The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
Paper ID #42250

Exploring Diverse Work Personas of Engineering Design Graduates through Cluster Analysis

Dr. Xiao Ge, Stanford University

Xiao Ge's design research focuses on understanding creative work theory and practice to improve practices of creativity, interdisciplinary teamwork, and engineering education. Her research at Stanford spans across disciplines in the Mechanical Engineering Department and the Psychology Department at Stanford University.

Dr. Mark Schar, Stanford University

The focus of Mark's research can broadly be described as "pivot thinking," the cognitive aptitudes and abilities that encourage innovation, and the tension between design engineering and business management cognitive styles. To encourage these thinking patterns in young engineers, Mark has developed a Scenario Based Learning curriculum that attempts to blend core engineering concepts with selected business ideas. Mark is also researches empathy and mindfulness and its impact on gender participation in engineering education. He is a Lecturer in the School of Engineering at Stanford University and teaches the course ME310x Product Management and ME305 Statistics for Design Researchers.

Mark has extensive background in consumer products management, having managed more than 50 consumer driven businesses over a 25-year career with The Procter & Gamble Company. In 2005, he joined Intuit, Inc. as Senior Vice President and Chief Marketing Officer and initiated a number of consumer package goods marketing best practices, introduced the use of competitive response modeling and "on-the-fly" A|B testing program to qualify software improvements.

Mark is the Co-Founder and Managing Director of One Page Solutions, a consulting firm that uses the OGSP® process to help technology and branded product clients develop better strategic plans. Mark is a member of The Band of Angels, Silicon Valley's oldest organization dedicated exclusively to funding seed stage start-ups. In addition, he serves on the board of several technology start-up companies.

Dr. Helen L. Chen, Stanford University

Helen L. Chen is a Research Scientist in the Designing Education Lab in Mechanical Engineering and co-founder of the Integrative Learning Portfolio Lab in Career Education at Stanford University. She earned her undergraduate degree from UCLA and her PhD in Communication with a minor in Psychology from Stanford. Her scholarship is focused on engineering and entrepreneurship education, portfolio pedagogy, reflective practices, non-degree credentials, and reimagining how learners represent themselves through their professional online presence.

Prof. George Toye

Ph.D., P.E., is adjunct professor in Mechanical Engineering at Stanford University. While engaged in teaching project based engineering design thinking and innovations at the graduate level, he also contributes to research in engineering education, effect

Dr. Sheri D. Sheppard, Stanford University

Sheri D. Sheppard teaches both undergraduate and graduate design-related classes, conducts research on fracture mechanics and finite element analysis, and on how people become engineers. From 1999 to 2008, she was a Senior Scholar at the Carnegie Foundation for the Advancement of Teaching, leading its engineering study. Sheppard has contributed to significant educational projects, including the Center for the Advancement of Engineering Education and the National Center for Engineering Pathways to Innovation (Epicenter). Her industry experience spans Detroit's Big Three: Ford, General Motors, and Chrysler. At Stanford, she has served as faculty senate chair, associate vice provost for graduate education, founder of and adviser to MEwomen, and leads the Designing Education Lab (DEL), which aims to revolutionize engineering education. Her achievements have earned her numerous honors, such as Stanford's Walter J. Gores Award and the American Society for Engineering Education's Chester F. Carlson and Ralph Coats Roe Awards. She earned her PhD from the University of Michigan.

Exploring Diverse Work Personas of Engineering Design Graduates through Cluster Analysis

ABSTRACT

(Paper type: Research) Engineering schools increasingly endeavor to diversify their educational offerings to meet students' needs of pursuing various career trajectories. However, students' aspirations are often shaped by engineering stereotypes and less informed by profiles of real-world engineering practices. Moreover, little research has been done to portray diverse profiles of engineering graduates. This study aims to fill this gap by conducting a cluster analysis to synthesize archetypical work profiles (i.e., work personas) based on a survey sample of 719 graduates of Mechanical Engineering with on average 16.8 years after completing a graduate-level design program from a U.S. university. *What are the different work profiles of engineering design graduates and how can these different profiles be characterized?* To answer the question, we used Principal Component Analysis (PCA) in combination with *k*-means Cluster Analysis which constructed five types of personas based on self-reported behaviors in engineering, design and innovation activities. We used job functions and a variety of self-efficacies to further validate and analyze the distinct personas that the engineering graduates have pursued. Of the five clusters, we find many attributes of Cluster One (C1) (6.1% of the survey population) diverge from engineering, design and innovation activities; we have focused on analyzing the Early Career Engineers C2 (25.3%), Expected Engineers C3 (29.8%), Managerial Engineers C4 (17.4%) and the Innovative Engineers C5 (21.34%). Innovative Engineers C5 shows consistently high engagement with all the different engineering design-affiliated work activities. Early Career Engineers C2 and Expected Engineers C3 exhibit similar shapes of engagement, with relatively lower scores in marketing & sales behaviors than the other behaviors. Managerial Engineers C4 have the most MBAs and expand what engineers do. The various engineering design routes engineering designers have taken as evidenced from the dataset expand our understanding of what engineers could do. The current study provides concrete insights for guiding engineering students in exploring various career options and contributes to engineering education research on career pathways.

Keywords: Workplace, Engineering Professions, Quantitative Analysis, Cluster Analysis, Graduate Education, Career Paths

1 Introduction

One often neglected aspect of engineering education is career education. Along with the accumulation of disciplinary knowledge and skills, students also develop an identity as an engineer [34]—a sense of "who I want to become" and "what I want to do". Career choice is arguably one of the most important decisions students make as they navigate through their engineering education. However, senior engineering students are often found unsure about or struggling with career decisions [34]-[36]. Moreover, while engineering schools endeavor to diversify their educational offerings to meet students' needs of pursuing various career trajectories, engineering students' aspirations are often shaped by technical course experiences and engineering stereotypes [37] and less informed by personas of real-world engineering practices [38]-[39], [43].

The various real-world experiences and practices from professionals with engineering degree backgrounds could serve as a valuable source of insights. From the perspective of social cognitive career theory [40], professionals over the years would traverse multiple iterations of the career development algorithms. Their work pathways may illuminate prospective career decision-making for engineering students, offering them new perspectives of what their future could hold. There has been limited empirical research that portrays diverse profiles of work professionals who graduated from engineering programs [41], [42]-[43]. To contribute to this research, we examine behaviors and beliefs of professionals with engineering degree backgrounds. As an initial step, we analyze work profiles based on a rich alumni dataset of graduates who completed one to two engineering design innovation and entrepreneurship courses (hereafter referred to: design program) as part of their graduate education at a U.S. university between 1987 and 2018. We argue that a data-driven approach such as cluster analysis is suitable for this study by identifying different engineering profiles with similar work behaviors.

In the following sections, we review literature on engineering career challenges, outline our empirical cluster analysis approach and its background, present findings of five work personas, and discuss the findings to inform student career planning. We address study limitations and propose future applications for data-driven approaches.

2 Engineering education research on career pathways

Student career preparedness has been a central concern of engineering education and research. Efforts for bridging the gaps between education and career are seen in various research initiatives from studies unpacking students' experiences (e.g., [30], [33]-[35], [53]-[54]) to transformative educational programs (e.g., [47]-[48], [61], [66]) to institutional assessments (e.g., [45]) to examination of engineering practices (e.g., [31], [49]-[51]). Most empirical work on career development is found to examine perceptions and experiences of students instead of reflections from alumni/professionals [52]. Due to page limitations, we focus on reviewing relevant research of engineering alumni and professionals.

2.1 Research about engineering alumni and professionals

Engineering alumni research has been largely driven by the question of education effectiveness and empirically through collecting qualitative or quantitative data of alumni's subjective perceptions to assess the impact of the programs they graduated from [55]-[59]. While some research has started to focus on alumni's work activities, often focusing on specific mindsets or behaviors, such as entrepreneurial tendencies [60]-[63] and research engagements [65], less has set out to understand general career behaviors. Brunhaver and colleagues [31], [41], [61] acted upon this critical gap by examining the various employment engagements of early-career engineers. For instance, [41] analyzed engineers' different career choices as well as personal, experiential, and affective outcomes based on a large cross-institutional database of alumni from various engineering majors. More recently, [31] explored how relationships with coworkers and supervisors shape early-career engineers' work experiences and identity. Another related area of research is design research situated in real-world engineering practices. Instead of career/job pathways, most research of engineering professionals aims to provide insights about the design process, design expertise or other design qualities [50]-[51], [64]. Little of this research is positioned to draw career insights by analyzing various kinds of work profiles.

2.2 Beyond engineering and non-engineering pathways

Not surprisingly, post-graduation work and careers are broad and diverse [41]. This is supported by empirical evidence that has been consistently found from large-scale, cross-institution survey studies [67]-[68] as well as small-scale, mixed-methods studies [44]. However, many studies are found to use generic labels such as engineering versus non-engineering to categorize which pathways students choose after graduation. Research on engineering versus non-engineering pathways is partly driven by the call to retain engineers. While this categorization is useful, such framing may be misleading as to not acknowledge non-engineering work as an important direction for engineers to flourish and contribute to society. Students who leave engineering majors are often called "non-persisters" [69]. Similarly, "success" is sometimes implicitly defined by engineering students completing their engineering degree programs (e.g., [5]) or pursuing the expected career directions.

Overall, we argue that while it can be useful to learn about broad career trends and patterns, generic labels often oversimplify the diverse paths that graduates take, neglecting the nuanced nature of their roles and the dynamic ways they identify with engineering. Instead of framing the research to address who persists and who does not in engineering, we explore various possibilities of engineering education outcomes—there is no right or wrong, desirable or undesirable career choices. Due to the dynamic landscape of modern professions, we argue that we can get a better picture of work types by relying on what they actually do in their jobs rather than patterns of job titles. Little career research is grounded in what engineers actually do, which may hinder a comprehensive understanding of their work and how they flourish in their careers.

The current work attempts to address this gap in research, and contributes to growing the body of research on engineering alumni.

3 Analysis overview

Cluster Analysis (or clustering) is a family of unsupervised or semi-supervised learning procedures that seek to organize a dataset into homogeneous clusters [3]. Various cluster analysis models have been developed with numerous applications across research fields [27]. Cluster analysis methods primarily rely on a selection of variables with the objective of maximizing within-cluster similarity and in the meantime enhancing between-cluster dissimilarity [3]. We argue that cluster analysis is suitable for the current study that explores the un-predefined profiling of professionals from a large, diverse pool of alumni data. We also note that the unsupervised or semi-supervised learning algorithms-based cluster analysis methodology is synergetic with the qualitative methods of persona research (see [9] for a review) that has also seen emerging applications in engineering education research.

3.1 Cluster analysis in engineering education research

Below, we give a brief overview of cluster analysis methods and applications within engineering education research. In-depth reviews about cluster analysis techniques can be found in [1], [27], [28]. Within engineering education research, studies applying cluster analysis are rather limited. [1] only identified five articles that have applied cluster analysis in the Journal of Engineering Education by 2017. We have only found three empirical papers using cluster analysis within ASEE Engineering Research and Methods (ERM) division [6]-[8]. Although applications are limited, engineering education researchers have used this exploratory approach in various ways.

The majority of studies have applied cluster analysis to understand student learner types in certain courses or other learning contexts [2], [9], [6], [14], [17], [23], in terms of academic success [4], [5], anticipation of future goals [11], or other attitudinal and behavioral factors [14], [23]. A few studies have used higher-education institutions as a unit of analysis to understand different types of educational systems [20], [21], [26]. We have also identified a few novel applications of cluster analysis, such as to understand engineering journal usage [16]. Cluster analysis has also been used as a sampling technique for qualitative research [15], [24]. We have only found one article that examined engineering alumni data using cluster analysis [24].

Like in other research fields, *k*-means methods are widely applied in engineering education research (e.g., [1], [11], [16], [21], [26]). On the other hand, several papers have compared different cluster analysis methods for their research contexts [11], [18], [23]. However, there is not an emerging agreed-upon method across these different studies (see appendix S1 for more review). The three empirical papers we found in the ASEE ERM division have each applied a different clustering method. [6] applied *k*-means clustering to understand levels of data analysis training amongst engineering undergraduates. [7] used hierarchical clustering for creating engineering student typologies based on self-reported talents and abilities to better prepare students for future work. Finally, [8] used topological data analysis to capture latent diversities of first-year undergraduate engineering students in terms of attitudes, beliefs, and mindsets.

4 The current study

We propose that there could be diverse work profiles in graduated professionals even from the same educational programs. To empirically explore this idea, we analyze a large alumni dataset of engineering design graduates who have finished their programs about engineering design innovation and entrepreneurship between 1987 and 2018. More specifically, we aim to address the following research question: *What are the different work profiles of engineering design graduates and how can these different profiles be characterized?*

Because of the large size of the alumni dataset, we started by experimenting with the most widely adopted method—*k*-means clustering. The simple *k*-means clustering method starts with an assigned number of clusters (i.e., k); partitions are then created as a function of the underlying data. Then, the algorithm optimizes the partitioning to obtain a solution to best represent the underlying structure of the data. The steps are repeated to compute optimized cluster centers. The quality of clustering outcome is often evaluated by variability of observations within each cluster (e.g., within cluster sum of squares that measures the squared error between the mean of each cluster and the data points in the cluster) and variation between clusters (e.g., between cluster sum of squares). Elbow method, silhouette score and gap statistics method are some of the most popular cluster metrics to help choose the optimal value of k.

To derive a better clustering outcome, we iterated with hybrid approaches to improve the algorithms' fit with our dataset. This has led us to choose a two-step varied *k*-means method, using principal component analysis (PCA) for data pre-processing and *k*-means for cluster analysis [12], [29]. PCA-based *k*-means clustering is a heuristic clustering approach in which researchers have developed various algorithms depending on research needs. In this study, we have followed a simple two-step approach [12], [29] which proves to dramatically improve clustering results. The selection of variables for inclusion in the analysis plays a critical role in the resulting solution [1]. Instead of sorting out patterns in job titles, we argue that we can get a more accurate and nuanced picture of work types by studying what they actually do in their jobs.

5 Methods

5.1 Dataset

The cleaned-up dataset comprises data about 719 engineering alumni who have attained master's degrees in Mechanical Engineering from a U.S. university, with an average of 16.8 years since the completion of a graduate-level design program—one to two engineering design innovation and entrepreneurship courses. Due to page limitation, please find details about Participants and Procedures, as well as data collection and cleanup in our previous publications [51], [61], [62].

5.2 Measures

Engineering design-affiliated work activities: We included a set of 11 items to capture participants' work engagement with various aspects of engineering design. Six items were drawn from the innovation behavior scale [78]; two engineering-task items came from [73]; the other three items were based on [76] that measures marketing and product improvement behaviors. A 0-4 point Likert scale was used, i.e., from Never to Very Often. See appendix S2 for details.

Workplace factors: We included a set of eight workplace factor items (see appendix S2) about ways engineers may engage with their work, from documentation, to solving ambiguous problems, to leadership and so on. The measure was developed by [73] which captures work environmental influences on engineers' workplace behaviors. The same Likert scale was used.

Job roles: A few variables were included to measure job roles, including: working for a medium- or large-size business (relative to all the alternatives, coded as a 0-1 dummy variable), and multiple choice questions about specific work functions (e.g., working in R&D, Design, Manufacturing, or Management roles) and career choices (e.g., Startup career).

Self-efficacy measures: Self-efficacy measures people's perceived confidence in their ability to successfully perform tasks and activities in certain domains, and have shown to be important predictors of their work outcomes [74]. We use pre-established scales as detailed in [61] to capture participants' beliefs about their personal efficacies in four interrelated domains that were the main focuses of the design programs, namely *design thinking self-efficacy, innovation self-efficacy, entrepreneurial self-efficacy, and engineering tasks self-efficacy* (see appendix S2).

Demographics: A few demographic variables (see appendix S2) are included following our prior work [61], including gender, race and ethnicity and years since graduation from the design program (hereafter referred to: years of experience). We also included graduates' educational background (e.g., have obtained MBA or not).

5.3 Analysis

We first computed the principal components based on the selected 11 variables. Based on eigenvalues that are equal or bigger than 1, we selected principal components to form the PCA-subspace. PCA dimensionality reduction is beneficial for *k*-means clustering as PCA allows us to search for the optimal, and usually better, solution more effectively [12]. With the PCA subspace, we performed *k*-means clustering and determined the value of k by comparing the elbow method and gap stat charts, and assessed the model quality using within-/between-sum-of-squares statistics. We then make sense of the result and analyze the structures of the clustered data. Cluster differences are validated with externally variables [80]-[81] and analyzed using one-way ANOVA, or MANOVA (Multiple Analysis of Variance) in case of shared dependencies between variables. Post-hoc Tukey's HSD pairwise comparison tests are conducted to find out where the cluster differences come from. Here, parametric methods are used to analyze Likert-scale data subsequent to verifying model assumptions. The use of parametric methods is recommended in previous literature [82], [83] and is based on the assumption that the current variables can be construed as continuous variables.

6 Results

6.1 Clustering results of work profiles

The PCA subspace formed based on the first four principal components that capture the majority of the variability in the data. The results of *k*-means clustering suggests k = 5 (see Figure S3.1 for elbow method and gap stat charts; see Figure S3.2 for the scatter plot). The ratio of between-sum-of-squares and total-sum-of-squares is 0.62. The resulting cluster sizes are as follows—44 participants in C1, 182 in C2, 125 in C3, 214 in C4, and 154 in C5. The five clusters are ordered by average scores of the 11 engineering design-affiliated work activities.

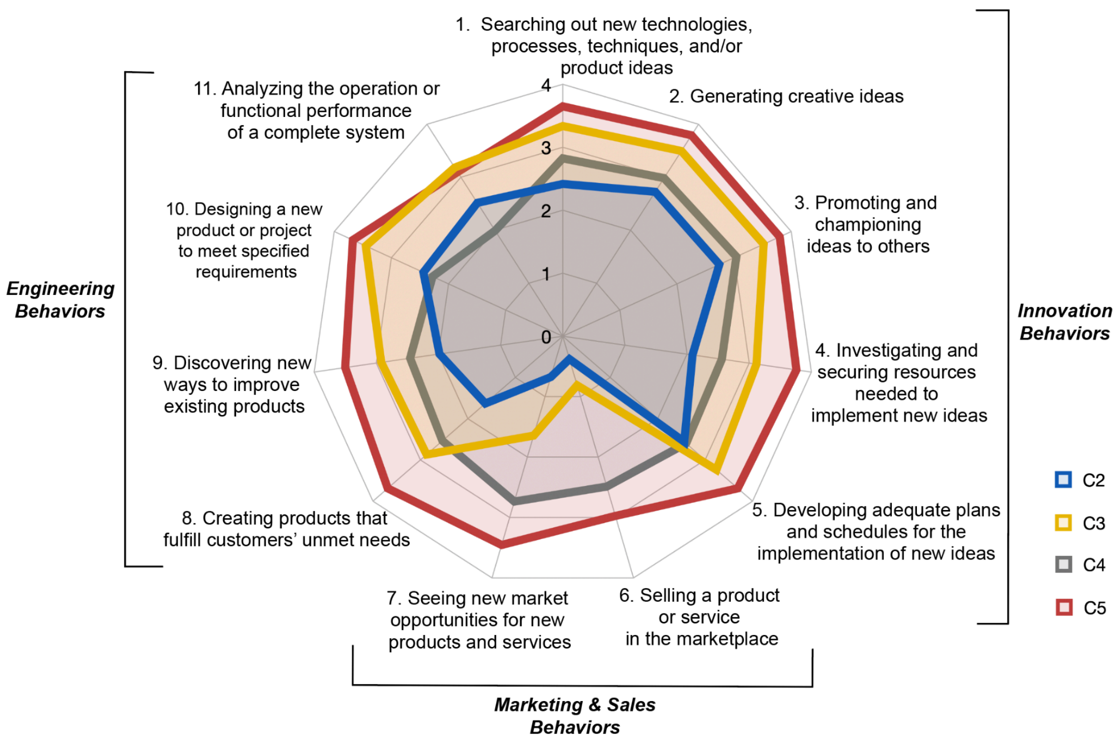

Figure 2. Cluster profiles of C2 - C5 showing group-average scores on the 11 engineering design-affiliated work items (see Table S3.5 for detailed statistics).

In terms of common demographic variables (see Table S3.3), we find a higher percentage of male participants across the board, and there are no cluster differences in the ratio of gender. As for the racial make-up of each cluster, Whites make up from 45% to 55% of participants within each cluster, followed by Asian or Asian Americans (27% to 40%). There are no cluster differences in terms of race and ethnicity. All five clusters are quite experienced, with average years of experience ranging from 13.6 to 20.7 years; significant cluster differences are found, $F(4, 714) = 13.99$, $p < .001$ (see Table S3.4).

6.2 Characterizing work profiles

C1 is the smallest sample of all groups, making up only 6.1% of the full sample. Compared with the other four groups, C1 scores low across the board in engineering design affiliated activities (Table S3.5), as well as self-efficacy measures and workplace factors (Table S3.6). Job role distribution in C1 shows a more varied picture of work than other groups (Table S3.7). These observations may indicate that those in C1 engage less and identify less with innovative engineering design domains than other clusters. It is not clear how well the survey captures the professional engagement of C1. Therefore, we decide to focus on the other four work cluster profiles—the *Expected Engineers C3*, *Innovative Engineers C5*, *Managerial Engineers C4* and the *Early Career Engineers C2,* which make up 93.9% of the sample. We will come back to considering C1 in the Discussion (Section 7). The clustering results for the four clusters are illustrated in Figure 2, showing cluster-average scores on the 11 engineering design-affiliated work activities.

There is an overall consistency of between-cluster patterns across measures. Many of the cluster differences in engineering design-affiliated activities (Figure 2) are found to match with the differences in related yet different measures about self-efficacies (Figure 3) and workplace factors (Figure 4). One-way ANOVA tests yield significant differences across clusters in all measures (Table S3.8 & S3.9). Due to potential shared dependencies between measures of self-efficacies and between items of workplace factors, we have also conducted one-way MANOVA tests combining all relevant measures and the results do not change.

In order to gain further insights into the diverse work profiles, we also map out how the clusters vary in terms of specific job roles and other educational and work backgrounds. Figure 5 (and Table S3.7) illustrates the results that measure and compare the percentage of participants of the cluster they belong to in a given category of job role/work background. By triangulating the data, we further analyze each work profile. We highlight some patterns next.

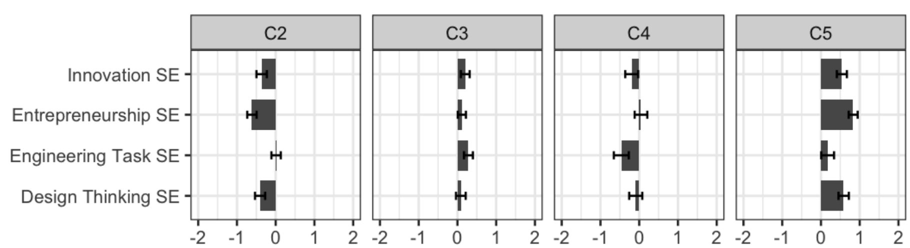

Figure 3. Comparison of clusters C2 - C5 based on z-scores of self-efficacy measures. Error bars represent 95% Confidence Intervals (see Table S3.6 & S3.8 for detailed statistics).

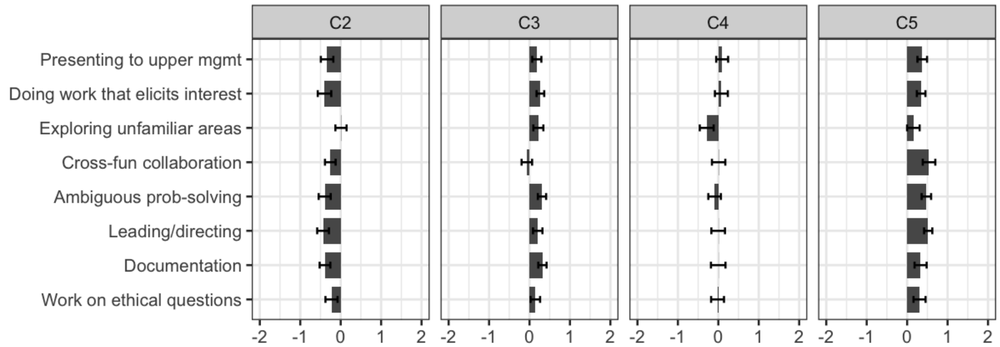

Figure 4. Comparison of clusters C2 - C5 based on z-scores of workplace factors. Error bars represent 95% Confidence Intervals (see Table S3.6 & S3.9 for detailed statistics).

6.2.1 Expected Engineers C3 (29.8% of the sample)

Many of C3's attributes fit the stereotypical image of engineering and suggest that C3 might describe the traditionally-expected engineers. **1. Excelling in engineering tasks**: C3 stands out as the only cluster that highly engages with engineering activities and yet at the same time not engaging much with marketing & sales activities (Figure 2). Along with C5, C3 is the most confident cluster in engineering tasks—the differences of C3 from C2 and C4 are significant, $p = .04$ and $p < .001$ respectively (Table S3.8). Unlike C5, however, C3 is not as confident in other self-efficacy domains. **2. Taking on typical engineering design roles**: C3 has taken typical engineering design roles—Design, R&D, Manufacturing and Project management (Figure 5, left), with more than 50% of them working in mid-to-large organizations and unlikely to have obtained MBA or become a founder (Figure 5, right). C3 on average often engages with various workplace factors except for cross-functional collaboration (Figure 4).

6.2.2 Innovative Engineers C5 (21.4% of the sample)

The clustering algorithms have assigned C5 as the only group that has high levels of engagement in all work activities of innovation, engineering tasks, and marketing & sales (Figure 2). **1. Excelling in engineering, innovation and marketing & sales**: As the older group (M_{year} = 18.53, SD_{year} = 8.15, Table S3.4) in the "high engineering task pair", C5 may represent the group that the Expected Engineers C3 (M_{year} = 16.12, SD_{year} = 8.75) is growing into. Though comparable in engineering task self-efficacy, C5 is more confident than C3 in innovation (adjusted $p < .001$), entrepreneurship (adjusted $p < .001$), and design thinking (adjusted $p = .005$) (Table S3.8). **2. More founders and managerial backgrounds than C3**: A closer examination of job roles and work backgrounds (Figure 5) reveals that C5 has a much higher percentage of founders (30.5%) than C3 (4.2%) or any others (C4: 23.2%, C2: 2.2%); C5 also has a much higher percentage of having obtained MBA (16.9%) compared with C3 (4.6%). C5 has less percentage of people to work in mid-to-large companies than C3 (33.8% compared to 56.5%). **3. Taking on traditional but also diverse job roles**: C3 and C5 have similar involvement in Design and R&D; yet, C5 reports more managing (Project Management, Functional Management) as part of their job function than C3 (53.9% and 30.5% for C5, as compared with 40.2% and 15.0% for C3), as well as more percentage of people in other job roles that are traditionally not regarded part of engineering, such as Sales (13.6% vs 4.7%), Public Relations (16.9% vs 4.7%), and Fiance (7.1% vs 0.9%) than C3. **4. High workplace factors**: C5's work involves frequently engaging in areas of solving ambiguous problems, cross-functional collaboration, leadership and work with upper management, and so on. Notably, C5's engagement with working on ethical questions is significantly higher than all other groups, adjusted *ps* < .001 (Table S3.9).

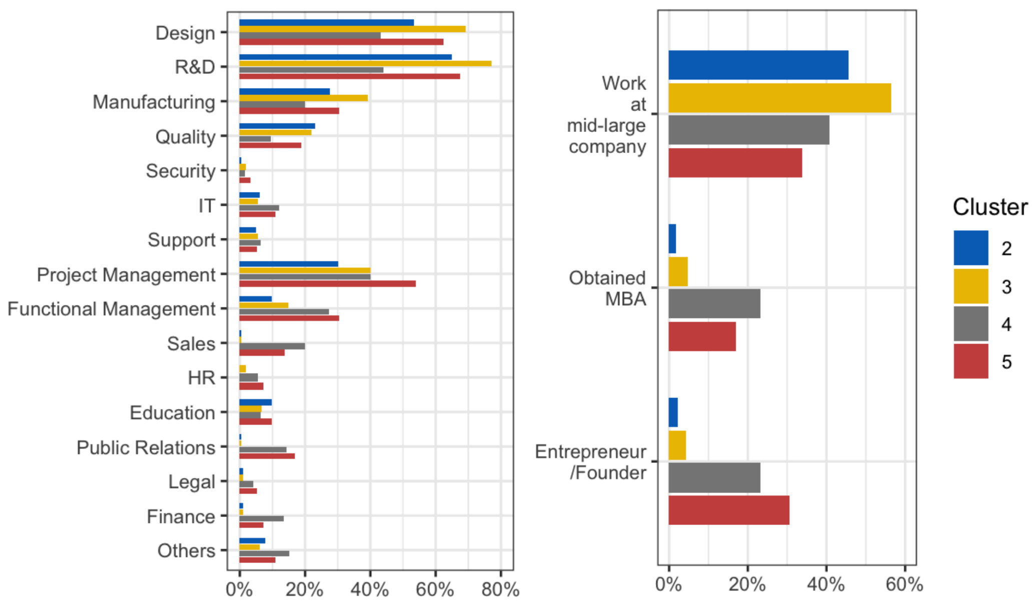

Figure 5. Comparison of clusters C2 - C5 based on job roles and work backgrounds (see more in Table S3.7). Percentages are calculated by dividing frequency of occurrence by cluster size for each of the five clusters. Note that the job roles are not mutually exclusive from each other; participants were asked to choose all options that apply.

6.2.3 Managerial Engineers C4 (17.4% of the sample)

As one of the oldest groups (M_{year} = 20.77, SD_{year} = 8.43, Table S3.4), C4 seems to have taken a slightly different route than the C3-C5 path and are less likely to see themselves as engineers. **1. Most MBAs with less engineering engagement**: C4 has acquired most MBAs (23.2%), compared with C2 (1.7%), C3 (4.7%) and even C5 (16.9%). A moderate percentage of C4 works in mid-to-large companies (40.8%), and at the same time it also has quite a few founders (23.2%). According to the clustering results (Figure 2), unlike the Expected Engineers C3 and Innovative Engineers C5, C4 has a much lower engagement in traditional engineering tasks including analyzing the operation or functional performance of a complete system; C4 is also less engaged with designing new products or projects to meet specified requirements. **2. Expanding what engineers do**: Compared with C3 and C5, C4 has less percentage of people in Design, R&D and Manufacturing. On the other hand, its occupancies in nontraditional engineering roles, such as Sales (20%), Public Relations (14.4%), and Finance (13.6%) are comparable to C5. Because a large percentage of C4 choose Others (Figure 5, left), we take a closer examination of the open-ended responses. Some 12 of them share that they work in managerial positions from product management to executive leadership (e.g., CEO, COO), three participants are business owners, and another nine participants identify with business strategy & planning. A few share that they work in non-engineering jobs, yet three participants explicitly acknowledge that their non-engineering work “requires an engineering background”. For instance, one works in a “non-engineering job that requires some technical knowledge and giving feedback/useful insights to engineering teams.”

6.2.4 Early Career Engineers C2 (25.3%)

While Managerial Engineers C4 and Innovative Engineers C5 are on the more experienced side compared with C3, the Early Career Engineers C2 (M_{year} = 13.64, SD_{year} = 9.26) is on the less experienced side (Table S3.4). **1. Similar work, less experience than C3**: C2 has an even smaller percentage of founders (2.2% vs 4.2%) and MBAs (2.2% vs 4.2%) than the Expected Engineers C3. Some 45.6% of C2 works in mid-to-large companies. Figure 5, left, shows that the job composition of C2 is similar to that of C3, with only slightly fewer percentages in Design or R&D. C2 is also the only group that exhibits the same *shape* of engagement with the 11 engineering design-affiliated work activities as C3—i.e., similar levels for all except for much lower levels in marketing & sales (Figure 2). Across the board, C2 consistently demonstrates lower levels of engagement in each measure of engineering design-affiliated activities compared to C3. **2. Less confident in engineering, design thinking, innovation and entrepreneurship than C3**: Though Early Career Engineers C2 is more confident than the Managerial Engineers C4 in engineering tasks, adjusted $p < .001$, C2 is less confident than C3 in engineering tasks, adjusted $p = .04$, as well as in all other self-efficacy measures, adjusted $ps < .001$ (Table S3.8).

7 Discussion

To fill the gap in research that lacks empirical insights into how engineering students’ career pathways may evolve in the future, we conducted a PCA-based *k*-means cluster analysis study to portray profiles of work professionals who graduated from engineering design programs. Rather than relying on traditional job labels, we identify work personas based on what engineers actually do. We selected a set of creative engineering design-affiliated work activities to identify five clusters, and further analyzed and validated the work profiles by comparing them across other variables about self beliefs, workplace factors, job roles and other characteristics.

7.1 Understanding career routes based on the diverse work profiles

To synthesize the findings, we map out the different work personas based on the *shapes* and *areas* of the radar chart Figure 2, to illustrate the cluster differences and potential relationships in Figure 6. Both Managerial Engineers C4 and Innovative Engineers C5 land in the right two quadrants, as “roundness” largely reflects whether the work cluster engages with marketing & sales activities. On the other hand, both Expected Engineers C3 and Innovative Engineers C5 land in the top two quadrants, based on their higher engagements in engineering design-affiliated activities than the other two groups.

Those in Innovative Engineers C5 are engaged in both developing technical ideas and in their scaling in terms of business opportunities, whereas Expected Engineers C3 is more focused on developing technical ideas, and Managerial Engineers C4 on business opportunities. This diaspora is illustrative of how a common engineering design focused education can lead to be the basis of professional work that employs different combinations of technical/engineering and business development/management acumen. We do know that the average experience of those in Expected Engineers C3 is significantly less than for Innovative Engineers C5 (16.12 yrs vs 18.53 yrs, Table S3.4); this leads us to wonder if some individuals in C3 are on a pathway to join those in C5 (indicated by the dashed orange arrow in the Figure 6).

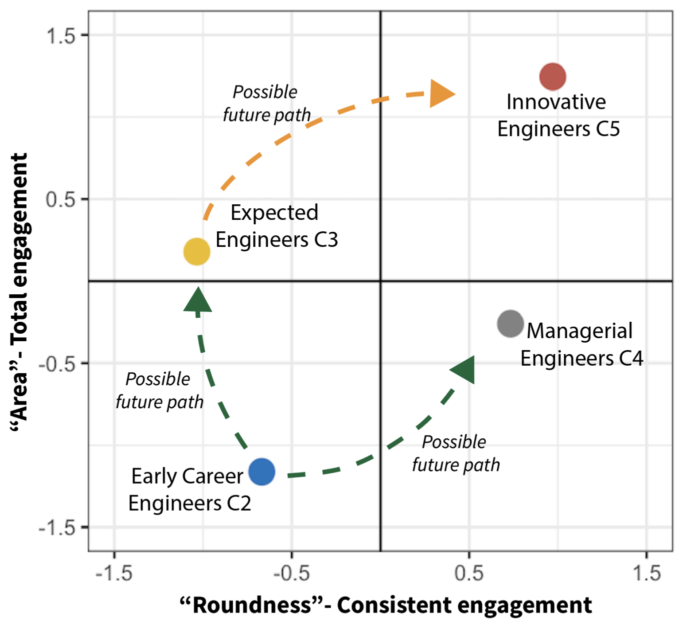

Figure 6. Mapping group engagements in various engineering design-affiliated work activities. X axis "Roundness" refers to: to what extent the responses within a cluster resemble a circle (i.e., cluster shapes in Figure 2), measured by the reversed z-score of the maximum difference in the responses. Y axis "Area" refers to: overall engagement across the 11 variables, measured by the z-score of the aggregated scores in the responses.

The question of age can also be brought into consideration for Early Career Engineers C2; those in C2 are significantly younger than C3 and C4 (Table S3.4). This leads us to speculate on whether the career path of some in C2 may lead into the C4 quadrant (by, for example, acquiring more business knowledge to round-out their professional capabilities) or into the C3 quadrant (by accumulating more engineering experiences). These possible paths are indicated by green dashes in Figure 6. Lastly, we have not focused on C1 in the study. As we analyzed earlier in section 6.2, it is possible that C1 has ventured beyond "the linear progression" from what their engineering design education offered. Though only accounting for 6.1% of the current sample, C1 could help students understand alternative career pathways that have engineering as a central focus.

7.2 Implications for student career development

Mechanical Engineering students of today and the future are confronted with some of the biggest societal problems to create a sustainable and caring future. At the same time, they are also experiencing some of the biggest advances in information and communication technologies that are eroding the boundaries of traditional disciplines and challenging the very meaning of "mechanical engineering". This study provides the first empirical evidence that there are indeed diverse engineering work profiles drawing on data of real-world professional practices. The findings are a potential source of insights for career education. Prior work shows that students can grapple with career decisions well into their senior/final year and even beyond [35], [44]. The diverse work personas concretize the dynamic ways how professionals with engineering design backgrounds identify with engineering. For instance, insights from C4 reject the simple dichotomy of engineering versus non-engineering. These findings offer new perspectives for current engineering students to contemplate how their future career might unfold.

7.3 Study limitation and future direction

The study findings should be interpreted in the context of a few limitations. First, the analysis lacks grounding in literature/theory, which is a common limitation for data-driven research [19]. We attempted to draw on the engineering epistemology [36] to map out work profiles potentially along the two axes of *people versus matter* and *practice versus theory*. Unfortunately, the data and method do not match well with this particular theory. Second, more accurate characterization of the clusters would benefit from triangulation with qualitative interview or observation data. Finally, caution must be exercised in generalizing the findings as the dataset is obtained from engineering design graduates from a particular well-regarded U.S. university. For instance, there could be significant institutional differences in career pathways [33]. Overall, the current data-driven study still illuminates diverse possibilities of career trajectories. The data-driven approach calls for more data—there has been boatloads of data collected, such as from ABET and other research efforts, yet we do not have adequate access to these valuable data sources to draw insights about alumni work profiles. Future work should also consider taking a longitudinal approach, from the first job to the most recent job [74], to gain a deeper understanding how the clustered profiles are developed over time. Relatedly, prior work also suggested teasing out other types of temporal changes, such as change of motivations, and their effects on the clusters [23].

8 Conclusion

To inform prospective career decision-making for engineering students, we have conducted a data-driven cluster analysis study to understand what different work profiles there are. The study is based on an alumni dataset from work professionals who graduated between 1987 and 2018 from a graduate-level engineering design program focusing on innovation and entrepreneurship from a U.S. university. There had been limited empirical work of this kind. Using PCA-based *k*-means cluster analysis, we analyzed the dataset to identify and subsequently validate diverse work profiles. The emergence of these diverse profiles sheds light on varied trajectories that engineering students can pursue. By extrapolating real-world practices, our study contributes valuable perspectives to inform and enhance students' strategic career planning endeavors. We call for more research to build upon our work and continue to examine how to use empirical knowledge from real-world practices to enhance students' career planning.

Acknowledgements

Our research was supported by the Hasso Plattner Design Thinking Research Program led by Professor Larry Leifer and Dr. Prof. Christoph Meinel. We thank the following individuals who were involved in designing and carrying out the research: Dr. Felix Kempf, Hung Pham, Nicole Salazar, Aya Mouallem, Johannes Lamprecht, and Stanford's School of Engineering Alumni Relations & Student Engagement. Others who provided critical feedback and background information include Niclas-Alexander Mauss, Dr. Lawrence Domingo, Dr. Jan Aurenhammer, Dr. Claudia Liebethal, Professor Helmut Schönenberger, Professor Mark Cutkosky (ME 310), Dr. Ed Carryer (ME 218), Crystal Pennywell, Gosia Wojciechowska, Elizabeth Mattson, Tammy Liaw, Sharon Nemeth and members of the Designing Education Lab (DEL) at Stanford. We gratefully acknowledge the many graduates of ME 218 and ME 310 who participated in this research project.

Reference

[1] A. Jackson and N. Mentzer, "Cluster analysis in engineering education," in *2017 ASEE Annual Conference & Exposition*, June 2017.

[2] R. Cabedo, T. C. Edmundo, and M. Castro, "A benchmarking study of clustering techniques applied to a set of characteristics of MOOC participants," in *2016 ASEE Annual Conference & Exposition*, June 2016.

[3] N. Grira, M. Crucianu, and N. Boujemaa, "Unsupervised and semi-supervised clustering: a brief survey," *A review of machine learning techniques for processing multimedia content*, vol. 1, no. 2004, pp. 9-16, 2004.

[4] D. M. Dorris, J. L. Swann, and J. S. Ivy, "A Data-Driven Approach for Understanding and Predicting Engineering Student Dropout," in *ASEE Annual Conference proceedings*, January 2021.

[5] F. Marbouti, J. Ulas, and C. H. Wang, "Academic and demographic cluster analysis of engineering student success," *IEEE Transactions on Education*, vol. 64, no. 3, pp. 261-266, March 2020.

[6] E. Kim, N. M. Hicks, and M. L. Sanchez-Pena, "Assessing the Data Analysis Training of Engineering Undergraduates," in *2019 ASEE Annual Conference & Exposition*, June 2019.

[7] D. Pistrui and N. E. Rayess, "Creating Data-Driven Undergraduate Student Engineering Typologies to Shape the Future of Work," in *2021 ASEE Virtual Annual Conference Content Access*, July 2021.

[8] A. Godwin et al., "Using topological data analysis in social science research: Unpacking decisions and opportunities for a new method," in *2019 ASEE Annual Conference & Exposition*, June 2019.

[9] A. Minichiello, J. Hood, and D. Harkness, "WIP: Methodological considerations for constructing nontraditional student personas with scenarios from online forum usage data in calculus," in *ASEE annual conference proceedings*, January 2017.

[10] M. Scheidt et al., "Engineering students' noncognitive and affective factors: Group differences from cluster analysis," *Journal of Engineering Education*, vol. 110, no. 2, pp. 343-370, February 2021.

[11] J. Chasmar and K. M. Ehlert, "Cluster analysis methods and future time perspective groups of second-year engineering students in a major-required course," in *2018 ASEE Annual Conference & Exposition*, June 2018.

[12] Q. Xu, C. Ding, J. Liu, and B. Luo, "PCA-guided search for k-means," *Pattern Recognition Letters*, vol. 54, pp. 50-55, January 2015.

[13] J. W. Herrmann et al., "Evaluating clustering algorithms for identifying design subproblems," *Journal of Mechanical Design*, vol. 140, no. 8, p. 081401, August 2018.

[14] C. Dungs, S. Sheppard, and H. Chen, "Extracurricular College Activities Fostering Students' Innovation Self-efficacy," in *Proceedings of the American Society for Engineering Education*, 2017.

[15] J. Chasmar and L. Benson, "Future time perspective and self-regulated learning: Multiple case studies in industrial engineering," in *2016 ASEE Annual Conference & Exposition*, June 2016.

[16] Q. Zhang, "Getting Tired of Massive Journal Usage Statistics: A Case Study on Engineering Journal Usage Analysis Using k-means Clustering," in *2020 ASEE Virtual Annual Conference Content Access*, June 2020.

[17] S. Maller et al., "Normative typologies of EPICS students on ABET EC Criterion 3: A multistage cluster analysis," in *2007 Annual Conference & Exposition*, June 2007.

[18] K. M. Ehlert et al., "Utilizing cluster analysis of close-ended survey responses to select participants for qualitative data collection," in *2017 ASEE Annual Conference & Exposition*, June 2017.

[19] Q. Liu and G. Evans, "Unleashing the Power of Data Analytics to Examine Engineering Students' Experiences and Outcomes," in *2020 ASEE Virtual Annual Conference Content Access*, June 2020.

[20] N. O. Ozaltin et al., "Understanding the technical entrepreneurship landscape in engineering education," in *2011 ASEE Annual Conference & Exposition*, June 2011.

[21] P. R. Norton, K. A. High, and W. Bridges, "Towards creating motivationally supportive course structures for introductory calculus," in *2020 ASEE Virtual Annual Conference Content Access*, June 2020.

[22] C. T. Schimpf and J. B. Main, "The distribution of family-friendly benefits policies across higher education institutions: a cluster Analysis," in *2014 ASEE Annual Conference & Exposition*, June 2014.

[23] A. Dillon and J. Stolk, "The students are unstable! Cluster analysis of motivation and early implications for educational research and practice," in *2012 Frontiers in Education Conference Proceedings*, October 2012.

[24] I. B. Mena, S. E. Zappe, and T. A. Litzinger, "Preparing the engineer of 2020: Analysis of alumni data," in *2012 ASEE Annual Conference & Exposition*, June 2012.

[25] K. A. Davis et al., "Characterizing students' intercultural competence development paths through a global engineering program," in *2018 ASEE Annual Conference & Exposition*, June 2018.

[26] L. F. Z. Rivera and M. M. Larrondo-Petrie, "An initial study applying data analysis and machine learning techniques to analyze dissertations and theses in the engineering education field," in *2017 ASEE Annual Conference & Exposition*, June 2017.

[27] C. C. Aggarwal, "An introduction to cluster analysis," in *Data clustering, Chapman and Hall/CRC*, 2018, pp. 1-28.

[28] L. Kaufman and P. J. Rousseeuw, "*Finding groups in data: an introduction to cluster analysis*," John Wiley & Sons, 2009.

[29] A. Nasser, D. Hamad, and C. Nasr, "K-means clustering algorithm in projected spaces," in *2006 9th International Conference on Information Fusion*, July 2006.

[30] H. Matusovich et al., "I'm graduating this year! So what is an engineer anyway?," in *2009 Annual Conference & Exposition*, June 2009.

[31] R. Korte, S. Brunhaver, and S. M. Zehr, "The socialization of STEM professionals into STEM careers: A study of newly hired engineers," *Advances in Developing Human Resources*, vol. 21, no. 1, pp. 92-113, February 2019.

[32] L. R. Lattuca, L. C. Strauss, and J. F. Volkwein, "Getting in sync: Faculty and employer perceptions from the national study of EC2000," *International Journal of Engineering Education*, vol. 22, no. 3, pp. 460, 2006.

[33] F. M. Lintl et al., "Starter or joiner, market or socially-oriented: Predicting career choice among undergraduate engineering and business students," in *2015 ASEE Annual Conference & Exposition*, June 2015.

[34] R. Stevens et al., "Becoming an engineer: Toward a three dimensional view of engineering learning," *Journal of Engineering Education*, vol. 97, no. 3, pp. 355-368, July 2008.

[35] Cynthia J. Atman et al., "Enabling Engineering Student Success: The Final Report for the Center for the Advancement of Engineering Education. CAEE-TR-10-02," *Center for the Advancement of Engineering Education*, 2010.

[36] R. Adams et al., "Multiple perspectives on engaging future engineers," *Journal of Engineering Education*, vol. 100, no. 1, pp. 48-88, January 2011.

[37] U.S. Department of Labor, "Mechanical Engineers," *Occupational Outlook Handbook*, at https://www.bls.gov/ooh/architecture-and-engineering/mechanical-engineers.htm (visited January 4, 2024).

[38] C. Striolo, M. Pollock, and A. Godwin, "Staying or leaving: contributing factors for UK engineering students' decisions to pursue careers in engineering industry," *European Journal of Engineering Education*, vol. 46, no. 3, pp. 364-388, June 2021.

[39] D. P. Dannels, "Learning to be professional: Technical classroom discourse, practice, and professional identity construction," *Journal of Business and Technical Communication*, vol. 14, no. 1, pp. 5-37, January 2000.

[40] R. W. Lent, S. D. Brown, and G. Hackett, "Toward a unifying social cognitive theory of career and academic interest, choice, and performance," *Journal of vocational behavior*, vol. 45, no. 1, pp. 79-122, August 1994.

[41] S. R. Brunhaver et al., "Not all the same: A look at early career engineers employed in different sub-occupations," in *2013 ASEE Annual Conference & Exposition*, June 2013.

[42] J. Trevelyan, "Transitioning to engineering practice," *European Journal of Engineering Education*, vol. 44, no. 6, pp. 821-837, November 2019.

[43] J. Trevelyan, "*The making of an expert engineer*," CRC Press, 2014.

[44] G. Lichtenstein et al., "An engineering major does not (necessarily) an engineer make: Career decision making among undergraduate engineering majors," *Journal of Engineering Education*, vol. 98, no. 3, pp. 227-234, July 2009.

[45] L. R. Lattuca, "Engineering change: A study of the impact of EC2000," Executive Summary, (Center for the Study of Higher Education, The Pennsylvania State University), pp. 1-20, 2006.

[46] A. Jain, "Data clustering: 50 years beyond k-means," *Pattern Recognition Letters,* vol. 31, no. 8, pp. 651-666, June 2010.

[47] C. B. Zoltowski and W. C. Oakes, "Learning by doing: Reflections of the EPICS program," *International Journal for Service Learning in Engineering, Humanitarian Engineering and Social Entrepreneurship*, pp. 1-32, 2014.

[48] T. Carleton and L. Leifer, "Stanford's ME310 course as an evolution of engineering design," in *Proceedings of the 19th CIRP design conference–Competitive design*, 2009.

[49] S. R. Barley and J. E. Orr, "*Between craft and science: Technical work in US settings*," Cornell University Press, 1997.

[50] W. G. Vincenti, "*What engineers know and how they know it*," Johns Hopkins University Press, 1990.

[51] L. L. Bucciarelli, "*An ethnographic perspective on engineering design*," Design studies, vol. 9, no. 3, pp. 159-168, July 1988.

[52] K. E. Winters et al., "From freshman engineering students to practicing professionals: Changes in beliefs about important skills over time," in *2013 ASEE Annual Conference & Exposition*, June 2013.

[53] M. M. Grau et al., "What do you want to do with your life? Insights into how engineering Students think about their future career plans," in *2016 ASEE Annual Conference & Exposition*, June 2016.

[54] T. S. Henderson, K. A. Shoemaker, and L. R. Lattuca, "Career calculus: Assessing the psychological cost of pursuing an engineering career," *Journal of Engineering Education*, vol. 111, no. 4, pp. 770-791, October 2022.

[55] D. L. Evans et al., "Attributes of engineering graduates and their impact on curriculum design," *Journal of Engineering Education*, vol. 82, no. 4, pp. 203-211, September 1993.

[56] J. L. Huff, C. B. Zoltowski, and W. C. Oakes, "Preparing engineers for the workplace through service learning: Perceptions of EPICS alumni," *Journal of Engineering Education*, vol. 105, no. 1, pp. 43-69, January 2016.

[57] S. D. Sheppard et al., "Decades of Alumni: Designing a Study on the Long-Term Impact of Design Education," in *Design Thinking Research: Achieving Real Innovation*, Springer International Publishing, 2022, pp. 247-269.

[58] C. L. Cobb et al., "What alumni value from new product development education: A longitudinal study," *Advances in Engineering Education*, vol. 5, no. 1, p. n1, January 2016.

[59] H. J. Passow and C. H. Passow, "What competencies should undergraduate engineering programs emphasize? A systematic review," *Journal of Engineering Education*, vol. 106, no. 3, pp. 475-526, July 2017.

[60] J. Rodriguez, H. L. Chen, S. Sheppard, L. Leifer, and Q. Jin, "Exploring the interest and intention of entrepreneurship in engineering alumni," in *2015 ASEE Annual Conference & Exposition*, June 2015, pp. 26-738.

[61] F. Kempf, N. Elfiki, A. Mouallem, H. L. Chen, G. Toye, M. Lande, K. Hysi, X. Ge, and S. D. Sheppard, "The Nexus of Entrepreneurship and Innovation: A new approach to looking at the creative contributions of engineering graduates," in *2023 ASEE Annual Conference & Exposition*, 2023.

[62] S. D. Sheppard, H. L. Chen, G. Toye, F. Kempf, and N. Elfiki, "Measuring the impact of project-based design engineering courses on entrepreneurial interests and intentions of alumni," *Design thinking research: Translation, prototyping, and measurement*, pp. 297-313, 2021.

[63] S. R. Brunhaver, R. F. Korte, S. R. Barley, and S. D. Sheppard, "Bridging the gaps between engineering education and practice," in *US engineering in a global economy*, University of Chicago Press, 2017, pp. 129-163.

[64] X. Ge, L. Leifer, and L. Shui, "Situated emotion and its constructive role in collaborative design: A mixed-method study of experienced designers," *Design Studies*, vol. 75, p. 101020, 2021.

[65] A. L. Zydney, J. S. Bennett, A. Shahid, and K. W. Bauer, "Impact of undergraduate research experience in engineering," *Journal of Engineering Education*, vol. 91, no. 2, pp. 151-157, 2002.

[66] G. Tryggvason, M. Thouless, D. Dutta, S. L. Ceccio, and D. M. Tilbury, "The new mechanical engineering curriculum at the University of Michigan," *Journal of Engineering Education*, vol. 90, no. 3, pp. 437-444, 2001.

[67] National Center for Science and Engineering Statistics, "Recent College Graduates - Workforce," [Online]. Available: https://ncsesdata.nsf.gov/recentgrads/2010/. [Accessed: January 4, 2023].

[68] National Center for Science and Engineering Statistics, "*Doctorate Recipients from U.S. Universities: 2021*," [Online]. Available: https://ncses.nsf.gov/surveys/doctorate-recipients/2021. [Accessed: January 4, 2023].

[69] H. M. Matusovich, R. A. Streveler, and R. L. Miller, "Why do students choose engineering? A qualitative, longitudinal investigation of students' motivational values," *Journal of Engineering Education*, vol. 99, no. 4, pp. 289-303, 2010.

[70] L. Perlow and L. Bailyn, "The senseless submergence of difference: Engineers, their work, and their careers," in *Between craft and science: Technical work in US settings*, pp. 230-243.

[71] D. A. Schön, "*The reflective practitioner: How professionals think in action*," New York: Basic Books, 1983.

[72] S. D. Sheppard, A. Colby, K. Macatangay, and W. Sullivan, "What is Engineering Practice?" *International Journal of Engineering Education*, vol. 22, no. 3, pp. 429–438, 2006.

[73] S. K. Gilmartin, H. L. Chen, M. F. Schar, Q. Jin, G. Toye, A. Harris, E. Cao, E. Costache, M. Reithmann, and S. D. Sheppard, "Designing a longitudinal study of engineering students' innovation and engineering interests and plans: The Engineering Majors Survey Project (EMS 1.0 and 2.0 Technical Report)," *Stanford University Designing Education Lab*. Stanford, CA, 2017.

[74] S. Brunhaver, R. Streveler, C. Carrico, H. Matusovich, P. Boylan-Ashraf, and S. Sheppard, "Professional engineering pathways study: A longitudinal study of early career preparedness and decision-making," in *2015 IEEE Frontiers in Education Conference (FIE)*, pp. 1-5, IEEE, October 2015.

[75] M. Schar, S. Gilmartin, A. Harris, B. Rieken, and S. Sheppard, "Innovation Self-Efficacy: A very brief measure for engineering students," in *Proceedings for the American Society for Engineering Education Annual Conference*, June 25-28, 2017, Columbus, OH.

[76] A. DeNoble, D. Jung, and S. Ehrlich, "Entrepreneurial self-efficacy: The development of a measure and its relationship to entrepreneurial action," in *Frontiers of Entrepreneurship Research*, P. D. Reynolds, Ed. Stanford, CA: Center for Entrepreneurial Studies, 1999, pp. 73–87.

[78] J. H. Dyer, H. B. Gregersen, and C. M. Christensen, "Entrepreneur Behaviors, Opportunity Recognition, and the Origins of Innovative Ventures," *Strategic Entrepreneurship Journal*, vol. 2, no. 4, pp. 317–338, 2008.

[79] M. Schar, "*Design Thinking Self-Efficacy"*, [presentation, unpublished], 2020.

[80] J. Clatworthy, D. Buick, M. Hankins, J. Weinman, and R. Horne, "The use and reporting of cluster analysis in health psychology: A review," *British Journal of Health Psychology*, vol. 10, no. 3, pp. 329-358, 2005.

[81] K. Zakharov, "Application of k-means clustering in psychological studies," *Tutorials in Quantitative Methods for Psychology*, vol. 12, no. 2, pp. 87-100, 2016.

[82] G. Norman, "Likert scales, levels of measurement and the 'laws' of statistics," *Adv. Health Sci. Educ.*, vol. 15, pp. 625-632, 2010.

[83] Hjalmarson, M. A., Herman, G. L., & Douglas, K. A. (2023). Considerations for Engineering Education Research Using Quantitative Methods. In *International Handbook of Engineering Education Research* (pp. 638-659). Routledge.

Appendix

S1 Additional literature review

S1.2 How cluster analysis is different from traditional statistical approaches

Traditional quantitative analysis methods (e.g., t tests, analysis of variance) typically examine data groupings based on hypotheses developed by the researcher. In contrast, many cluster analysis methods are unsupervised and identify data groupings solely based on the given dataset [1]. As a person-centered approach, cluster analysis is also advantageous for capturing variations in large datasets with heterogeneous observations, whereas traditional statistical techniques focus on the "average" or other central measurements, which may hide potentially important information [23] or overshadow underrepresented individuals as "anomalies" [8]. For instance, Dillon and Stolk [23] found that compared with traditional analytics that would categories students into "intrinsic" and "extrinsic" motivational categories, cluster analysis allow them to discover a large percentage of their engineering student sample simultaneously adopting a range of motivations that do not fall neatly into the conventional "intrinsic" and "extrinsic" categories.

S1.1 Cluster analysis in engineering education research

Prior work in engineering education research has compared *k*-means clustering against other methods. [11] found that *k*-means delivered best and strongest results in their study context, [18] found different methods comparable, whereas [23] endorsed different methods without choosing one over another and analyzed the nuanced differences between similar clusters derived from them. In the adjacent field of design research, [13] compared various methods for grouping subproblems in designers' problem decomposition process, where a novel hierarchical method—spectral clustering method was found to be a more accurate approach than basic hierarchical clustering for their dataset.

Some engineering education researchers advocated or developed more complicated methods. For instance, [5] used a hybrid approach to examine educational barriers based on student demographic and other background information through bisecting *k*-means clustering. This hybrid approach combines the ideas of *k*-means and hierarchical clustering, another widely applied cluster analysis method, to address potential problems caused by the random initialization of centroids in *k*-means clustering. Similarly, [17] also applied a hybrid approach combining *k*-means and hierarchical clustering, namely multistage Euclidean grouping, to profile normative typologies of students enrolled in a design program.

Despite the popularity of the above two basic clustering methods—*k*-means and hierarchical clustering, researchers have advocated newer tools. For instance, [10] applied a probabilistic clustering method that has been less used in engineering education—Gaussian mixture modeling to group the multifaceted psychosocial factors of students that may affect their holistic success.

[20] combined a neural network model, Kohonen networks, with *k*-means to categorize entrepreneurial educational programs and institutions to understand the technical entrepreneurial landscape.

Despite the variety of clustering techniques, there seems to be no consensus which technique is better in engineering education research. This is not surprising as researchers similarly observe a lack of consensus of clustering methods across research domains [46]. The lack of consensus is partly due to dataset idiosyncrasies in the different study settings and ranging research questions that researchers attempted to address with cluster analysis.

S2 Additional measures

The 11 engineering design-affiliated work activity items selected as variables for clustering analysis capture various aspects of innovative engineering product design. For this set of questions, participants are asked based on a 0-4 point Likert scale, i.e., from Never to Very Often, with a "Not Applicable" option (In data cleaning, we decided to code "Not Applicable" as "Never"): *In your current or most recent job, how often are you engaged in the following activities?*

1. Searching out new technologies, processes, techniques, and/or product ideas
2. Generating creative ideas
3. Promoting and championing ideas to others
4. Investigating and securing resources needed to implement new ideas
5. Developing adequate plans and schedules for the implementation of new ideas
6. Designing a new product or project to meet specified requirements
7. Analyzing the operation or functional performance of a complete system
8. Selling a product or service in the marketplace
9. Seeing new market opportunities for new products and services
10. Creating products that fulfill customers' unmet needs
11. Discovering new ways to improve existing products

The 8 workplace factor items included in the survey are shown below. Participants are asked based on the same Likert scale: *How often are you engaged in the following activities?*

1. Presenting your work to senior managers and leaders
2. Working on something that generates interest and feedback from others in your organization
3. Working in areas or domains that are unfamiliar to you
4. Cross-functional collaboration (working with different business units, etc.)
5. An ambiguous ill-defined problem
6. Leading or directing others
7. Documentation

8. Ethical questions or considerations

The self-efficacy measures included in the survey are shown below. All the measures below are based on a 0-4 point Likert scale. Participants are asked: *How confident are you in your ability to do each of the following at this time?*

Innovation Self-Efficacy: Innovation Self-Efficacy items are drawn from [75]. Item 4 was added to the original five-item measure [61], as it was the second highest loading item from the most significant factor (i.e., Experimenting) out of the five original factors [75].

1. Ask a lot of questions
2. Generate new ideas by observing the world
3. Experiment as a way to understand how things work
4. Actively search for new ideas through experimenting
5. Build a large network of contacts with whom you can interact to get ideas for new products or services
6. Connect concepts and ideas that appear, at first glance, to be unconnected

Entrepreneurial Self-Efficacy: The 10 Entrepreneurial Self-Efficacy items are drawn from [76]. The current scale makes up two of the six sub-dimensions of the original ESE measure, namely "developing new product and market opportunities" (item 1-7) and "coping with unexpected challenges" (item 8-10). It was decided that these two sub-dimensions were most relevant to the engineering design program from which the engineers graduated [61].

1. See new market opportunities for new products and services
2. Design products that solve current problems
3. Discover new ways to improve existing products
4. Create products that fulfill customers' unmet needs
5. Identify new areas for potential growth
6. Determine what the business will look like
7. Bring product concepts to market in a timely manner
8. Tolerate unexpected changes in business conditions
9. Work productively under continuous stress, pressure and conflict
10. Persist in the face of adversity

Design Thinking Self-Efficacy: The Design Thinking Self-Efficacy was originally developed by [79] for evaluating people's perceived confidence in design thinking. Design thinking was operationalized as a five-step process including empathize, reframe, ideate, prototype and test. The measure was based on a 0-100 slider scale and was rescaled to 0-4.

1. Sense how another person feels and what they might be thinking
2. Look at problems in the world from different angles
3. Generate a wide variety of ideas
4. Build a prototype solution that satisfies user needs
5. Accept feedback on your work and make changes
6. Enhance the lives of people by finding a better way to do things

Engineering Task Self-Efficacy: The 5-item Engineering Task Self-Efficacy was developed by [73].

1. Design a new product or project to meet specified requirements
2. Conduct experiments, build prototypes, or construct mathematical models to develop or evaluate a design
3. Develop and integrate component sub-systems to build a complete system or product
4. Analyze the operation or functional performance of a complete system
5. Troubleshoot a failure of a technical component or system

Demographic measures:

Demographic variables include gender (Male, Female, I prefer not to answer), race and ethnicity (American Indian or Alaska Native, Asian or Asian American, Black or African American, Hispanic or Latino/a, Native Hawaiian or Pacific Islander, White, Other, I prefer not to answer), and years since graduation from the design program.

S3 Additional results

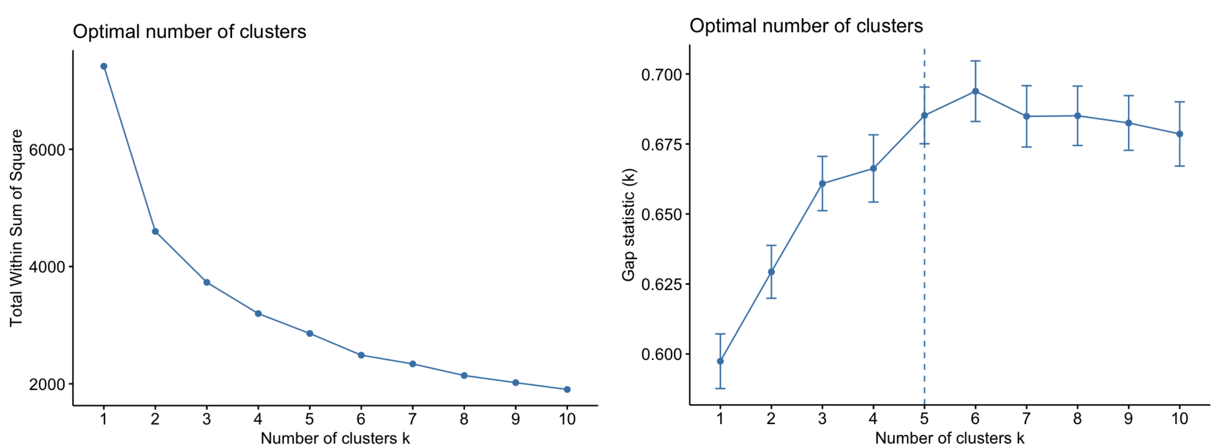

Figure S3.1. The elbow method (left) and gap statistics (right) suggest k = 5.

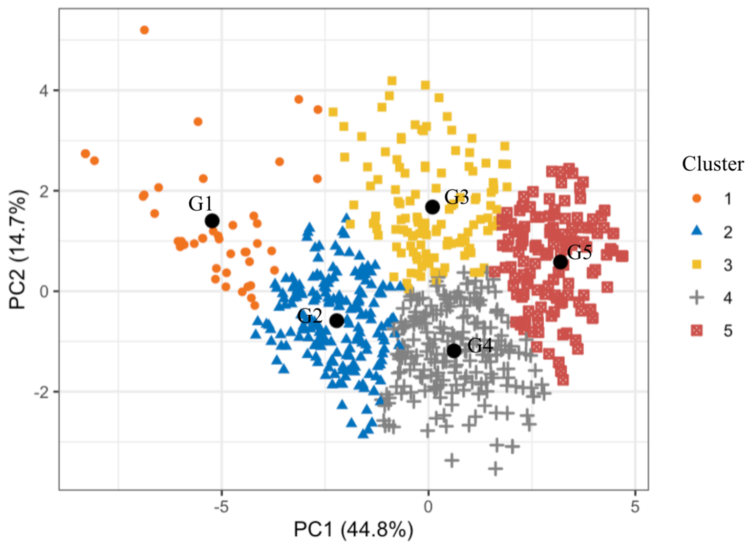

Figure S3.2. Scatter plot.

Table S3.3a. Descriptives of demographics based on five clustering results.

Demographics Weighted membership		C1 44 (%)	C2 182 (%)	C3 214 (%)	C4 125 (%)	C5 154 (%)
Gender	Female	12 (27.3%)	40 (22.1%)	48 (22.4%)	18 (14.4%)	28 (18.4%)
	Male	32 (72.7%)	141 (77.9%)	166 (77.6%)	107 (85.6%)	124 (81.6%)
Race	American Indian or Alaska Native	NA	1 (0.5%)	2 (0.9%)	NA	NA
	Asian or Asian American	14 (31.8%)	50 (27.5%)	71 (33.2%)	47 (37.6%)	61 (39.6%)
	Black or African American	1 (2.3%)	10 (5.5%)	4 (1.9%)	NA	5 (3.2%)
	Hispanic or Latino/a	6 (13.6%)	14 (7.7%)	18 (8.4%)	5 (4%)	8 (5.2%)
	Middle Eastern or North African	NA	NA	NA	NA	1 (0.6%)
	Native Hawaiian or Other Pacific Islander	NA	1 (0.5%)	NA	1 (0.8%)	NA
	White	20 (45.5%)	101 (55.5%)	116 (54.2%)	66 (52.8%)	73 (47.4%)
	Multiracial and Other	3 (6.8%)	5 (2.7%)	3 (1.4%)	6 (4.8%)	6 (3.9%)
Years since completing the design program	0-4	6 (13.6%)	28 (15.4%)	14 (6.5%)	1 (0.8%)	7 (4.5%)
	5-8	7 (15.9%)	48 (26.4%)	36 (16.8%)	10 (8%)	12 (7.8%)
	9-12	2 (1.4%)	28 (15.4%)	41 (19.2%)	18 (14.4%)	26 (16.9%)
	13-16	6 (13.6%)	13 (7.1%)	27 (12.6%)	12 (9.6%)	19 (12.3%)
	17-20	8 (18.2%)	15 (8.2%)	23 (10.7%)	15 (12%)	18 (11.7%)
	21-34	15 (34.1%)	50 (27.5%)	73 (34.1%)	69 (55.2%)	72 (46.8%)

Table S3.3b. Descriptives of demographics based on five clustering results.

Measures	Sample		C1		C2		C3		C4		C5	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Years since completing the design programs	16.81	9.04	16.25	9.08	13.64	9.26	16.12	8.75	20.77	8.43	18.53	8.15

Table S3.4: ANOVA and Tukey's HSD analysis on years since completing the engineering design innovation and entrepreneurship course(s) (design program).

Measures	ANOVA		Tukey's HSD pairwise comparison				
	F(4, 714)	***p***	**Between cluster comparison**	**Diff.**	**Lower***	**Upper***	***p* adj.**
Years since completing the design program	13.99	< .001	C4-C1	4.42	0.24	8.60	0.03
			C3-C2	2.47	0.07	4.88	0.04
			C4-C2	7.03	4.26	9.80	< .001
			C5-C2	4.88	2.27	7.50	< .001
			C4-C3	4.56	1.87	7.24	< .001
			C5-C3	2.41	-0.11	4.93	0.07

*Note: Lower and upper ends of the 95% Confidence Interval.

Table S3.5. Comparison of five clusters based on the 11 engineering design-affiliated work activities (0-4 pt. scale).

Measures	Sample		C1		C2		C3		C4		C5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1. Searching out new tech, processes, techniques, and/or product ideas	2.95	1.00	1.14	0.98	2.42	0.89	3.34	0.70	2.82	0.77	3.65	0.59
2. Generating creative ideas	3.18	0.86	1.95	1.01	2.73	0.81	3.50	0.66	2.98	0.72	3.79	0.44
3. Promoting and championing ideas to others	3.20	0.88	1.86	1.07	2.74	0.80	3.52	0.63	3.04	0.70	3.80	0.54
4. Investigating and securing resources needed to implement new ideas	2.79	1.12	1.30	1.00	2.09	1.03	3.12	0.91	2.56	0.88	3.76	0.47
5. Developing adequate plans for new idea implementation	2.94	0.97	1.41	0.92	2.56	0.89	3.24	0.80	2.60	0.68	3.69	0.57
6. Selling product/ service in the marketplace	1.43	1.37	0.36	0.89	0.38	0.60	0.81	0.69	2.49	1.10	2.97	1.02
7. Seeing new market opportunities for new products and services	1.90	1.31	0.39	0.69	0.68	0.64	1.64	0.85	2.74	0.78	3.45	0.68
8. Creating products that fulfill customers' unmet needs	2.52	1.21	0.43	0.66	1.63	0.98	2.87	0.87	2.54	0.80	3.69	0.54
9. Discovering new ways to improve existing products	2.60	1.12	0.91	0.94	1.98	1.03	2.92	0.87	2.46	0.84	3.50	0.72
10. Designing a new product/project to meet specified requirements	2.87	1.12	0.75	0.97	2.44	1.00	3.45	0.68	2.28	0.89	3.68	0.52
11. Analyzing the operation/functional performance of a complete system	2.65	1.14	0.86	0.98	2.52	0.99	3.18	0.78	1.99	1.07	3.10	1.00

Table S3.6. Comparison of five clusters in terms of self-efficacy measures and workplace factors (0-4 scale).

Measures		Sample		C1		C2		C3		C4		C5	
		Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Self Efficacy measures	*Innovation Self-Efficacy*	2.96	0.63	2.61	0.81	2.70	0.56	3.01	0.58	2.90	0.60	3.32	0.52
	Entrepreneurship Self-Efficacy	2.67	0.70	1.93	0.77	2.24	0.58	2.75	0.55	2.70	0.65	3.25	0.51
	Engineering Task Self-Efficacy	2.91	0.78	2.36	0.89	2.92	0.64	3.13	0.66	2.56	0.84	3.05	0.81
	Design Thinking Self-Efficacy	3.07	0.52	2.65	0.75	2.88	0.48	3.18	0.44	2.97	0.49	3.35	0.42
Workplace factors	*Presenting to upper mgmt*	3.16	1.01	2.05	1.24	2.77	0.91	3.49	0.77	3.16	1.03	3.50	0.95
	Doing work that elicits interest	3.30	0.83	2.25	1.08	2.97	0.84	3.56	0.63	3.22	0.73	3.69	0.61
	Exploring unfamiliar areas	3.00	0.85	2.34	1.12	2.81	0.87	3.12	0.74	2.98	0.76	3.26	0.79
	Cross-fun collaboration	3.47	0.90	2.50	1.32	3.11	1.02	3.71	0.63	3.54	0.81	3.78	0.60
	Ambiguous prob-solving	3.24	0.87	2.32	1.29	2.95	0.92	3.40	0.74	3.33	0.73	3.57	0.66
	Leading/directing	3.19	0.99	2.20	1.32	2.76	0.98	3.39	0.85	3.18	0.94	3.70	0.63
	Documentation	3.03	0.93	2.27	1.09	3.03	0.88	3.23	0.86	2.76	0.90	3.17	0.91
	Work on ethical questions	2.00	1.07	1.43	1.15	1.73	0.96	1.93	1.03	2.01	1.00	2.58	1.04

Table S3.7. Comparison of five work profile clusters based on job roles and work backgrounds. Percentages are calculated by dividing frequency of occurrence by cluster size for each of the five clusters. Participants were asked to choose all options that apply.

Job roles and work backgrounds	**Clusters**				
	C1	**C2**	**C3**	**C4**	**C5**
Design	25.0%	53.3%	69.2%	43.2%	62.3%
R&D	36.4%	64.8%	77.1%	44.0%	67.5%
Manufacturing	13.6%	27.5%	39.3%	20.0%	30.5%
Quality	9.1%	23.1%	22.0%	9.6%	18.8%
Security	0.0%	0.5%	1.9%	1.6%	3.2%
IT	2.3%	6.0%	5.6%	12.0%	11.0%
Support	9.1%	4.9%	5.6%	6.4%	5.2%
Project Management	18.2%	30.2%	40.2%	40.0%	53.9%
Functional Management	9.1%	9.9%	15.0%	27.2%	30.5%
Sales	4.5%	0.5%	0.5%	20.0%	13.6%
HR	2.3%	0.0%	1.9%	5.6%	7.1%
Education	15.9%	9.9%	6.5%	6.4%	9.7%
Public Relations	4.5%	0.5%	0.5%	14.4%	16.9%
Legal	0.0%	1.1%	0.9%	4.0%	5.2%
Finance	2.3%	1.1%	0.9%	13.6%	7.1%
Work at mid-large company	31.8%	45.6%	56.5%	40.8%	33.8%
Obtained MBA	2.3%	1.6%	4.7%	23.2%	16.9%
Entrepreneur/Founder	6.8%	2.2%	4.2%	23.2%	30.5%
Others	11.4%	7.7%	6.1%	15.2%	11.0%

Table S3.8. ANOVA and Tukey's HSD analysis of self efficacy measures (0-4 scale), focusing on results of the four clusters C2 to C5.

Measures	ANOVA		Tukey's HSD pairwise comparison				
	F(4, 714)	*p*	Between cluster comparison	Diff.	Lower*	Upper*	*p* adj.
Innovation Self-Efficacy	28.37	< .001	C3-C2	0.31	0.15	0.47	< .001
			C4-C2	0.20	0.01	0.38	0.03
			C5-C2	0.62	0.45	0.80	< .001
			C5-C3	0.31	0.14	0.48	< .001
			C5-C4	0.42	0.23	0.62	< .001
Entrepreneurship Self-Efficacy	82.53	< .001	C3-C2	0.51	0.35	0.67	< .001
			C4-C2	0.46	0.28	0.65	< .001
			C5-C2	1.02	0.84	1.19	< .001
			C5-C3	0.51	0.34	0.67	< .001
			C5-C4	0.55	0.36	0.75	< .001
Engineering Task Self-Efficacy	19.47	< .001	C3-C2	0.21	0.01	0.42	0.04
			C4-C2	-0.37	-0.60	-0.13	< .001
			C4-C3	-0.58	-0.81	-0.35	< .001
			C5-C4	0.49	0.25	0.73	< .001
Design Thinking Self-Efficacy	32.78	< .001	C3-C2	0.29	0.16	0.43	< .001
			C5-C2	0.47	0.33	0.61	< .001
			C4-C3	-0.21	-0.35	-0.06	< .001
			C5-C3	0.18	0.04	0.31	0.01
			C5-C4	0.38	0.22	0.54	< .001

*Note: Lower and upper ends of the 95% Confidence Interval.

Table S3.9. ANOVA and Tukey's HSD analysis on workplace factors (0-4 pt. scale), focusing on C2 to C5.

Measures	**ANOVA**		**Tukey's HSD pairwise comparison**				
	F(4, 714)	***p***	**Btw cluster comparison**	**Diff.**	**Lower***	**Upper***	***p* adj.**
Presenting to upper mgmt	35.81	< .001	C3-C2	0.72	0.46	0.97	< .001
			C4-C2	0.39	0.10	0.69	< .001
			C5-C2	0.73	0.45	1.01	< .001
			C4-C3	-0.33	-0.61	-0.04	0.02
			C5-C4	0.34	0.03	0.65	0.02
Doing work that elicits interest	49.34	< .001	C3-C2	0.58	0.38	0.79	< .001
			C4-C2	0.25	0.02	0.49	0.03
			C5-C2	0.72	0.50	0.94	< .001
			C4-C3	-0.33	-0.56	-0.11	< .001
			C5-C4	0.47	0.23	0.71	< .001
Exploring unfamiliar areas	14.83	< .001	C3-C2	0.31	0.09	0.54	< .001
			C5-C2	0.45	0.21	0.70	< .001
			C5-C4	0.28	0.01	0.54	0.04
Cross-fun collaboration	34.13	< .001	C3-C2	0.60	0.37	0.83	< .001
			C4-C2	0.43	0.16	0.69	< .001
			C5-C2	0.67	0.42	0.92	< .001
Ambiguous prob-solving	28.73	< .001	C3-C2	0.45	0.23	0.68	< .001
			C4-C2	0.38	0.12	0.64	< .001
			C5-C2	0.62	0.38	0.86	< .001
Leading/directing	39.08	< .001	C3-C2	0.63	0.38	0.88	< .001
			C4-C2	0.43	0.14	0.71	< .001
			C5-C2	0.94	0.67	1.21	< .001
			C5-C3	0.31	0.05	0.57	0.01
			C5-C4	0.52	0.22	0.81	< .001
Documentation	14.27	< .001	C4-C3	-0.47	-0.74	-0.19	< .001
			C5-C4	0.41	0.11	0.70	< .001
Work on ethical questions	19.47	< .001	C5-C2	0.85	0.55	1.16	< .001
			C5-C3	0.65	0.35	0.94	< .001
			C5-C4	0.57	0.24	0.90	< .001

*Note: Lower and upper ends of the 95% Confidence Interval.