What do words convey? A thematic and linguistic analysis of undergraduates' reasons for choice of major and associations with science motivation

Dr. Nayssan Safavian, Dr. Anna-Lena Dicke, Yannan Gao, Dr. Jacquelynne S. Eccles

Abstract:

To better understand underrepresented undergraduates' motivation for pursuing their major, we examined undergraduates' articulated explanations for their major choice and its association with their motivational beliefs within an introductory chemistry course. Students' (n=503, 68% female, 56% First-Generation College-Going, 34% Hispanic) open-ended explanations for their major choice was examined a) by examining prevalent content themes, and b) by examining linguistic features of their written responses (word count, use of affective words). Undergraduates most frequently referenced occupation-related and enjoyment-related reasons for their major choice. Undergraduates in Life science-related majors that mentioned occupation-related reasons also reported lower levels of interest in Chemistry. We also found a positive association between students' surveyed motivational beliefs in Chemistry and the length of their major choice explanations.

Safavian, N., Dicke, A-L., Gao, Y., Eccles, J. (2021, April). *What do words convey? A thematic and linguistic analysis of undergraduates' reasons for choice of major and associations with science motivation* [Paper Session]. Annual Meeting of the American Educational Research Association, virtual.

*Analyses are ongoing and manuscript is in draft.

For more information, please contact: nsafavia@uci.edu

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A170160 to the University of California, Irvine. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

1. Objectives and theoretical framework

Underrepresented minorities (URM) and first-generation college-going (FG) students enter college just as likely to major in science fields as their white middle class peers, but are much less likely to complete a science major than their white peers (Hayes, 2007; National Science Board, 2014). High attrition rates are particularly common in gateway science courses (Gultice et al., 2015). To better understand why this attrition occurs and how to best counteract it, more needs to be understood about the motivation of these students for pursuing science majors in the first place and how this is related to their motivation for the field they are pursuing.

We know that the motivation for students' educational choices is complex. For instance, Expectancy-Value Theory postulates that students' engagement experiences, course-taking decisions, and aspirations for STEM careers are most directly influenced by students' expectations for success and the value students attach to pursuing such careers. The value they attach is multifaceted and comprises their intrinsic value (i.e., enjoyment), attainment value (i.e., personal relevance), utility value (i.e., perceived usefulness for personal short- as well as long-term goals and plans), and cost (i.e., anticipated negative aspects of engaging in the task). Studies have documented robust associations of students' expectancy and value beliefs and their course-taking patterns, college major and occupational choices related to STEM fields (Jacobs, 2005; Lauermann et al., 2015; Wang, 2012; Wang et al., 2013).

Much of the current work on motivational beliefs and science-related choices, including college major selection and persistence, has been quantitative in nature and generally reliant on survey methodology. Thereby the complex motivation behind the college major selection is not as well understood, particularly for undergraduates who are most at risk of attrition and disengagement. Scientists, however, are expanding their methodological and disciplinary approaches to broaden how they examine and understand such choices. For instance, social psychologists are turning to psycholinguistics for clues about the mechanisms through which language reflects psychological disposition, emotionality, cognition, and social orientation and the associations of language use (word use pattern, production, style, etc.) and motivation and cognition (Tausczik & Pennebaker, 2010; Friberg et al., 2015; Pennebaker & Francis, 1996; Pennebaker & King, 1999). A number of studies suggest that certain types of language use might be an indicator of students' motivation (e.g,. Canning & Harackiewicz, 2015; Harackiewicz et al., 2015). This includes, for instance, the use of valence (i.e., positively vs. negatively valenced words) or the use of personal and social themes (i.e., more personal pronouns words, fewer long words, and more social, family, friends, and human words). These themes might be particularly important for FG and URM undergraduates.

How language use can be leveraged to understand the science-related motivations and choices is relatively understudied. However, the study of students' language use, particularly in the context of self-generated explanations of choice of major offers valuable insights about students' psychological construction of motivation for their major. In this paper, we apply a combination of qualitative and quantitative content analysis as complementary approaches to describe and analyze students' science major choices by examining their written expressed rationale for their major choices. The purpose of this exploratory effort is to describe what students are thinking about broadly when articulating their rationale for their choice of major. We will examine whether themes that arise organically and are observed linguistically are related to students' motivational beliefs as captured through survey methodology.

To this end, we asked the following research questions using data from ethnically and socio-economically diverse undergraduates enrolled in a gateway chemistry course:

1) How do undergraduates enrolled in a gateway chemistry course communicate their rationale for their choice of major in terms of a) prevalent content themes and b) linguistic features (word count and use of affective words)? Are there differences by gender or type of STEM major?

2) Are the chemistry-related motivational beliefs (affective interest, behavioral interest and attainment value) of undergraduates within life science and health-related majors associated with the way they express their major choice in terms of a) content themes and b) linguistic features (word count and use of affective words)?

2. Methods and Data Sources

The current study used data from 503 students (68% female, 56% First-Generation College-Going, 34% Hispanic) enrolled in two sections of a ten-week gateway chemistry course at a large public Hispanic- and Asian-serving (HSI-designated) university in Southern California. Students were surveyed about their attitudes online at the beginning of the course and socio-demographic background information was collected from university records.

Major choice. In open-ended questions, students were asked to report their current (or expected) major and to explain why they chose that particular major. Reported majors were coded into three STEM categories: Life sciences (including health-related majors; e.g., biology, chemistry, nursing), physical sciences (e.g., physics, engineering) and non-STEM. Students' reason for their choice of major was coded in two ways. First, a deductive approach was used to group explanations thematically into five categories (see Table 1 for coding scheme and examples): Competence, Enjoyment/Emotion, Altruistic/Future public good, Occupation orientation, and Other.

Second, students' responses were coded for linguistic features using the Linguistic Inquiry and Word Count text analysis program (LIWC)—a text analysis program that allows for the analysis of of text files and extracts linguistic features of written texts through keyword spotting. Text analysis utilizing LIWC is a commonly used and validated approach to studying word use in social and personality psychology (Tausczik & Pennebaker, 2010). LIWC derives frequency values for a large number of words that are pre-defined and pre-sorted into psychological and linguistic categories. The LIWC output is a ratio of total word count. Along with word count, students' use of affective words (e.g., happy) and more specifically, positive (e.g., love, nice, sweet) and negative emotion words (e.g., hurt, ugly, nasty), were assessed.

Motivational beliefs. To capture students' motivational beliefs, students' affective interest (4 items, α=.91, e.g., 'Chemistry fascinates me.'), behavioral interest (3 items, α=.76, e.g., 'I like to read about chemistry topics in my spare time.') and attainment value (4 items, α=.83, e.g., 'The study of chemistry is personally meaningful to me.') were assessed.

Socio-demographic background characteristics. Students reported their gender and ethnicity. Both were dichotomously-coded (1=Female/Hispanic). Students also reported their parents'

educational background, which was coded into three categories: 1= High school or less, 2= Some college, 3=Four years or more of college.

To answer our research questions, we first examined the frequency of coded themes and linguistic features of undergraduates' major choice explanations. T-tests were used to test for potential differences by gender or type of STEM major. Levene's test for equality of variances was conducted and t-test results not assuming equal variances were reported where appropriate. Next, the associations of chemistry motivational beliefs with coded themes and linguistic features of undergraduates' major choice explanations were examined using regression analysis[1]. For this subset of analyses, only students with life science majors were considered given that chemistry motivational beliefs were assessed. Socio-demographic background characteristics (gender, ethnicity and parental educational background) were included as covariates.

3. Results

RQ1: Thematic and linguistic analysis of coded major choice explanations. Descriptive statistics of thematically coded major choice explanations are presented in Table 2. A higher proportion of females were enrolled in majors related to Life Sciences (78% female), whereas enrollments by gender in majors related to Physical Sciences (47% female) was almost equal. Undergraduates most frequently mentioned *Enjoyment/Emotion* and *Occupation orientation* as reasons for their major choice. References to their personal *Competence* and the importance of contributing to the *Altruistic/Future public good* were mentioned far less frequently. Undergraduates majoring in Life science-related majors referenced occupation-related and enjoyment-related reasons for their choice of major in greatest frequency. Interestingly, however, undergraduates majoring in Physical sciences-related majors predominantly referenced enjoyment-related reasons. One interesting gender difference also emerged: Proportionally, females mentioned reasons related to their *Occupation orientation* more often than males (t (255) = -2.611, p = .010). In addition, undergraduates' reasons for major choice varied by the type of STEM major they were enrolled in. Undergraduates enrolled in Life science majors less frequently made references to *Enjoyment /Emotion* (t (147) = -3.180, p = .002) and contributing to the *Altruistic/Future public good* (t (120) = -2.084, p = .039), but more frequently mentioned that their major choice was related to their future occupation (*Occupation orientation*, t (221) = 8.578, p = .000).

Descriptive statistics for linguistic features of major choice explanations are presented in Table 3. In proportion to their overall word count (M=15.1 words), undergraduates used affective words in 6.50% of their writing. The vast majority of these affective words were positive emotion words (6.27%). Compared to males, females wrote significantly longer explanations (t (297) = -3.025, p = .003), but used a significantly lower proportion of affective words (t (223) = 2.087, p = .038). In addition, undergraduates in Physical Sciences majors used a higher proportion of affective words (t (149) = -4.924, p = .000).

RQ 2: Associations of motivation and major choice explanations. The chemistry-related motivational beliefs of undergraduates within Life Science-related majors were associated with some of the content themes that emerged in their major choice explanations using logistic

[1] The psychometric validity of the motivation scales was examined prior to analysis and indicated satisfactory model fit.

regression analysis (see Table 4). We find that when undergraduates attributed their choice of major to occupation-related reasons, they also reported lower levels of affective and behavioral interest. References to *Enjoyment/Emotion* were not associated with undergraduates' surveyed motivational beliefs.

Lastly, chemistry-related motivational beliefs of undergraduates within Life science majors were also associated with one of the linguistic features of their major choice explanations (see Table 5). Undergraduates with higher levels of affective interest and attainment value wrote significantly longer major choice explanations. No association of motivational beliefs and affective word usage was found.

4. **Scientific significance**

We presented an exploratory examination of students' explanations of their selection of major by simply asking them "Why did you choose this major?" Results from the thematic analyses documented relatively higher proportions of occupation-related and enjoyment-related reasons for students' major choices. This is supported by scholarship on the associations of math and science interest and occupational aspirations and college majors. The prevalence of enjoyment and occupation-oriented reasons for choosing a major suggest two understanding about college education: to explore and find your passion and to prepare for future career. This finding extends previous scholarship on the duality how students' approach their college studies and major selection (Mullen, 2014). Gender and social class factors are potential explanations for the differences (Goyette & Mullen, 2006; Ma, 2009). The fact that altruistic and prosocial motives emerge in explanations underscores the value students attach to future careers in health/medical field and also validates the importance of accommodating students with communal strives in STEM fields (Diekman et al., 2016), which can be particularly important for girls, FG students and students coming from certain cultural background. We further observed that physics majors referenced enjoyment-related reasons in highest frequency relative to other themes. This finding is particularly interesting in light of the significantly higher rates of undergraduate attrition within the field of physics. This contradicts the research that generally supports a positive correlation between interest and persistence.

We also implemented a relatively novel approach to studying undergraduates' explanation of their major choice by examining their use of language. In accordance with recent findings, we documented a positive association between students' surveyed motivations and their linguistic articulation of their reasons in terms of number of words written. The application of psycholinguistics to understanding students' rationale for choice of science-major is relatively new. Findings for this study offer insights into how we can leverage the power of students' language use as a tool for further understanding their motivations and cognitions. In contrast to traditional survey methodology, this is a unique approach to exploring students' rationalization and STEM-related choices by focusing on linguistic cues to identify and describe language use in communicating their motives. Research and educational implementations will be discussed.

1991 words

5. References

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

Chen, X., & Ho, P. (2012). STEM in postsecondary education: Entrance, attrition, and coursetaking among 2003–04 beginning postsecondary students (NCES 2013-152). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*, 21(8), 1051-1057

Friberg, F., Lidén, E., Hakanson, C., & Öhlén, J. (2015). Communicating bodily changes: Physicians' ways of enabling patient understanding in gastrointestinal cancer consultations. *Palliative and Supportive Care*, 13(3), 661-671. doi:10.1017/S1478951514000352

Gultice, A., Witham, A., & Kallmeyer, R. (2015). Are your students ready for anatomy and physiology? Developing tools to identify students at risk for failure. *Advances in Physiology Education*, 39(2), 108-115.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2015). Closing Achievement Gaps With a Utility-Value Intervention: Disentangling Race and Social Class. *Journal of Personality and Social Psychology*, No Pagination Specified. doi: 10.1037/pspp0000075

Hayes, R. (2007, September). Science, technology, engineering and mathematics: Review of the latest CSRDE STEM retention study. Paper presented at the meeting of the National Symposium on Student Retention, Milwaukee, WI.

Jacobs, J. (2005). Twenty-five years of research on gender and ethnic differences in math and science career choices: What have we learned? *New Directions for Child and Adolescent Development, 110*, 85-94.

Goyette, K. A., & Mullen, A. L. (2006). Who studies the arts and sciences? Social background and the choice and consequences of undergraduate field of study. *The Journal of Higher Education*, 77(3), 497-538.

Lauermann, F., Tsai, Y-M., Epstein, A. & Eccles, J. S. (2015). Reciprocal Links Between Adolescents' Math-Related Expectancy/Value Beliefs and Career Plans, Multiplicative Expectancy-Value Associations, and Long-Term Implications for Having a Math-Related Career. *Manuscript submitted for publication.*

Ma, Y. (2009). Family socioeconomic status, parental involvement, and college major choices—Gender, race/ethnic, and nativity patterns. *Sociological Perspectives*, 52(2), 211–234.

Mullen, A. L. (2014). Gender, social background, and the choice of college major in a liberal arts context. *Gender & Society*, 28(2), 289-312.

National Science Board. (2014). Science and Engineering Indicators 2014. Arlington VA: National Science Foundation.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic Inquiry and Word Count: LIWC [Computer software]. Austin, TX: LIWC.net.

Radford, A. W., Berkner, L., Wheeless, S. C., and Shepherd, B. (2010). *Persistence and attainment of 2003-04 beginning postsecondary students: After 6 years* (NCES 2011-

151). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch

Simpkins, S.D., Davis-Kean, P.E., & Eccels, J.S. (2006). Math and Science motivation: A longitudinal examination of the links between choices and beliefs. Developmental Psychology, 42(1), 70-83.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. Journal of Language and Social Psychology, 29(1), 24–54. https://doi.org/10.1177/0261927X09351676

Wang, M. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs and interests. *Developmental Psychology, 48,* 1643-1657

Wang, M., Eccles, J.S. & Kenny, S. (2013). Not lack of ability but more choice: Individual and gender differences in choice careers in science, technology, engineering and mathematics. *Psychological Science, 24,* 770-775.

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 6,* 49-78.

Table 1. Coding scheme for major choice explanation

Category	Definition	Example
Competence	Refers to being or building competence in the chosen field	*"I chose this because I am good at biology."*
Enjoyment/Emotion	Refers to enjoyment of/interest in the field/subject matter	*"Because I am interested in biology and microorganisms"*
Altrusitic/Future public good	Refers to wanting to help others/society	*"I chose this major because I really want to help others."*
Occupation orientation	Refers to their chosen (field of) occupation	*"I chose this major to have a career in the Medical Field."*
Other	Refers to other reasons (mentioned with low frequency)	*"family expectations", "For set up a strong basic"*
No decision	No major chosen yet	*"I don't know what I want to do yet."*

Table 2. Coded major choice explanation by gender and major

	Total	By gender		By major	
		Male (N=154)	Female (N=349)	Life Sciences (N=339)	Physical Sciences (N=102)
	N (%)	N (%)	N (%)	N (%)	N (%)
Competence	17 (3)	7 (4)	10 (3)	10 (3)	7 (7)
Enjoyment/Emotion	164 (33)	52(34)	112 (32)	103 (30)	49 (48)
Altruistic/Future public good	36 (7)	10 (6)	26 (7)	18 (5)	13 (13)
Occupation orientation	191 (38)	44 (29)	147 (42)	163 (48)	13 (13)
Other	28 (6)	12 (8)	16 (5)	13 (4)	11 (11)
No decision	10 (2)	6 (4)	4 (1)	0 (0)	1 (1)
Missing	57 (11)	23 (15)	34 (10)	32 (10)	8 (8)

Table 3. Linguistic features of major choice explanation by gender and major

	Total	By gender		By major	
		Male (N=131)	Female (N=315)	Life Sciences (N=331)	Physical Sciences (N=100)
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Word count	15.18 (11.30)	12.89 (9.60)	16.13 (11.83)	15.54 (11.89)	14.30 (9.15)
Affect	.85 (1.01)	.89 (1.04)	.84 (1.00)	.76 (1.03)	1.18 (.90)
Positive Emotion	.81 (.93)	.82 (.92)	.80 (.93)	.73 (.94)	1.11 (.84)
Negative Emotion	.04 (.22)	.05 (.23)	.04 (.22)	.03 (.21)	0.06 (.24)

Note. Means for linguistic categores Affect, Positive and Negative Emotion are in reference to the ratio of words within the category to word count (e.g., ratio of affective words used in proportion to the word count within the statement).

Table 4. Logistic regressions of major choice themes on motivational beliefs

	Enjoyment/Emotion			Occupation oriented		
	OR	*OR*	*OR*	*OR*	*OR*	*OR*
Affective Interest	0.98			0.83†		
Behavioral Interest		1.13			.74*	
Attainment Value			0.98			0.9
Gender	1.1	1.13	1.1	0.95	0.92	0.98
Hispanic	1.14	1.11	1.14	0.89	0.91	0.87
Parental Education	0.85	0.84	0.85	1.07	1.06	1.05

Note. Due to sample size limitations, analyses was limited to Enjoyment/Emotion and Occupation themes.

Table 5. Regressing linguistic features of major choice explanation on motivational beliefs

	Word count			Affect		
	b	*b*	*b*	*b*	*b*	*b*
Affective Interest	0.13*			0.02		
Behavioral Interest		0.08			-0.05	
Attainment Value			0.11*			0.02
Gender	0.08	0.07	0.06	0.03	0.02	0.03
Hispanic	0.02	0.00	0.01	0.00	0.01	0.00
Parental Education	0.03	0.01	0.01	0.04	0.04	0.04

Note. Standardized beta coefficients reported.