2023 Annual Conference & Exposition
Baltimore Convention Center, MD | June 25 - 28, 2023
The Harbor of Engineering Education for 130 Years
ASEE
Paper ID #39265

Measuring Biomedical Engineers' Self-Efficacy in Generating and Solving Provocative Questions about Surgery

Mr. Nathan Zhang, Vanderbilt University

Nathan Zhang is a undergraduate studying biomedical engineering at Vanderbilt University working on biomedical engineering education in conjunction with the Vanderbilt Institute for Surgery and Engineering.

Dr. Stacy S. Klein-Gardner, Vanderbilt University

Dr. Stacy Klein-Gardner serves as an Adjunct Professor of Biomedical Engineering at Vanderbilt University. She is the co-PI and co-Director of the NSF-funded Engineering For Us All (e4usa) project. She is also the co-PI and co-Director of the Youth Engineering Solutions (YES) Middle School project focusing on engineering and computational thinking. Dr. Klein-Gardner is a Fellow of ASEE.

Dr. Michael I. Miga, Vanderbilt University

Michael I. Miga, Ph.D. received his B.S. and M.S. from the University of Rhode Island in Mechanical Engineering and Applied Mechanics, respectively. He received his Ph.D. from Dartmouth College specializing in biomedical engineering. He joined the facul

Measuring Biomedical Engineers' Self-Efficacy in Generating and Solving Provocative Questions about Surgery

Abstract

Self-Efficacy has shown to be positively correlated with academic success [1-3]. A previous study by Mamaril (2016) found significant positive correlations between general engineering self-efficacy and academic success [4]. With an often-cited need for biomedical engineers to engage more closely with the medical field, this study seeks to create an instrument to determine how self-efficacy in biomedical engineering is related to a subject's abilities to identify and solve provocative questions relevant in a clinical environment and ability to write grant proposals related to those questions.

To create the instrument, 35 broad survey questions related to self-efficacy were generated, separated into 4 broad categories: General Self-Efficacy (GEN) a unidimensional scale taken from Mamaril's paper, Engineering Problem Identification in Surgery (IDENT), Engineering Problem Solving (SOLVE), and Engineering Proposal Writing Skills (WRITE). Participants were asked to rate their level of certainty with which they believe they can perform each task on a Likert scale from 1 (Strongly Disagree) to 5 (Strongly Disagree). To ensure content-related validity, this instrument was reviewed by two professors in engineering education who have expertise in self-efficacy research and survey design. To gather validity evidence based on response processes, think aloud protocols were used with two students to improve the wording of the survey before its broader use. Undergraduate and graduate biomedical engineers from a variety of universities across the United States were asked to participate in our survey, with a total of 50 responses. An exploratory factor analysis in the form of principal axis factoring was performed on the categories of IDENT, SOLVE, and WRITE. Correlation values were used to relate these categories to Mamaril's general self-efficacy instrument. Cronbach's alpha was used on each category to determine reliability.

Survey questions that demonstrated low communality with other variables on the pattern matrix from IDENT were removed in future statistical analyses. Our three underlying factors ended up accounting for about 64% of the variation between variables, with each construct having a moderate communality with their main factor and low communality for other factors. Through reliability statistics, IDENT, SOLVE, and WRITE, each showed moderately high internal consistency. Correlative statistics were determined between the averages of the variables in GEN with the averages of IDENT, SOLVE, and WRITE, with positive and moderately high Pearson's correlation coefficients for each test.

We found statistical evidence of reliability and validity of our self-efficacy instrument. Our instrument can use biomedical engineers' self-efficacy to measure their ability to identify and solve provocative questions relevant in a clinical environment as well as write grant proposals related to those questions.

Introduction

There have been increasing efforts to incorporate novel clinical immersive techniques into both graduate and undergraduate biomedical engineering curricula in universities across the United States [5]. For example, programs like Johns Hopkins University's two week long clinical observation program for undergraduates [6] and the authors' institution's unique two semester long Engineering in Surgery and Intervention Course for graduate students [7] attempt to incorporate a clinically immersive aspect to the traditional biomedical engineering core curriculum. Through the use of grades, survey data, reflections and other measures, these programs show promising improvements to students' abilities to understand and identify clinical and surgical needs as biomedical engineers. However, few testable instruments exist that can be used across different programs to assess their efficacy easily. We seek to develop an instrument that can correlate an individual's self-efficacy regarding general engineering skills to their ability to identify and solve provocative questions relevant in a clinical environment and write grant proposals related to those questions.

Self-Efficacy, coined by Bandura [1-3], refers to an individual's personal beliefs in their own capacity to perform certain actions. This theory has shown promise in predicting trends in academic success. One such study found positive and significant correlation between self-efficacy and academic achievement in secondary education [8]. Another influential study by Mamaril sought to measure undergraduate engineering student's self-efficacy, finding significant positive correlations between general engineering self-efficacy and academic success, measured by GPA [4]. We want to test how general engineering self-efficacy might correlate with biomedical engineers' needs identification skills in a clinical setting.

Our main goal was to determine a quantitative method of measuring the self-efficacy of undergraduate and graduate level biomedical engineers that can also be used to assess their abilities in many of these clinical immersion programs. Because the ultimate goal of many of these programs is to improve biomedical engineers' needs-based assessments in clinical settings, we define a provocative question (PQ) based on Miga et al.'s (2021) definition : A provocative question must identify a procedural barrier and/or missing area of knowledge that affects the procedural delivery of care, inspire analysis and design approaches that are focused at solutions have strong engineering needs, and if solved will have a significant clinical impact on health or the understanding of disease. The clear distinction in this framework is that PQs have high clinical significance, considerable scope, and substantive scale. More specifically, PQs are the equivalent of major scientific inquiry arising from the procedural medicine community. It should be noted that, while PQs are a somewhat broad concept, the focus here is quite granular as they are constrained to the surgical and interventional environment and the impact of engineering approaches, i.e., questions and solutions that have reasonable translational timelines.

While the above defines the needs assessment extension resulting from the addition of clinical immersion, it does not fully provide a complete rationale for engineering trainees to be trained in forming PQs or the need of a clinical immersion experience. As discussed in Miga and Labadie

[7], the authors describe the nature of "real-domain experiences" and the goal of establishing an interactive dialogue where capability meets need with better exchange. They also discuss that it is important for trainees to not only to understand the physiology and nature of malady but the mechanism of treatment, and the procedural aspects of application to include the typical experiences of clinical colleagues within the domain, in this case the operating room or interventional suite. Equally important is the recognition that truly PQs are collaborative in nature and intrinsically arise due to the constituent participants with engineers serving in a key role as innovator. More specifically, clinicians are typically trained from an apprenticeship model whereas engineers typically learn from the perspective of invention or innovation. The dialogue that arises among them is a unique interaction involving one training paradigm resistant to change but steeped in clinical experience, and a second paradigm focused at exploratory design and novel approaches. The delicate balance between explicit and implicit knowledge is foundational for the creation of PQs and requires both constituents.

With that backdrop established, this work concentrates on the experience of the engineering trainee and designates three main broad categories toward enabling biomedical engineers to succeed in addressing issues in a clinical setting: (1) Engineering Problem Identification in Surgery (IDENT) focuses on participants' confidence in identifying meaningful PQs in a clinical environment. (2) Solving Engineering Problems (SOLVE) focuses on participants' confidence in their ability to begin to design and solve an identified provocative question. (3) Engineering Proposal Writing Skills (WRITE) which focuses on participants' confidence in their ability to write a sufficient grant proposal based off the provocative question that they chose to answer. An additional fourth category, General Engineering Self-Efficacy (GEN) was taken from Mamaril's (2016) instrument [4], which we will use to develop a link between our own variables and general engineering self-efficacy. It should be noted that there is an equally interesting extension of this work where clinical learning could be assessed within the context of engineering immersive experiences, but at this time is outside the scope of this paper.

In order to draw a statistical relationship among different factors of self-efficacy, an instrument must be developed with validity to the tested topics. Moskal and Leyden's manuscript on the validity of rubrics details three forms of validity: content, criterion related, and construct validity [9]. Content validity focuses on how well a specified instrument can measure desired effects and is often achieved through a system of expert reviewed surveys and questionnaires. Criterion-related validity refers to the instrument's ability to predict related criterion and is often achieved through factors such as rubrics and grades. Construct validity is the ability for the instrument to detect relationships between variables and can be ideally achieved through statistics.

Method

Participants

Biomedical engineers with educational experience ranging from undergraduate to graduate students were asked to participate in the survey study which was approved by the authors' Institutional Review Board. Undergraduate and graduate students across the United States from the authors' university, another university's graduate level biomedical engineering program, and three Research Experiences for Undergraduate (REU) programs topically related to surgery and/or biomedical devices were asked to participate in the survey. The surveyed population consisted of 50 participants who consented and completed the entire survey.

Procedures

All surveys were completed online through the REDCap software [10],[11]. Participants willing to take the survey were given a REDCap link where they could anonymously fill out the survey. Items were presented in the same order to all participants, with items grouped together based on category.

Self-Efficacy Instrument Design

In total, our initial instrument contained 35 survey questions split into the categories of GEN, IDENT, SOLVE, and WRITE. Each survey question was designed to begin with phrases such as "I can identify" to reflect the individual's belief in their own ability to complete that task. A Likert scale of 1 (*Strongly Disagree*) to 5 *(Strongly Agree)* for quantitative statistical analysis was used for each item. Students were asked to rate their level of certainty that they can perform certain tasks. All items are listed in the Appendix.

To ensure the content validity of our experiment, our instrument was intensively reviewed by two professors in engineering education who have expertise in self-efficacy research. One engineering education expert has thirty years of experience and the other has nearly ten years in the field and also expertise in survey design. These reviews resulted in minor edits being made to the draft instrument. Several of these edits were around clarity and conciseness of the prompt. In some places, we were encouraged to break apart a survey item into multiple items because the tasks were unique, (e.g. design and analysis).

To gather validity evidence based on response processes, we asked two students: a BME graduate student and a BME senior undergraduate student to take the survey in a think-aloud protocol, commenting on wording and clarity as they took the survey. The BME graduate student had few edits to the survey overall, making changes to only five questions that included clarifying two terms. The BME senior undergraduate student made additional phrasing changes to improve clarity.

Principal axis factoring, reliability analyses, and correlative analyses were used to determine construct validity. We chose principal axis factoring instead of principal component analysis as we wanted to look at underlying factor structure. Because the survey was designed with three new categories, we hypothesized that three main factors would exist to explain most of the variance among constructs. Ideally, a moderate proportion of each construct's variance, also known as communality, should be explained by one of the three corresponding underlying factors.

Analyses

Principial Axis Factoring, Reliability Analyses, and Correlative Analyses were calculated using IBM SPSS Version 28 [12].

Validity

Our instrument was reviewed multiple times by two professors in engineering education with expertise in self-efficacy and survey design prior to use for content validity. Correlational statistics of each category compared to Mamaril's GEN category were used to determine criterion-related validity. Principal axis factoring was used to determine a degree of construct validity. Reliability statistics were also calculated to ensure consistency of our instrument's scores.

Principal Axis Factoring

Principal axis factoring (PAF) is a type of exploratory factor analysis that looks to identify unknown latent variables within data in order to explain the observed variance. We chose to use PAF instead of the more commonly used method of principal component analysis (PCA) as our goal was not to reduce our data set to three unique factors, but to see how much the variance of IDENT, SOLVE, and WRITE could be explained by underlying factors. Communality, often abbreviated h^2, is defined as the proportion of each constructs' variance that can be explained by an underlying factor, and an $h^2 > |0.500|$ is interpreted to mean that the factor is responsible for a majority of the variance, while a communality of $h^2 < |0.300|$ is ideally desired for factors that do not correspond to the construct's category. Because we had three main categories of survey questions, PAF was conducted with the expectation that three underlying factors would be responsible for our dataset, with the factors corresponding to IDENT, SOLVE, and WRITE respectively.

Correlative and Reliability Statistics

Correlation values were calculated between the average of constructs in GEN with the averages of the constructs in IDENT, SOLVE, and WRITE. We would expect to see a moderate positive correlation as having higher general engineering self-efficacy should in theory correlate to engineering ability to identify, solve, and write about needs in biomedical engineering. Cronbach's alpha was calculated for IDENT (12 variables), SOLVE (6 variables), WRITE (10 variables) in order to confirm the internal consistency of each category. We expect to see Cronbach's alpha values $\alpha > 0.7$ to confirm that our instrument is reliable.

Results

PAF was performed, generating the pattern matrix values displayed below in Table 1. As we would expect, Factors 1, 2, and 3 have high communalities with IDENT, WRITE and SOLVE respectively, with low communalities for the other categories. Our three unique factors ended up accounting for about 64% of the variation between variables, with each construct having a communality > |0.500| with its main factor. These results give statistical credibility that our three underlying factors are responsible for a majority of the variance that we see between our three categories of IDENT, WRITE, and SOLVE.

Table 2 displays general statistical information about IDENT, WRITE, and SOLVE as well as the results of correlative and reliability statistical tests. Pearson's correlation coefficients of $r = 0.449$, $r = 0.608$, and $r = 0.483$ for IDENT, WRITE, and SOLVE, respectively, show that these constructs have a moderately positive correlation with general engineering self-efficacy. Reliability statistical tests return Cronbach's alpha values of $\alpha = 0.930$, $\alpha = 0.861$, $\alpha = 0.948$, respectively, indicating a high degree of reliability for each category.

Table 1 Principal Axis Factoring Pattern Matrix for IDENT, SOLVE, and WRITE.

Factor	1	2	3
IDENT3	.874*	.006	-.023
IDENT7	.855*	.004	-.014
IDENT1	.824*	-.024	.024
IDENT5	.779*	.001	.015
IDENT4	.756*	.019	.173
IDENT6	.752*	-.007	-.084
IDENT2	.679*	-.060	.015
IDENT8	.571*	-.178	-.103
IDENT10	.549*	-.070	.207
IDENT9	.543*	.004	.031
IDENT12	.509*	.181	.272
IDENT11	.508*	-.184	.235
WRITE7	-.108	-.905*	.055
WRITE1	.097	-.898*	-.135
WRITE2	.062	-.881*	-.013
WRITE9	.073	-.866*	.001
WRITE3	-.110	-.856*	.151
WRITE5	-.109	-.747*	.152
WRITE8	.098	-.746*	.014
WRITE4	.089	-.738*	.039
WRITE6	-.042	-.725*	-.004
WRITE10	.230	-.550*	-.077
SOLVE6	-.196	-.096	.826*
SOLVE4	-.009	-.032	.736*
SOLVE3	.095	-.041	.689*
SOLVE1	.254	.063	.645*
SOLVE5	.076	-.177	.611*
SOLVE2	.151	.010	.558*

Note: * Indicates a communality > |0.500| with respective factor.

Table 2 Reliability and Descriptive Statistics for IDENT, SOLVE, and WRITE

Group	Items	Mean Total	SD Total	Correlation to GEN	Cronbach's Alpha
IDENT	12**	3.443	0.7435	0.446*	0.931
SOLVE	6	3.8833	0.6073	0.608*	0.861
WRITE	10	3.4584	0.906	0.483*	0.948

Note: Taken from Sample of 50 participants. One variable, IDENT13 was removed from IDENT category due to low correlation with other factors. * denotes a p value less than 0.001. ** is used to note that one element (IDENT13) was removed from the category

Discussion

Principal Axis Factoring

In our initial calculation of PAF, we found IDENT13 to have low communality with the factor corresponding to IDENT, and thus we removed this construct from further statistical analysis. After its removal PAF was performed again (Table 1). Ultimately three underlying factors accounted for over 64% of the variance observed in our participants. This result strongly implies that each category had a relatively unique and distinct factor to explain the variance. The communalities are consistent with the idea that each of our categories of IDENT, SOLVE, and WRITE had a corresponding factor with communality > |0.500| and a communality < |0.300| for the other confounding factors. A notable exception occurred in the item IDENT13, which was found to have a low communality with all three factors. It's possible that this variation is due to the different structure of the question compared to other questions in IDENT. Whereas all other items in IDENT contain words such as "clinical", "surgical" or "disease" (see Appendix below), IDENT-13's broad wording of "area of missing knowledge" and "an engineering solution" likely is more related to general engineering rather than biomedical engineering.

Correlations between General Engineering Self-Efficacy and IDENT, SOLVE, and WRITE

Overall, general engineering self-efficacy was significant and positively correlated (r > 0.4) for all three categories. These findings are consistent with the common idea that self-efficacy correlates with higher success in academics and professions. While the results of this study may not present a surprise, it does confirm that we have created a reliable and valid instrument that can correlate general engineering self-efficacy to useful skills necessary for biomedical engineers to succeed in clinical environments. Using this instrument periodically throughout the duration of special courses and programs may provide a general evaluation of how effective a particular immersion program is.

Limitations

As both a survey and a cross-sectional study, our study serves to provide a snapshot of a single time point in participants' careers. While we can draw a correlation between IDENT, SOLVE, and WRITE skills and a participants general engineering self-efficacy, a definite causal link cannot be established due to the nature of the study.

While the statistics used in this study demonstrated significant results, a greater sample size would help to make the instrument more generalizable to a wider variety of biomedical engineers. Additionally, because participants were sampled from three topically related REU programs, one other institution, and the authors' institution, our study may represent a population with generally high levels of self-efficacy in both engineering and biomedical engineering ability. However, many clinical immersion programs are specifically targeted towards these types of students. Biomedical engineering is also a diverse field with many subdisciplines. Because our IDENT and SOLVE categories use specific wording such as "clinical impact" or "surgical barrier", our instrument would likely not be very reliable for students in areas in biomedical engineering that do not interact directly with a clinical environment. Still, even with only fifty participants, our instrument should be reflective of the population we intend to measure.

Conclusions

We found statistical evidence of reliability and validity of our self-efficacy instrument designed to assess undergraduate and graduate level biomedical engineers in their abilities to identify, solve and write provocative questions relevant to a clinical environment. Future work might also design broader instruments designed to target a wider variety of biomedical engineers in a wider array of disciplines. Researchers may use our work as a measure of biomedical engineer's self-efficacy in their own clinical immersion programs.

Appendix

Label	Item
General Engineering Self Efficacy (GEN)	
GEN-1	I can master the content in the engineering-related courses I am taking this semester.
GEN-2	I can master the content in even the most challenging engineering course if I try.
GEN-3	I can do a good job on almost all my engineering coursework if I do not give up.
GEN-4	I can do an excellent job on engineering-related problems and tasks assigned this semester.
GEN-5	I can learn the content taught in my engineering-related courses.
GEN-6	I can earn a good grade in my engineering-related courses.
Engineering Problem Identification (IDENT)	
IDENT-1	I can identify a surgical procedural barrier through observing surgery or an intervention.
IDENT-2	I can identify a surgical procedural barrier through participating in rounds or a clinical conference.
IDENT-3	I can identify an area of missing knowledge that affects a surgical procedure through observing surgery or an intervention.
IDENT-4	I can identify an area of missing knowledge that affects a surgical procedure through participating in rounds or a clinical conference.
IDENT-5	I can identify a surgical procedural barrier that, if solved, will have a significant clinical impact on patient health.
IDENT-6	I can identify an area of missing knowledge that, if filled, will have a significant clinical impact on patient health.
IDENT-7	I can identify a surgical procedural barrier that, if solved, will have a significant impact on the understanding of disease.
IDENT-8	I can identify an area of missing knowledge that, if filled, will have a significant impact on the understanding of disease.
IDENT-9	I can identify an engineering approach that can address procedural barriers or missing knowledge identified by a surgeon.
IDENT-10	I can identify precise problems that will have significant clinical impact that require more than creating a lower cost device or an incremental advance in technology.
IDENT-11	I can identify precise problems whose solutions would advance the field beyond incremental improvements in treatment, outcomes, or understanding of human disease/dysfunction.
IDENT-12	I can identify a surgical procedural barrier that requires an engineering solution.
IDENT-13*	I can identify an area of missing knowledge that requires an engineering solution.
Solving Engineering Problems (SOLVE)	
SOLVE-1	I can design engineering approaches and solutions for a surgical procedural barrier.
SOLVE-2	I can analyze engineering approaches or solutions for a surgical procedural barrier.

SOLVE-3	I can design engineering approaches or solutions to fill an area of missing knowledge.
SOLVE-4	I can analyze engineering approaches or solutions to fill an area of missing knowledge.
SOLVE-5	I can integrate a new technology to help solve a surgical procedural barrier.
SOLVE-6	I can integrate a new technology to help fill an area of missing knowledge.
Engineering Proposal Writing Skills (WRITE)	
WRITE-1	I can write a grant proposal's abstract or project summary clearly and completely.
WRITE-2	I can write a grant proposal's specific aims clearly and succinctly.
WRITE-3	I can write a grant proposal's hypothesis that is brief, clear, states an expected relationship or difference, can be tested, and is grounded with sufficient rigor of prior research.
WRITE-4	I can write a grant proposal's significance that explains the importance of a problem and the impact of its solution.
WRITE-5	I can use literature to support a grant proposal effectively.
WRITE-6	I can use tables and figures in a grant proposal effectively.
WRITE-7	I can write a grant proposal's methodological approach as well-reasoned and complete.
WRITE-8	I can write a grant proposal's analytical approach as well-reasoned and complete.
WRITE-9	I can write a grant proposal's experimental design as well-reasoned and complete.
WRITE-10	I can persuasively describe in a grant proposal the resources such as money, time, facilities, and personnel needed to conduct research such that a project will be completed.

Acknowledgements
This work was supported by the National Institutes of Health, National Institute of Biomedical Imaging and Bioengineering, Grant Number T32EB021937.

References

[1] Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 84(2), 191–215.

[2] Bandura, A. (1986). Social foundations of thought and actions: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

[3] Bandura, A. (1994). Self-efficacy. In R. J. Corsini (Ed.), Encyclopedia of psychology (2nd ed., 3, 368–369). New York, NY: Wiley.

[4] Mamaril, N.A., Usher, E.L., Li, C.R., Economy, D.R. and Kennedy, M.S. (2016), Measuring Undergraduate Students' Engineering Self-Efficacy: A Validation Study. J. Eng. Educ., 105: 366-395. https://doi.org/10.1002/jee.20121

[5] M. Kotche, A. E. Felder, K. Wilkens, and S. Stirling, "Perspectives on Bioengineering Clinical Immersion: History, Innovation, and Impact," *Ann Biomed Eng*, vol. 48, no. 9, pp. 2301–2309, Sep. 2020, doi: 10.1007/s10439-020-02508-x.

[6] D. E. A. Logsdon and J. H. University, "A Team Leader Model for Biomedical Engineering Design Team Project-Definition Training and Scalable Clinical Observation (Work in Progress)".

[7] M. I. Miga and R. F. Labadie, "A Novel Clinically Immersive Pre-doctoral Training Program for Engineering in Surgery and Intervention: Initial Realization and Preliminary Results," *Biomed Eng Education*, vol. 1, no. 2, pp. 259–276, Jul. 2021, doi: 10.1007/s43683-021-00051-2.

[8] F. Doménech-Betoret, L. Abellán-Roselló, and A. Gómez-Artiga, "Self-Efficacy, Satisfaction, and Academic Achievement: The Mediator Role of Students' Expectancy-Value Beliefs," *Front. Psychol.*, vol. 8, p. 1193, Jul. 2017, doi: 10.3389/fpsyg.2017.01193.

[9] B. M. Moskal and J. A. Leydens, "Scoring Rubric Development: Validity and Reliability", doi: 10.7275/Q7RM-GG74.

[10] P. A. Harris *et al.*, "The REDCap consortium: Building an international community of software platform partners," *Journal of Biomedical Informatics*, vol. 95, p. 103208, Jul. 2019, doi: 10.1016/j.jbi.2019.103208.

[11] P. A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, and J. G. Conde, "Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support," *Journal of Biomedical Informatics*, vol. 42, no. 2, pp. 377–381, Apr. 2009, doi: 10.1016/j.jbi.2008.08.010.

[12] IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp.