physiol Educ 40: 26–31, 2016. doi:10.1152/advances.00091.2015.
24. Tamir P, Zohar A. Anthropomorphism and teleology in reasoning about biological phenomena. Sci Educ 75: 57–67, 1991. doi:10.1002/see.3730750106.
25. U.S. Department of Commerce Economics and Statistics Administration. Census Regions and Divisions of the United States. https://www.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf [18 Feb 2018].
26. The Carnegie Classification of Institutions of Higher Education. About the Carnegie Classification (Online). http://carnegieclassifications.ui.edu/ [https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf [18 Feb 2018].