

Article

Linguistic Analysis of Project Ownership for Undergraduate Research Experiences

D. I. Hanauer,^{*†} J. Frederick,[‡] B. Fotinakes,^{*} and S. A. Strobel[§]

^{*}Department of English, Indiana University of Pennsylvania, Indiana, PA 15705; [†]PHIRE Program, University of Pittsburgh, Pittsburgh, PA 15260; [‡]Center for Scientific Teaching at Yale and [§]Department of Molecular Biophysics & Biochemistry, Yale University, New Haven, CT 06520

Submitted April 11, 2012; Revised September 10, 2012; Accepted September 10, 2012
Monitoring Editor: Hannah Sevian

We used computational linguistic and content analyses to explore the concept of project ownership for undergraduate research. We used linguistic analysis of student interview data to develop a quantitative methodology for assessing project ownership and applied this method to measure degrees of project ownership expressed by students in relation to different types of educational research experiences. The results of the study suggest that the design of a research experience significantly influences the degree of project ownership expressed by students when they describe those experiences. The analysis identified both positive and negative aspects of project ownership and provided a working definition for how a student experiences his or her research opportunity. These elements suggest several features that could be incorporated into an undergraduate research experience to foster a student's sense of project ownership.

INTRODUCTION

Undergraduate research experiences are an important factor in attracting students to the sciences and helping them build a foundation for a scientific career (Nagda *et al.*, 1998; Seymour *et al.*, 2004; Villarejo *et al.*, 2008). Undergraduate research experiences can offer students an opportunity to personalize and own science, ultimately increasing the chances of scientific careers for such students. A study of the learning cycles of Nobel laureates identified the progression of events for several illustrious scientific careers that began with “nurturing” a sense of ownership through a passionate commitment to scientific inquiry (Stewart, 1994). In a discussion of factors that contribute to effective scientific inquiry experiences, Graham Hatfull, the lead scientist of a bacteriophage isolation and genomic description educational program, stated that “Isolating phages from environmental samples proves encouraging

and exciting for students, giving them a sense of accomplishment and a feeling of ownership of their new virus. These feelings help to motivate them to continue analyzing their new phage” (Hatfull, 2010, p. 244). This statement suggests that research experiences are especially motivating and encouraging when students gain a “feeling of ownership” over their research projects.

Despite recognition of the relationship between positive research experiences and project ownership, no studies explore this connection for undergraduate science, technology, engineering, and mathematics (STEM) research. We undertook this question with the goal of addressing four objectives. First, our study defines and offers theoretical clarification of the concept of project ownership and its relevance to science education. Second, we describe a quantitative methodology for measuring project ownership by using interview data. Third, we used linguistic analysis to characterize the elements of project ownership expressed in relation to different types of educational research experiences. Finally, we offer suggestions based on the linguistic analysis for designing science education programs that promote project ownership.

DOI: 10.1187/cbe.12-04-0043

Address correspondence to: Scott A. Strobel (scott.strobel@yale.edu), David I. Hanauer (Hanauer@iup.edu).

© 2012 D. I. Hanauer *et al.* CBE—Life Sciences Education © 2012 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution-Noncommercial-Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®” and “The American Society for Cell Biology®” are registered trademarks of The American Society for Cell Biology.

Defining Project Ownership

As a theoretical construct, the concept of project ownership integrates several auxiliary terms related to the concepts of choice, control, personal responsibility, self-determination,

commitment, engagement, emotional involvement, and identification (Kennedy, 1994; Chung *et al.*, 1998; Downie and Moore, 1998; Mason *et al.*, 2004; Nail, 2007; Wiley, 2009). To explicate the concept of ownership of learning from these adjacent terms, Wiley has proposed three conceptual definitions (Wiley, 2009, pp. 25–33):

1. *Ownership as responsibility* relates to the student's role in his or her educational activity. Students have ownership over their learning process when they take personal responsibility for the choice of study and the outcome of their work.
2. *Ownership as buy-in* refers to the commitment, engagement and loyalty that a student has towards his or her educational program or project. Students expressing ownership have "bought into" the importance of the educational project.
3. *Ownership as self-identification* refers to a sense of belonging and identification with the educational program or project. A student with ownership has a deep personal connection and sees the project as an extension of him/herself.

Combining Wiley's ideas leads to a more nuanced definition of ownership. We assume that development of project ownership is not just an individual student characteristic, but one that results from a more complex interaction between the student and the educational environment. Accordingly, features of an educational program may increase or decrease the degree of project ownership that a student experiences. Milner-Bolotin defines student educational ownership along a spectrum and notes that student-centered educational approaches should elicit much higher levels of ownership (Milner-Bolotin, 2001). Importantly, some evidence suggests that engagement, one aspect of project ownership, is a factor that retains students in the sciences and reduces the number of declared science majors who leave to study other fields after taking introductory science courses (Dickinson and Flick, 1998; Hake, 1998; National Research Council, 2003; Cech and Kennedy, 2005; Jones *et al.*, 2010; Thiry *et al.*, 2011).

Despite agreement that research experiences add value to a science education and can help promote student long-term interest in science, these experiences vary widely. Laboratory training ranges from rote experiments with known outcomes, amounting to little more than demonstration exercises, to a research experience in a faculty member's laboratory. Even capstone experiences vary significantly in the extent of autonomy and intellectual control provided to the undergraduate student (Wenning, 2005, 2007; Hanauer *et al.*, 2009). Characteristics of the educational research experience, such as control over decision making, personal agency, and amount of responsibility, are likely to be reflected in the degree to which a student experiences and develops project ownership. This is the connection that we set out to explore.

Measuring Project Ownership

One approach to evaluating text is through content analysis (Tausczik and Pennebaker, 2010), a qualitative research method in which verbal data are assigned to categories relevant to the research topic (Stemler, 2001). Linguistics researchers develop reliable content categories via established

methods and use the content categories to code a collection of texts. The percentage of statements categorized this way can yield insights about the text, such as common features of student descriptions of their undergraduate research experiences. We used this approach to develop a definition of ownership as expressed by undergraduate students engaged in a research project.

We also wanted to measure degrees of project ownership by using a quantitative approach. This involved addressing two methodological problems. The first issue involved finding suitable comparative research experiences that differ in relation to choice, control, agency, and responsibility. The second problem concerned finding a way to measure a student's sense of project ownership. Project ownership is a complex term that integrates attitudes and emotions related to a student's personal connection to an educational experience. It involves a response to an educational environment and is not a disembodied feature of the student. We anticipated that project ownership would be evident in a student's description or discussion of his or her scientific project. In other words, students who feel a sense of ownership will reveal a sense of personal connection through the attitudinal and emotive language they use to discuss their work. This expression provides a measure of how ownership is developed and conceptualized by students.

Recent developments in the field of computational linguistics offer powerful tools for analyzing speech and texts. Computational linguistics—the analysis of language using computer-based methodologies—allows analysts to go beyond intuitive, comprehension-based approaches to text analysis, revealing systematic language use patterns that reflect a range of attitudinal and emotive stances (Hunston, 2002). The Linguistic Inquiry and Word Count (LIWC) program has been used extensively to analyze issues of attitude, stance, and emotive responses (Hunston, 2002; Alpers *et al.*, 2005; Kahn *et al.*, 2007; Tausczik and Pennebaker, 2010). The LIWC program calculates frequencies for linguistic categories (such as pronouns or verbs) and semantic categories (such as emotive words, words designating social processes, or cognitive words) in a given collection of texts. As a basic principle, researchers who use word counts and content analysis hypothesize that what someone is thinking and feeling is reflected in the language he or she uses. The LIWC program offers a way of systematically assessing that information.

Several different categories of information are relevant to project ownership. If a student is an active agent, in control and personally connected to a project, we expected that this sense of ownership would be reflected in increased use of first-person pronouns (I, me, my, and mine). Additional features of project ownership include emotional engagement, identification, and commitment. In linguistic terms, these features would be reflected in the semantic category of words that denote an emotional response. Students who are emotionally engaged with a program would be expected to express a heightened sense of emotion through increased use of emotive words. But in contrast with the emotional lexicon, when students discuss the content of scientific inquiry experiences, they would express cognitive processes and insights related to their projects. We would not expect language in cognitive categories to fluctuate significantly with the degree of ownership.

METHODS

Our experimental approach incorporated content analysis and analytical techniques for language processing (the LIWC program) to define and measure ownership as expressed in student interviews. We identified elements of project ownership by using content analysis and employed computational linguistic analysis to measure the degree of project ownership for groups of students in different science educational experiences. For the question dealing with differing degrees of project ownership, we compared student descriptions of different types of research experiences. A comparison between the content analysis and the linguistic analysis provides a means for assessing the compatibility of the linguistics findings with content statements concerning degrees of project ownership.

Participants

Three groups of students were identified based on their self-selected participation in different types of undergraduate laboratory experiences. We were interested in comparing a research experience with ownership as an explicit design feature to the more variable category of undergraduate research experiences in faculty labs around campus. As a contrast to the two groups of students participating in a capstone undergraduate experience, we also interviewed a third group of students enrolled in a standard undergraduate laboratory course. The three groups were identified as follows:

1. Group A: These students collectively participated in an inquiry-based learning course, Rainforest Expedition and Laboratory. This course involved the development of individual research questions relevant to diverse microbial samples that the students isolated from materials collected during a two-week field research trip. The course was structured to provide a very high degree of autonomy. Students developed their own research questions and directed their projects from field collection to assay development and chemical analysis.
2. Group B: These students were enrolled in an independent research course and did research in various faculty laboratories throughout the university. Students selected the laboratory in which they worked. A post-doctoral associate or senior graduate student mentor typically defined the actual research project and experimental design.
3. Group C: These students were enrolled in a standard laboratory course intended for science majors and premed students. These students were given an established protocol with a defined outcome. Other than enrolling in the course, students had little or no choice of topic and no autonomy in directing the project.

Group A students enrolled in the course, but were not selected for enrollment based upon previous research experience. Students in group B independently identified a lab with a research subject of interest to them and experienced a variable level of mentoring and autonomy available in each of the individual labs. Group C involved students enrolled in one of

two laboratory courses, Organic Chemistry Laboratory and Biochemistry Laboratory. Students interviewed for this study were sent targeted invitations inviting their participation in a brief interview protocol. Students in each of the different groups studied at Yale University; were drawn from a mixture of college sophomores, juniors, and seniors; and were mainly science or premed majors. Gender and ethnic balance in the groups was approximately representative of the overall university population, with roughly equal numbers of male and female students and two or fewer African-American or Hispanic participants in each group. A total of 49 students participated in this study: 17 in group A, 16 in group B, and 16 in group C.

Interview Tool

All students were interviewed by the same person, who used identical questions designed to invoke a general description of each student's laboratory experience. The interview protocol consisted of the following questions:

1. Tell me about the idea for your research project.
2. Can you give a little background about why this project is interesting in general?
3. What were some especially memorable moments during this research experience?
4. What were some of the challenges, and how did they get resolved?
5. What are some likely next steps for continuing the project, and how do you envision this happening?
6. Extrapolate out three years—where could this project go?

All answers were recorded, transcribed, and subjected to computational linguistic and content analyses.

Data Analysis

Two types of data analysis were conducted on the transcribed interviews, content analysis and computational linguistic analysis. The content analysis was conducted in three stages, following accepted guidelines (Neuendorf, 2002; Krippendorff, 2004). The initial stage consisted of constructing a coding system. Two applied linguistics researchers read the complete data set of interviews and proposed an initial set of codes for the interview data. After establishing agreement between researchers using standard practices, a 14-category coding system was defined. In the second stage of coding system development, reliability was established according to accepted guidelines (Cohen's kappa statistic, a measure of agreement among raters, was $K = 0.61$; $p < 0.01$; Landis and Koch, 1977). In the final stage of the content analysis, the entire interview data set was analyzed with this coding system, and the percentages of codes by group were calculated. The quantitative computational analysis was conducted using the LIWC program (Pennebaker *et al.*, 2007). Frequencies of first-person pronouns, emotive words, cognitive lexicon, and insight words were counted and analyzed using appropriate statistical methods. Details of the methods and analysis are provided in the Supplemental Material.

RESULTS

Content analysis of the interview data yielded a coding system that defines elements of project ownership. Interestingly, eight of the categories are positive indicators of project ownership (categories 1–8 below) and four are negative indicators (categories 9–12 below). The last two categories listed are neutral and pertain to question clarification or details of signing up for a research program (categories 13 and 14 below). This analysis provides an extensive definition of the elements of project ownership. The coding system categories are as follows:

1. Constructing connections between personal history and scientific inquiry: This category includes both statements and narratives that describe significant moments in a student's life that have shaped, influenced, and created the scientific work done in the laboratory and field-work. They involve a student bringing past experience in the form of personal stories and past educational experience into current and future research. An example statement is: "So I actually started last summer when I was in Ghana with some children in an orphanage and I had malaria twice and just the experience of having malaria and of being around all these children who experienced malaria at least once a year got me thinking about that."
2. Agency combined with mentorship: This category identifies moments when the student actively seeks advice, assistance, or direction from professors, teachers, and other students in order to overcome an issue or fulfill an aim in the student's research project. These moments represent joint contributions and knowledge-building between students and educators. An example statement is: "I was directed to J.M.G. who is a postdoc in M.'s lab, and just got in contact with him and he suggested this particular assay that I can do that would be easy and so that's sort of how the project came about."
3. Expressions of excitement toward scientific inquiry: This code shows emotional connections to the work the students are performing. The statements express positive emotional interaction relating to involvement in science. An example statement is: "I guess of course one of the coolest things is actually getting to go to the rainforest and actually do the research from start to finish. So that was really amazing."
4. Overcoming challenging moments in science: This category codes statements that address strategies for overcoming frustrating moments or problems encountered in research. Students discussed how they approached problems by adjusting their work or predicted how they could develop an entirely different approach. An example statement is: "So one of the challenges is that the way that I make the sample chambers that I put these particles in is not very consistent. . . . So one thing I'm trying to work on is to very consistently, very repeatably make these chambers for my samples. And it's a much more well-defined problem and something I just recently started to work on and it's coming up with a reproducible method for really systematically making chambers. And hopefully that will really cut down some variables."
5. Expressions of a sense of personal scientific achievement: This category describes a positive emotional expression upon achieving a specific goal. It captures a specific moment in a student's work. Students reference a specific finding or discovery, and this finding causes pride, happiness, or satisfaction for the student. An example statement is: "I guess what one of my high-est points was when I found a *Streptomyces* during the school year . . . I smelled it and it had that characteristic smell and I was like oh, yeah, I found one. I got really excited because they are, you know, it's cool stuff. So that was exciting."
6. Frustrating moments in the scientific process: This category describes situations in which students find difficulty in the procedures and nature of scientific inquiry. While these statements express frustration, they also reveal a student's moment of struggle with understanding the scientific processes. The majority of statements in this code relate to procedural frustrations with lab techniques or with other people in the lab. An example statement is: "So, yeah, just realizing like how frustrating PCR can be and like, yeah, experiencing that. I mean, you know, it's failed so many times and I never realized how . . ."
7. Scientific outcomes with real-world significance: This category identifies instances when students believe they have made a scientific discovery that may have the ability to affect the world outside the laboratory. These statements make a connection with a student's current work and its impact on humanity. An example statement is: "I really hope that one of the fungus [*sic*] does have like a product that it makes protein or whatever that helps women. And so it would be really great that there is like a cream or a pill so women won't have to always have the epidurals. . . ."
8. Expressions of understanding the unexpected aspects of science: This category codes moments when students realize that scientific inquiry can be unpredictable and this unpredictability can influence their research. These statements express the discovery that during scientific work, students cannot always know what to expect and that there is an uncertain aspect to doing science. An example statement is: "And you're like, well, you know, you probably are just not plugging it in right or something. But no, it's just like totally sometimes it's random. You do everything right and it just fails."
9. Prescribed scientific inquiry: This category describes instances when students encounter science by being told what to do, what results to expect, and/or procedures to follow with little or no student agency. An example statement is: "But because this is a lab class, and it's only so much time, our instructor did it for us, and then we just talked through to figure it out."
10. General statements of interest: This category identifies statements that express interest (but not excitement) in the science and reflect an appreciation for doing science as an academic subject without emotional connection. An example statement is: "The part we did last time was interesting because it's the first time that we got to see the two reactions, actually [inaudible] when the product was synthesized, and then what we're doing"

today is interesting because we're using techniques that can identify different products."

11. Indifference toward science: This category finds moments when students express apathy, passivity, or boredom toward their scientific work. Typically, these statements are short-phrase answers to questions concerning good or bad memorable moments in a student's work. An example statement is (in response to being asked if he/she had any memorable moments): "None really so far I don't think."
12. Descriptions of procedures: This category codes when students describe laboratory or field procedures. These are moments when students describe what they are doing and how they do it. An example statement is: "Okay, so Project One, we're looking at a protein [inaudible] which leaves DNA, and basically what we are doing, we went in and mutated the DNA using PCR mutagenesis."
13. Clarification of interview questions: These phrases occur when a student clarifies what is being asked or as linguistic fillers when the student thinks about how to answer a question. An example statement is: "The entire thing or our individual research project?"
14. Procedure of entering research experience: This code is a description of a procedure of how a student got to the position, such as, "I was in a class and heard the announcement. So then I applied..."

Table 1 summarizes the percentage of statements made by each research experience group for the 14 content analysis categories. All the statements made by participants were coded and categorized according to the coding system presented. The percentages were calculated in relation to the total number of statements made by each group.

Consistent with the various levels of autonomy and project development that were experienced by the three groups, the distribution of coding categories reveals some interesting differences in the content of their language usage. Students in group A had a higher percentage of statements in seven of the eight positive project ownership categories compared with both groups B and C. These statements suggest that they had a strong sense of project ownership, consistent with the formulation of the course. In contrast, students in group C exhibited higher frequencies of the four negative ownership indicators compared with both groups A and B (categories 9–12). Notably, group C students made no statements describing their experiences as involving agency (category 2) or personal connections (category 1). They also lacked statements about overcoming challenges or gaining unexpected understanding (categories 4 and 8), two critical elements of successful research.

Group B was generally situated between the category usage of group A and group C. These students expressed some elements of project ownership (such as agency) but not all (such as personal connection and emotional engagement). Interestingly, group B had the highest number of statements that deal with frustration with the scientific process (category 6). They made more than twice the number of statements concerning frustration than either of the other groups. This suggests that these students were conducting scientific research but may have lacked sufficient support from mentors or a community of peers.

Table 1. Percentage of statements by group and for each content analysis category

Code category	Group A	Group B	Group C
1. Constructing connections between personal history and scientific inquiry	4.5	0.4	0
2. Agency combined with mentorship	12.9	6.8	0
3. Expressions of excitement toward scientific inquiry	12.9	10.2	0.85
4. Overcoming challenging moments in science	3.3	2.4	0
5. Expressions of a sense of personal scientific achievement	6.2	3.4	0.85
6. Frustrating moments in the scientific process	7.3	16.3	7.6
7. Scientific outcomes with real-world significance	18	16.6	9.4
8. Expressions of understanding unexpected aspects of science	6.2	0.9	0
9. Prescribed scientific inquiry	0	1.9	11.9
10. General statements of interest	3.9	5.8	14.5
11. Indifference toward science	0	0.4	8.5
12. Descriptions of procedures	15.8	25.4	32.4
13. Statements of question clarification	8	9	14
14. Statements on signing up for research program	1	0.5	0

When describing their research experiences, group C students made twice as many statements describing the procedures of science (category 12) than members of group A. Conversely, group A made twice as many statements concerning the real-world significance of their scientific research (category 7) than group C. It seems that for group A, there was far greater understanding of the broader social value of their research, while group C students were more focused on the procedures. Furthermore, group A was the only group to make statements addressing the unexpected nature of science (category 8). This suggests that a student who has ownership over his or her research also faces, and ultimately understands, the uncertainty and unpredictability that characterizes scientific inquiry. Overall, the content analysis of the interviews of the three groups suggests a different understanding of their research experiences.

We also explored the degree of ownership expressed by the students, using a usage frequency of emotional and cognitive lexicon. The computational linguistic analysis of the four linguistic categories (first-person pronouns, emotional words, cognitive lexicon, and insight words) for the three scientific inquiry groups is summarized in Table 2. Group A had a higher usage frequency of first-person pronouns and emotional words. The frequency of cognitive lexicon and insight usage were very similar for all groups. Statistical analysis supported the differences between these groups.

Table 2. Means and SDs for three groups for four linguistic variables defining project ownership

Variable	Group A	Group B	Group C
Personal Pronouns	5.03 (1.33)	3.91 (1.12)	2.96 (1.53)
Emotional Words	4.3 (0.97)	3.88 (1.1)	3.37 (1.06)
Cognitive Lexicon	21.42 (1.5)	21.54 (2.2)	22.93 (2.82)
Insight Words	3.4 (1.14)	3.23 (0.94)	3.32 (1.58)

Tests of the homogeneity of variances were calculated in order to verify the appropriateness of the statistical procedure. This was necessitated by the small numbers of participants and the slight inequality of student numbers in the groups and as a precursor to the statistical analysis of the descriptive data. A Box's *M* test and Levene's test of the equality of variances were conducted. Both returned *p* values above 0.05, thus allowing the assumption of normality. For evaluation of the above analysis of the descriptive data, a one-way multivariate analysis of variance (MANOVA) was calculated using the student group as the independent variable and First-person Pronouns, Emotional Words, Cognitive Lexicon, and Insight Words as four dependent variables. Hotellings' T^2 multivariate generalization of the univariate *t* value was used. The MANOVA revealed a highly significant effect for the group variable (Hotellings' $T^2 = 0.80, p < 0.01$). Because a significant multivariate effect was found, univariate *F*-tests were calculated to determine which variables contributed to the overall difference. Significant differences were found for First-person Pronoun usage ($F(2, 46) = 10.33, p < 0.01$) and Emotional Words ($F(2, 46) = 3.62, p < 0.035$). The other two variables, Cognitive Lexicon and Insight Words, were not found to be significantly different.

To further understand the source of difference in the variable of Personal Pronouns and Emotional Words, we conducted post hoc comparisons of cell means, using the Scheffe method. Group A (rainforest group) was found to be significantly different than group C (standard lab group) for both Personal Pronoun usage and Emotional Words. No significant differences between groups A and B (independent research) were found on this test. However, on a less conservative, least significant differences post hoc test, group A was also found to be significantly different than group B on the variable of Personal Pronoun usage. This analysis also confirmed the significant differences between group A and C for both the Personal Pronoun and Emotional Word variables.

The statistical results support the conclusion that the three groups are significantly different based on the frequency of Personal Pronouns and Emotional Words. Group A students used first-person pronouns and emotional words with higher frequencies than did group C students. However, differences were also found between groups A and B for Personal Pronoun usage. These results suggest that there are differences in the levels of project ownership expressed in interview data concerning different types of research experiences. Higher levels of personal choice and agency in the research project design were reflected by increased first-person pronoun and emotional word use. This further demonstrates that the degree of project ownership felt by a student is reflected in the language the student uses to express that experience.

DISCUSSION

The results of this study suggest that features of an educational research experience do result in measurable differences in the way students describe these experiences. Significant differences were observed in both the content analysis and computational linguistic data. The content analysis has defined a set of elements that positively and negatively influence the degree of project ownership experienced by a student. Research experiences in which students felt agency, choice, and control elicited statements of excitement about doing scientific research, whereas experiences lacking these elements elicited statements of indifference toward science. The evidence from this study suggests that students develop and describe a sense of project ownership when research experiences are appropriately designed.

A second important finding is that there are measurable degrees of project ownership and these are systematically aligned to the types of education research students experience. Rather than considering project ownership a dichotomous variable that a student does or does not have, the data suggest that there are degrees of expressed ownership. Group B revealed agency in the content analysis data but did not express the degree of emotional commitment or personal connection of the group A participants. Our results emphasize that undergraduate research in general is better than a standard laboratory course for instilling project ownership and demonstrate that undergraduate research experiences are highly variable in design, incorporation of mentoring, and whether students are included within a scientific community.

An underlying question in this study concerned the validity of using computational linguistic data concerning first-person pronouns and emotional words as a way of measuring project ownership. In addressing this question, the same interview data were analyzed using two different linguistic analytical techniques: computational linguistic frequency counts and content analysis. The data presented in this study reveal a close correspondence between the frequency counts designating project ownership and the informational categories designating project ownership in relation to different research experiences. Group A expressed the highest levels of first-person pronouns and emotional words, and also had the highest levels of those categories relevant to project ownership (agency, personal connection, excitement, personal achievement, and a sense of scientific significance). This correspondence between the results suggests that our use of computational linguistic analysis was a valid measure of project ownership.

The interviews for this study were conducted in 2008. Enough time has passed that it is possible to ask whether the degree of ownership showed any correlation with retention in the sciences or persistence in a scientific career. We tracked student decisions for groups A and B, in which students were doing research as a capstone undergraduate experience. Of the group A students, 88% are now following a career and studies in science or medicine (18% to an MD program, 53% to a PhD or MD/PhD program, 18% to a master's program or lab work, 12% not doing science). Three-quarters (76%) continued to participate in research during their undergraduate education and after graduation. For group B students who reported future plans, 63% either continued to graduate

school in the sciences or went to a research-intensive medical school with a research/independent study requirement (25% to an MD program, 38% to a PhD or MD/PhD program, 6% to science consulting, and 31% indicated other/not doing science). The group of students with a higher level of project ownership also showed a higher level of persistence in science. Although these data derive from a small cohort of students, they hint that project ownership in an undergraduate research experience could encourage students toward longer-term career goals in the sciences. We did not track group C students because the courses we drew from were not capstone experiences like those of groups A and B.

Our results suggest that the type of educational program influences the way students describe their experiences and that intentionally designing educational programs to inculcate ownership results in students expressing increased levels of project ownership. We recognize that using three groups of self-selected students introduces the possibility that a participant's desire for project ownership motivated the group selection. We would expect a higher number of group A students to continue in science if their choice to participate in group A reflects a pre-existing interest. The inclusion of pre- and postdata on student education and career interests in future studies would address this limitation.

The content analysis identified both positive and negative expressions of project ownership that provide a working definition for how a student experiences a research opportunity. These elements suggest several features that could be incorporated into undergraduate research to foster student project ownership:

1. Facilitating personal agency: The undergraduate student should be allowed to make decisions concerning research questions and methods in consultation with more experienced researchers. This important feature is not typical of undergraduate research programs, in which students are often assigned to a small aspect of an existing project or asked to choose between several pre-designed mini-projects. Being able to make strategic decisions in a research project allows students to develop confidence as scientific scholars and to envision themselves as scientists. (Code categories: 2, Agency combined with mentorship; and 5, Expressions of a sense of personal scientific achievement)
2. Personal significance for the research project: Students should be given opportunities to pursue research questions in areas of personal interest, such as particular diseases or health issues. This might involve events from the student's personal or family life or a major social problem. (Code category: 1, Constructing connections between personal history and scientific inquiry)
3. Scientific value for the scientific inquiry: Research questions and findings from undergraduate research projects should be important to a broader scientific community. (Code category: 7, Scientific outcomes with real-world significance)
4. Social interaction and mentorship: Working closely with peers and having access to more experienced scientists is a crucial aspect of authentic scientific inquiry and may reduce feelings of frustration that may accompany novel research questions and procedures. Such mentoring is a balance between being overly prescriptive, which

inhibits creativity and agency, and being insufficiently supportive, which leads to uncertainty, frustration, and a sense of failure. (Code category: 2, Agency combined with mentorship)

5. Research that demands problem solving: Students with strong project ownership also exhibited more involvement in higher levels of problem solving and the ability and willingness to contend with problems that arose within the scientific inquiry process. Encouraging this cognitive investment builds intellectual challenge and deepens the personal significance of the work for the student. (Code category: 4, Overcoming challenging moments in science)

The components of project ownership defined by our linguistic analyses can be used to design undergraduate research programs in many settings with various levels of resources. Programs in which students experience strong project ownership make authentic science accessible to undergraduate students.

ACKNOWLEDGMENTS

D.I.H. was funded through a subaward (IUP RI log no.: 0910-028) from a Howard Hughes Medical Institute (HHMI) professorship award to Graham Hatfull. D.I.H. and S.A.S. were supported from a HHMI professorship award to S.A.S. J.F. was supported by an HHMI institutional award to Yale University and Dr. Robert Wyman.

REFERENCES

- Alpers GW, *et al.* (2005). Evaluations of computerized text analysis in an internet breast cancer support group. *Comput Hum Behav* 21, 361–376.
- Cech T, Kennedy D (2005). Editorial: doing more for Kate. *Science* 310, 1741.
- Chung H, Rodes P, Knapczyk D (1998). Using Web Conferencing to Promote Ownership in Distance Education Coursework. Paper presented at the Third Annual World Conference of the WWW, Internet, and Intranet, Orlando, FL, November 7–12, 1998.
- Dickinson VL, Flick LB (1998). Beating the system: course structure and student strategies in a traditional introductory undergraduate physics course. *Sch Sci Math* 5, 238–246.
- Downie M, Moore P (1998). Closing the gap: schools forge a bridge to community—in California: comprehensive services and realistic planning. *Perspect Educ Deafness* 16, 15–27.
- Hake R (1998). Interactive-engagement versus traditional methods: a six-thousand student survey of mechanics test data for introductory physics courses. *Am J Physics* 66, 64–74.
- Hanauer DI, Hatfull GF, Jacobs-Sera D (2009). *Active Assessment: Assessing Scientific Inquiry*, New York: Springer.
- Hatfull GF (2010). Bacteriophage research: gateway to learning science. *Microbe* 4, 243–250.
- Hunston S (2002). *Corpora in Applied Linguistics*, Cambridge, UK: Cambridge University Press.
- Jones MT, Barlow AE, Villarejo M (2010). The importance of undergraduate research to persistence in biology. *J Higher Educ* 81, 82–115.
- Kahn JH, Tobin RM, Massey AE, Anderson JA (2007). Measuring emotional expression with the Linguistic Inquiry and Word Count. *Am J Psychol* 120, 263–286.

- Kennedy M (1994). The Ownership Project: an experiment in student equity. *Soc Stud Rev* 33, 24–31.
- Krippendorff K (2004). *Content Analysis: An Introduction to a Methodology*, Thousand Oaks, CA: Sage.
- Landis JR, Koch GG (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174.
- Mason CY, McGahee-Kovac M, Johnson L (2004). How to help students lead their IEP meetings. *Teach Except Child* 36, 18–25.
- Milner-Bolotin M (2001). The effects of topic choice in project-based instruction on undergraduate physical science students' interest, ownership, and motivation. Unpublished doctoral dissertation, Austin: University of Texas at Austin.
- Nagda BA, Gregerman SR, Jonides J, von Hippel W, Lerner JS (1998). Undergraduate student-faculty research partnerships affect student retention. *Rev Higher Educ* 22, 55–72.
- Nail MH (2007). Reaching out to families with student-created newsletters. *Kappa Delta Pi Record* 44, 39.
- National Research Council (2003). *BIO2010: Transforming undergraduate education for future research biologists*, Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21st Century, Washington, DC: National Academies Press.
- Neuendorf KA (2002). *The Content Analysis Guidebook*, Thousand Oaks, CA: Sage.
- Pennebaker JW, Booth RJ, Martha FE (2007). *Linguistic Inquiry and Word Count: LIWC2007*, Austin, TX: LIWC.
- Seymour E, Hunter AB, Laursen SL, DeAntoni T (2004). Establishing the benefits of research experiences for undergraduates in the sciences: first findings from a three-year study. *Sci Educ* 88, 493–534.
- Stemler S (2001). An overview of content analysis. *Pract Assess Res Eval* 7.
- Stewart ED (1994). On developing gifts and talents: the Laureate Learning Cycle. *J Secondary Gifted Educ* 6, 1–7.
- Tausczik YR, Pennebaker JW (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *J Lang Soc Psychol* 29, 24–54.
- Thiry H, Laursen SL, Hunter AB (2011). What experiences help students become scientists? A comparative study of research and other sources of personal and professional gains for STEM undergraduates. *J Higher Educ* 82, 357–388.
- Villarejo M, Barlow AEL, Kogan D, Veazey BD, Sweeney JK (2008). Encouraging minority undergraduates to choose science careers: career paths survey results. *CBE Life Sci Educ* 7, 394–409.
- Wenning C (2005). Levels of inquiry: hierarchies of pedagogical practices and inquiry processes. *J Phys Teach Educ Online* 2, 3–11.
- Wenning C (2007). Assessing inquiry skills as a component of scientific literacy. *J Phys Teach Educ Online* 4, 21–24.
- Wiley J (2009). Student ownership of learning: an analysis. Unpublished master's thesis, University of Hawaii, Manoa.