Proceedings of the ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference
IDETC/CIE 2021
August 17 – 20, 2021

IDETC2021-70909

PREDICTING A PARADIGM SHIFT: EXPLORING THE RELATIONSHIP BETWEEN COGNITIVE STYLE AND THE PARADIGM-RELATEDNESS OF DESIGN SOLUTIONS

Courtney Cole
Department of Industrial and Manufacturing Engineering
The Pennsylvania State University,
University Park, PA, USA
Email: cmc6503@psu.edu

Jacqueline Marhefka
Department of Psychology
The Pennsylvania State University
University Park, PA, USA
Email: jtm40@psu.edu

Kathryn Jablokow
School of Engineering Design
The Pennsylvania State University
University Park, PA, USA
Email: KWL3@psu.edu

Susan Mohammed
Department of Psychology
The Pennsylvania State University,
University Park, PA, USA
Email: sxm40@psu.edu

Sarah Ritter
School of Engineering Design
The Pennsylvania State University
University Park, PA, USA
Email: scr15@psu.edu

Scarlett Miller
School of Engineering Design
The Pennsylvania State University,
University Park, PA, USA
Email: scarlettmiller@psu.edu

ABSTRACT

Nearly 60 years ago, Thomas Kuhn revolutionized how we think of scientific discovery and innovation when he identified that scientific change can occur in incremental developments that improve upon existing solutions, or it can occur as drastic change in the form of a paradigm shift. In engineering design, both types of scientific change are critical when exploring the solution space. However, most methods of examining design outputs look at whether an idea is creative or not and not the type of creativity that is deployed or if we can predict what types of individuals or teams is more likely to develop a paradigm-shifting idea. Without knowing how to identify who will generate ideas that fit a certain paradigm, we do not know how to build teams that can develop ideas that better explore the solution space. This study provides the first attempt at answering this question through an empirical study with 60 engineering design student teams over the course of a 4- and 8-week design project. Specifically, we sought to identify the role of cognitive style using KAI score, derived from Kirton's Adaption-Innovation (A-I) theory, on the paradigm-relatedness of ideas generated by individuals and teams. We also sought to investigate the role of crowdsourcing for measuring the paradigm-relatedness of design solutions. The results showed that KAI was positively related to a greater likelihood of an individual's idea being categorized as paradigm-breaking. In addition, the team KAI diversity was also linked to a greater likelihood of teams' ideas being categorized as paradigm-challenging. Finally, the results support the use of crowdsourcing for measuring the paradigm-relatedness of design solutions.

Keywords: design theory and methodology, design theory, decision making

INTRODUCTION

Thomas Kuhn revolutionized the way that we think about scientific discovery and innovation nearly 60 years ago in his landmark book, *The Structure of Scientific Revolutions* [1]. It was there that Kuhn defined two different types of scientific change: incremental developments or "normal science" and scientific revolutions that involve the ever-evasive "paradigm shift" [1], *"an important change that happens when the usual way of thinking about something is replaced by a new or different way"*[2]. In other words, incremental developments often lead to refined versions of existing solutions that excel by performing better in their primary or a related context [3-6], while paradigm shifting ideas lead to radical changes that allow us to approach a problem from unexpected angles or to connect concepts that at first seem unrelated [3-6]. In recent times, Kuhn's work has been commended for suggesting that these two creative problem-solving perspectives do not just coexist, but are interrelated and should be considered in combination [7]. In a design context, Kuhn's work is particularly influential when design problems are wicked problems [8], or problems that are societal and less structured, and the information needed to understand the problem depends on generating a vast array of ideas [9]. While we recognize the importance of both incremental and radical ideas, we focus on Kuhn's emphasis on radical, or "paradigm shifting" [1] ideas, for now.

By introducing the concept of paradigms, Kuhn encouraged a paradigm shift of his own by pushing the scientific community to tackle problems of various kinds in ways beyond typical methods [10, 11]. In Kuhn's book, he stated, "Under *normal conditions* the research scientist is not an *innovator* but a solver of puzzles, and the puzzles upon which he concentrates are just those which he believes can be both stated and solved within the existing scientific tradition," [1] (p. 170). When faced with a challenge outside of such *normal conditions*, Kuhn postulated that there may be some underlying attributes or experiences of people who are more likely to develop these radical ideas, foreshadowing (while not explicitly proposing) the concept of an individual's *cognitive style*. Specifically, he suggested that people who come up with paradigm shifting ideas are typically young or very new to the field whose paradigm they change – or people who are not committed by prior practice to the traditional rules of "normal science," [1]. In this way, he was describing former and future game-changing scientists like Tesla, who pioneered new disciplines, fueled innovations in the modern electrical industry [12], and more recent innovators like Steve Jobs, whose ideas transformed the computer industry into what it is today [13]. But, was it really their youth or naïveté that made them paradigm shifters?

If we know paradigm shifts are vital to technical discoveries, then a key question is: are there characteristics of individuals (or teams) that can predict paradigm-shifting ideas? One trait that may impact a designer's tendency to develop paradigm-shifting ideas is their cognitive style. Roughly a decade after Kuhn's book was published, Kirton's Adaption-Innovation (A-I) theory was validated, pointing to *cognitive style* as a factor that can influence the types of ideas and solutions a person generates through that individual's innate cognitive preference for structure [14]. Here, "A-I theory" refers to the theory itself and not the metric that is derived from it. Using A-I theory, an individual's cognitive style falls somewhere within the range of highly adaptive (i.e., strongest preference for structure) to highly innovative (i.e., weakest preference for structure) [15]. In practice, more innovative individuals are less structured thinkers who tend to approach tasks from unsuspected angles, challenge problem constraints, and are more disruptive risk-takers [15, 16]. In contrast, more adaptive individuals are more structured thinkers who refine current systems, focus on precision, reliability, and efficiency, and engage in prudent risk-taking [15, 16]. A-I theory is based on the assumption that all individuals, of all cognitive styles, are creative [15], which dovetails nicely with Kuhn's supposition that both normal science and paradigm shifts are necessary for science to progress [1].

To support its practical use in context, the Kirton Adaption-Innovation Inventory (KAI), which was derived from A-I theory [14], has been validated for the general population and for other sub-groups, including engineers [14, 15, 17]. Specifically, KAI has been used extensively in engineering design research [18-24], where cognitive style was shown to significantly predict "creative idea generation" using a rating method known as Consensual Assessment Technique (CAT) [25]. However, the CAT dismisses some ideas as "not creative," limiting a full interpretation of the data through the A-I theory lens. While relatively new, *paradigm-relatedness* as a rating technique overcomes this limitation by mirroring A-I theory and supporting a more diverse definition of creativity [26, 27].

Researchers first defined *paradigm-relatedness* as a measure of an idea's creative style, "independent of and orthogonal to the creativity level" [28] (p. 89). The concept was taken further by defining categories of *paradigm-relatedness* based on the elements, relationships, and focus of a design concept [29]. Although it can be more difficult to achieve high interrater-reliability when breaking up paradigm-relatedness into components such as *elements*, *relationships*, *constraints*, and *focus* [27], a *category-based* (which involves separating ideas into one of a few broad categories) metric approach is still recommended for assessing large sets of ideas, because it is faster to apply and more reliable [27].

In addition to exploring cognitive style and paradigm-relatedness at the individual level, it is also important to analyze the impact at the team level as well. This is important because the path to creative results is less clear at the team level, and there is much debate over *how* to promote team creativity [30, 31] due to the complex dynamics of teams [32-34]. Specifically, when team members' cognitive styles are diverse, cognitive gaps are created. A team can leverage this style diversity by approaching problems using different perspectives, or they can succumb to conflicts that disrupt the team's efforts [15]. Cognitive gaps can be measured in different ways, including the standard deviation of a team's KAI score distribution [22] (referred to as *cognitive diversity* [35]). Cognitive style can also be measured at the team level through average measures (e.g., a team's average KAI score). Team research shows that computing the average of team members' scores (referred to as *cognitive elevation* [35]) can be viewed as a collective value that represents the team as a whole [36, 37], as additive aggregation models assume that all team members' scores should be equally represented (e.g., [38]). Despite prior team research using KAI and the assessment of design solutions, the impact of cognitive style on the paradigm-relatedness of design outcomes remains unclear and largely uninvestigated. Understanding this impact is important, because the diversity of strategy and approach in generating both incremental and radical ideas within a team can help teams explore a wider solution space, and thus increase the potential for a successful design [39].

To be able to make such comparisons between KAI and paradigm-relatedness, or any kind of rating technique, human raters must be recruited. This can be costly in terms of time to evaluate hundreds or thousands of concepts [40], which makes it especially difficult to gather expert raters due to time constraints [41]. To alleviate this burden of manpower, crowdsourcing is a technique that has been used to divide the workload among many individuals [42], allowing results to be gathered quickly and at a lower cost [43]. Specifically, crowdsourcing has seen success in the social sciences [44], especially in various studies that focus on the evaluation of ideas [45-49]. However, paradigm-relatedness has only been applied in settings where trained raters rate ideas in-person [26, 27]. As such, it brings to question if crowdsourcing is a reliable method for rating the paradigm-relatedness of ideas in engineering tasks.

The objective of this paper was to explore the relationship between cognitive style and the paradigm-relatedness of design outputs during the *concept generation* phase of the design process. Specifically, we sought to understand this impact at both the individual and team level. In addition, we sought to understand if crowd-sourcing could be used to measure the paradigm relatedness of design solutions.

RELATED LITERATURE

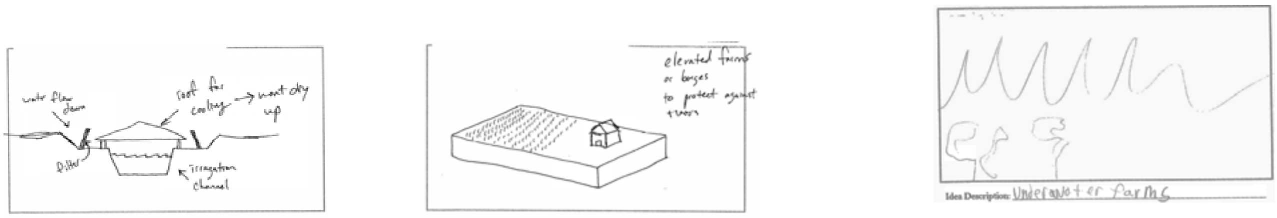
The engineering design process can be simplified into three phases consisting of generation, evaluation (e.g., concept screening), and communication [50-52]. During *concept generation*, team members are encouraged to produce creative ideas, or ideas that are both novel and useful [53]. This stage is critical to overall performance, as the availability of creative ideas is a precursor to evaluation and part of the formula for pushing innovation [54]. Concept generation practices tend to be dependent on individuals' background knowledge of the problem [55, 56], which is why engineering design students are given time to become familiar with their design prompt. Despite knowing such information about creative idea generation on the surface, the cognitive mechanisms behind the paradigm-relatedness of individuals' design solutions have seen limited exploration [28, 29, 57, 58], particularly in engineering design.

Kirton's Adaption-Innovation (A-I) Theory

Since its inception, the influence of Kirton's Adaption-Innovation (A-I) theory grew from a 32-item inventory (i.e., KAI) that could characterize an individual's preferences for problem solving [14] to its use in predicting constructs such as online discussion behaviors [59] and creative outputs [60]. Kirton's A-I theory and the KAI inventory are both built on the key assumption that there are people who prefer "to do things better" (Adaption) and those who prefer "to do things differently" (Innovation), and are both creative (p. 622) [14]. Varying amounts of adaption and innovation can be beneficial, depending on the problem-solving scenario. In this context, Adaption and Innovation exist on opposite sides of continuous spectrum of *cognitive style*, which is defined as the stable, characteristic cognitive preference that describes how people seek or respond to change [61]. Cognitive style differs from cognitive level, which defines an individual's capacity for engaging in problem-solving and creative behavior [15]. When generating solutions to problems, it is noted that innovators tend filter their ideas less, stretch the problem space boundaries, and rely less on group cohesion. Conversely, adaptors tend to screen their ideas more carefully, explore thoroughly inside the problem space boundaries, and promote group cohesion [18].

In terms of precise measurement, an individual's KAI score falls somewhere within the range of 32 (highly adaptive) to 160 (highly innovative) [15]. This total score is further broken down into three inter-related sub-scores that correspond to three sub-factors of cognitive style, namely: Sufficiency of Originality (SO), Efficiency (E), and Rule/Group Conformity (R/G) [14]. Specifically, adaptive individuals tend to offer highly detailed ideas that improve upon existing solutions and adhere to the problem definition, whereas innovative individuals tend to offer ideas that challenge the problem statement and solve problems more loosely with less details [15, 61, 62]. Any one of these three sub-factors could impact concept generation; prior research shows that individuals with higher SO and R/G sub-scores tended to perceive their ideas as less diverse when working in a team versus working alone. Conversely, those with more adaptive SO and R/G sub-scores perceived their ideas as more diverse when working with someone else [18].

When individuals of different KAI scores are placed in a team, diversity between individuals' cognitive styles and/or levels increases and the cognitive gaps grow [39]. Here, it is important to make the distinction between cognitive style and cognitive level. Cognitive style is an individual's stable, characteristic cognitive preference for managing structure in problem solving, while cognitive level describes an individual's cognitive capacity to solve problems and demonstrate creativity. Cognitive level is assessed through measures of both potential capacity (e.g., intelligence and talent) and manifest capacity (e.g., knowledge and skills) [15]. The just-noticeable difference between individuals' cognitive styles occurs at 10 points on the KAI scale [63, 64]. Gaps of 20 points or more can cause conflict in the form of poor communication, blaming one another, and misinterpreting differences in cognitive style as incompetence, for example [39]. Although failing to address conflict could diminish performance, as teams spend more time trying to figure out how to deal with one another, rather than coming up with solutions to the problem itself [15, 65], coping behavior can mitigate these effects [64, 66]. Specifically, coping behavior is a mechanism used by individuals to deal with the negative impacts of cognitive gap by adjusting one's behavior to solve problems in a way that is not consistent with their preferred style [15]. Such effects of coping behavior on actual problem-solving behavior have been expressed in prior literature, where the context (such as class or team) can impact how individuals manipulate their coping behavior [15, 18, 39]. Additionally, other forms of team communication can also



Irrigation channel (paradigm-consistent) | Elevated farms or barges to protect against floods (paradigm-challenging) | Underwater farms (paradigm-breaking)

Figure 1: Example of ideas categorized as paradigm-consistent, -challenging, or -breaking that satisfy the design challenge, "'Extreme weather conditions cause issues with farming in the Philippines. El Nino and La Nina (long dry seasons, long seasons of rainfall and flooding) is the major cause of this.' The end users being designed for are poor farming families in the Philippines."

impact how individuals project their behavior, as some individuals may have a greater impact on team outcomes through being dominating or charismatic, or may not say much at all and conform to group norms [67].

Measuring Paradigm Shifting Ideas

To identify whether underlying individual and team cognitive characteristics like cognitive style impact the creation of paradigm shifting ideas, we must first identify how to measure paradigm shifting ideas. While there is a wealth of creativity measurements available, here we focus on two different design rating methods: The Consensual Assessment Technique [68, 69] and Paradigm-Relatedness [26, 27].

One of the most widely used, albeit imperfect method for measuring design creativity is the Consensual Assessment Technique (CAT) [68, 69]. The underlying premise behind the CAT is that something is creative to the extent to which experts in the field agree, independently, that it is creative [68]. Additionally, to ensure high interrater reliability, it is standard practice to complete a practice set of ratings, and then the raters work independently to rate their assigned ideas separately. While the CAT is supported by over 30 years of research and is used extensively in the social science community [41], it also requires that raters be experts or quasi-experts (novice idea raters) in the domain [68, 69]. Specifically, quasi-expert raters must be trained by experts [70], which can be difficult when evaluating a large number of ideas across various domains. Issues with CAT are further complicated when observed from a global assessment of creativity (see [71, 72]), as a recent study showed lack of significant agreement on global ratings of creativity by experts [41]. This issue is further complicated with novice raters, as interrater reliability is typically lower [68, 73], and research has shown low correlation between the ratings of experts and novices [74]. Another issue that can occur is that the CAT can yield a negative relationship with other idea rating techniques for novelty, such as the Shah, Vargas- Hernandez, and Smith (SVS) method [41]. This implies that CAT does not necessarily yield results similar to other creativity rating schemes. Finally, and perhaps most problematically, this technique of rating creativity dismisses some ideas as "not creative," which directly contradicts A-I theory and limits the interpretation and use of KAI scores to predict the paradigm-relatedness of solutions.

In the last few years, the engineering community has made significant strides towards developing new rating methods that allow us to consider ideas on a continuum from incremental to radical [27], fully capturing the ideas behind Thomas Kuhn's pivotal work. Thus, the *paradigm-relatedness* creativity rating technique was developed within an engineering design context to evaluate design ideas [26, 27]. Specifically, a *paradigm* refers to the "ways of perceiving or acting in response to a situation or problem," (p. 31), whereas *relatedness* refers to "the extent that an idea operates within" (p. 31) or challenges that paradigm [27]. The first category used in this technique, *paradigm-consistent*, describes a solution that resembles an already existing, common design that stays within the problem constraints. The second category, *paradigm-challenging*, either integrates an uncommon element or relationship into the solution and begins to stretch the problem boundaries. The third and final category, *paradigm-breaking*, shifts the focus of the problem to a larger problem while violating some or all relevant constraints [27]. Examples of these are shown in Figure 1. It is important to note that no one category is better than any other; while some people mistakenly associate only radical ideas with higher levels of creativity [75], incremental ideas are creative as well [60, 76]. The distinction between these types of ideas is important in considering different types of creativity as identifying such differences allows us to generate a variety of ideas that more fully explore the problem space [77, 78]. Unlike those using the CAT, raters using the paradigm-relatedness technique are primed with the problem definition and where it will be used to achieve acceptable levels of interrater reliability [26, 27]. When using just a category-based approach, this can make it easier to rate larger quantities of ideas with higher interrater reliability [27]. Although it remains to be investigated for its relationship with KAI scores, the foundation of the paradigm-relatedness rating technique was based on concepts derived from A-I theory [26], which leads us to question how these two variables are related.

Even with the potential benefits of paradigm-relatedness, how can ideas be efficiently and effectively rated when there are hundreds or thousands of ideas rated generated for a given study? One way to scale design ratings is through the use of crowdsourcing – or by obtaining ratings from an (unknown) large online community [45, 79]. Crowdsourcing provides an economically and intellectually beneficial method of dividing the workload among many individuals [42, 43]. As studies show that crowdsourcing reduces demand characteristics due to a lack of face-to-face interactions [80], we see that crowdsourcing is a less intrusive method for gathering idea ratings compared to the laboratory or field environment. Additionally, crowdsourced workers represent the U.S. population relatively well [81], which is helpful when English proficiency is needed. However, crowdsourcing in idea evaluation comes with a caveat – i.e., achieving accuracy in evaluation among non-experts [45, 82]. In other words, you must for validate that the 'wisdom of the crowd' is an appropriate method for evaluating your design ideas. This is particularly important in engineering design, as previous research has shown that the 'who' does the evaluation part of the equation matters [41, 74]. Additionally, this can be problematic when using non-experts to rate concepts that are not within the scope of common knowledge of a demographic [83], and experience of the crowdsourced worker can complicate things, as inexperienced workers are more likely to misunderstand task requirements [84]. However, *paradigm-relatedness* does not require expert raters in a domain to perform the ratings and it has been shown to lead to high interrater reliability [26, 27]. Furthermore, *paradigm-relatedness* allows us to view the creativity of ideas on a continuum [27], which is critical when considering the benefits of incremental and innovative ideas when exploring the solution space [39]. As such, crowdsourcing - which provides a relatively low-cost and efficient method for rating large quantities of ideas – may be an effective means of rating the paradigm-relatedness of design ideas. However, research is needed to validate the validity of crowdsourcing for measuring paradigm-relatedness.

RESEARCH OBJECTIVES

The goal of this paper was to explore the relationship between cognitive style and the paradigm-relatedness of design outputs and to identify whether crowdsourcing could be used to measure paradigm-relatedness. Specifically, the following research questions (RQ) were explored:

RQ1: Can crowdsourcing be used to rate the paradigm-relatedness of design ideas? We hypothesized that crowdsourced workers employed by Amazon Mechanical Turk would be able to rate the paradigm-relatedness of design solutions when given the context of the idea. This is based on prior research that has found success in using crowdsourcing for rating design solutions [45-49]. Specifically, if we provide all workers with some context as to how to rate ideas according to paradigm-relatedness and the problem statements that describe each design challenge, we would be able to gather usable results as defined by the ICC values in [85]. This is a critical first step before conducting any other analyses with the crowdsourced data.

RQ2: Can an individual's cognitive style be used to predict the paradigm-relatedness of their design solutions? We hypothesized that higher cognitive style (*innovative* trend) would predict a greater likelihood for *paradigm-breaking* solutions, whereas lower cognitive style (*adaptive* trend) would predict a greater likelihood for solutions that are *paradigm-consistent*. This hypothesis was based on the fact that the paradigm-relatedness scale was developed based on cognitive style [26, 27], and KAI is representative of cognitive style [14, 15]. As such, we would expect to see KAI predict the likelihood of such design solution paradigms.

RQ3: Can the elevation or diversity of team cognitive style be used to predict the paradigm-relatedness of a team's design solutions? We hypothesized that higher team cognitive style *elevation* (average) [35] would predict greater likelihood for paradigm-breaking ideas, whereas teams with lower cognitive style elevation would predict a greater likelihood for paradigm-consistent ideas. This is based on prior work used to develop the paradigm-relatedness scale [26, 27], where cognitive style as defined by KAI [14] is thought to impact the paradigm-relatedness of solutions at the individual level. Through aggregating KAI scores to the team level, we can view cognitive style under a team construct lens, as mean values have been used to refer to members' attributes as a collective value [36, 37]. We also hypothesized that team cognitive style *diversity* (standard deviation) [35], which has been used in prior work in engineering design [22], would predict a greater likelihood for an even spread of ideas across the paradigm-relatedness bins. Particularly, prior work on cognitive gap emphasizes that complex problems require diversity of strategy and approach for generating both incremental and radical ideas [15, 39], which could be fostered by a larger cognitive gap among team members. Additionally, this could be shown by a greater likelihood of a team's ideas being categorized as paradigm-challenging, as this characterizes ideas that are contain incremental and radical elements [27]. Finally, we hypothesized that heterogeneous teams, as defined by individuals with KAI scores more than 10 points apart, would also predict a greater likelihood for an even spread of ideas across the bins, whereas homogeneous teams would not, based on prior literature [15].

METHODOLOGY

To answer the research questions, an empirical study was conducted at a large northeastern university over the first project of a cornerstone engineering design course over five semesters. Further study details and the experimental design are presented in the remainder of this section.

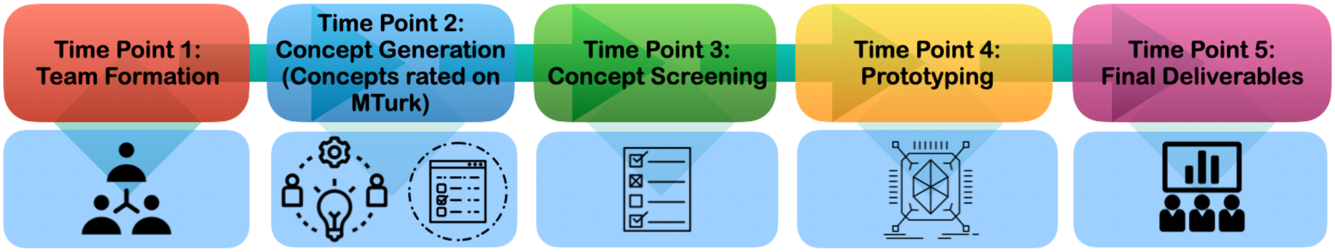


Figure 2: Study timeline – engineering design outputs were gathered at each time point (total time period: 8 weeks for Fall/Spring, 4 weeks for Summer); focus for this study is on Time Point 2

Participants

Sixty engineering design student teams, comprised of 213 participants (152 males and 61 females), participated in the study. All participants were enrolled in a first-year cornerstone engineering design class at a large northeastern university. Table 1 outlines the participants and their distributions over the data collection periods.

Procedure

The study was completed over the course of five semesters, with ten sections of a first-year engineering design course, see Table 1. Six of these courses took place over the course of a typical semester (15 weeks), and four occurred over a condensed summer session (6 weeks) (see Table 1). The course schedule remained consistent across all sections, where the same design outputs were gathered from the same five time points (see Figure 2) in all sections, following the same course schedule presented in [86]. All participants consented at the beginning of the study based on the Institutional Review Board guidelines established at the university. The remainder of this section emphasizes the methodologies used as part of the current investigation.

After consent was attained, at ***Time Point 1***, participants completed the 32-item KAI inventory to obtain a numerical representation of their cognitive styles—i.e., their individual KAI score. From there, 3- and 4-member teams were formed based on KAI. Specifically, approximately half of the teams were constructed to be homogeneous (team KAI scores within a 10-point range) while the other half were constructed to be heterogeneous. Next, newly-formed teams were given a design challenge within one of the categories presented in Table 1, which varied by instructor/semester. Then, teams researched the context of their design problem.

TABLE 1: DESCRIPTIONS OF DESIGN CHALLENGES BASED ON INSTRUCTOR AND SEMESTER

Project	Semester	Instructor	N
Tackle food insecurity in developing countries as a result of climate, conflict, unstable markets, food waste, and lack of investment in agriculture.	SU 2018	A	46 students; 12 teams
Ensure healthy lives and promote the well-being for all at all ages through addressing diseases, pollution, and traffic injuries.	SP 2019	A and B	41 students; 12 teams
	SU 2019	A	44 students; 12 teams
	FA 2019	A	26 students; 7 teams
	SP 2020	A and C	56 students; 15 teams

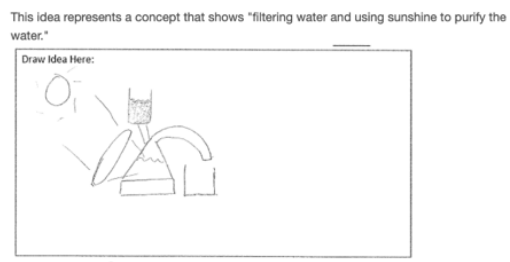


Figure 3: Example of survey interface when rating ideas. The participants were asked which category the idea best represented: paradigm-consistent, -challenging, or -breaking using radio buttons.

During ***Time Point 2***, students attended the same series of lectures in [86], helping them to form their problem statements. Next, the participants engaged in a 15-minute concept generation session with the goal of individually sketching as many ideas as possible. Then, teams joined together to combine their ideas and sketch new ones during a group brainstorming session. At the end the session, the instructor collected the ideas, and digital copies were scanned for analysis.

Design Ratings via Crowdsourcing

To evaluate the 1,395 ideas generated during the study, we recruited 600 from Amazon Mechanical Turk (MTurk) workers according to procedures set forth by the Institutional Review Board. Specifically, each MTurk worker was recruited and provided with a link to a Qualtrics survey where they were randomly assigned one team's ideas to rate. During the survey, the MTurk worker was provided information on paradigm-relatedness as well as an example of design ideas from a different task. They were then presented with the design prompt from the team, where all of the ideas were displayed on the same page of the survey. The order in which solutions were presented to each rater were randomized. Then, they were asked to evaluate the same ideas being shown based on the paradigm-relatedness construct [26, 27], where they binned an idea into one of three categories: paradigm-consistent, paradigm-challenging, or paradigm-breaking, see Figure 3. We provided definitions of these categories to the participants at the top of each page: *paradigm-consistent*, "ideas that adapt existing versions of a product or idea, but remain relatively similar to the standard solution", *paradigm-breaking* ideas "challenge the standard solution and lean towards breaking constraints for the sake of being inherently different", and paradigm-challenging falls between these two bins, where it is a "combination of paradigm-consistent and paradigm-breaking." Because the ideas were all on the same survey page, they could change their ratings. Finally, each rater was provided with a random code at the end of the survey. The survey took 8.45 ± 4.85 minutes to complete and workers were paid $2 for participating in the study. Each team was rated by 7-11 workers.

METRICS

To measure the impact of cognitive style on the paradigm-relatedness of the ideas generated, KAI was the primary metric used. Because paradigm-relatedness is based on rating each concept as one of the three categories discussed in the "Related Literature" section, we turn our focus to KAI.

Kirton Adaption-Innovation Inventory (KAI)

To measure the impact of cognitive style on the paradigm-relatedness of design solutions from concept generation, the KAI inventory was used to assess problem-solving preferences at both the individual and team level [14]. In this study, the total KAI score, or the sum of the three sub-scores, was used in the analyses. Furthermore, because KAI scores fall on a continuous scale, comparisons of cognitive style are relative, such that lower KAI scores correspond to an individual having a more adaptive cognitive style, whereas individuals with a higher KAI score have a more innovative cognitive style [15]. A certified KAI practitioner collected and scored the student responses at the beginning of each semester that this study was conducted, and the participants received feedback on their results.

The KAI total scores of the 213 engineering design students in this study showed a normal distribution, with the scores ranging from 63 to 127 (M=91.82, SD=13.76). KAI scores follow a normal distribution when analyzed across large general populations (across cultures) within a theoretical range of 32 to 160, with an observed range of 43 to 149 (M=95, SD=17) [18].

RESULTS AND DISCUSSION

During the study, 213 individuals generated an average of 6.31±2.55 ideas. These individuals composed of 60 engineering design teams generated an average of 23.23±7.92 ideas per team. The remainder of this section presents the results in reference to our research questions. The statistical data were analyzed via the SPSS v.27. A value of $p < .05$ was used to define statistical significance [87].

RQ1: Can crowdsourcing be used to rate the paradigm-relatedness of design ideas?

The objective of our first research question was to examine if crowdsourcing could be used to reliably evaluate the paradigm relatedness of engineering design solutions. To answer this research question, we examined the degree to which the 600 MTurk workers provided similar paradigm-relatedness ratings of the 1,395 total ideas generated by study participants, also known as interrater reliability. As a reminder, MTurk workers were each assigned to one team each and were asked to rate all of the ideas within that team. Ratings took MTurkers approximately 10 minutes to complete and data from the 600 raters was completed in approximately 2 hours.

To determine the inter-rater agreement, intraclass correlation coefficient (ICC(2)) were calculated. As noted by LeBreton and Senter [85] ICC(2) represents an understanding of "the extent to which the mean rating assigned by a group of judges is reliable" (p. 824). The values range from 0 – 1 with zero being low (or no) reliability and 1 being perfect reliability or consistency in ratings. Values above .7 are typically considered acceptable for assessing consistency across judges (i.e., judges are rating things similarly) [88, 89]. Our results found that the MTurk workers provided similar ratings of paradigm-related to one another as judged by ICC(2) = 0.699. In addition, our results showed that ICC(1) = 0.189, which is considered a medium effect (around ICC(1)=.10 is considered as such) [85]. These results support the use of MTurkers for evaluating engineering design tasks according to paradigm relatedness, as we found that the ICC values meet the criteria for consistency among raters [85].

RQ2: Can an individual's cognitive style score be used to predict the paradigm-relatedness of their design solutions?

The objective of our second research question was to examine if individual cognitive style could be used to predict the paradigm-relatedness of design outputs. Specifically, we hypothesized that lower KAI scores (more adaptive) would predict a greater likelihood for ideas being rated as paradigm-consistent ideas, as this bin would represent ideas that are more incremental [26, 27]. Conversely, we hypothesized that higher KAI scores (more innovative) would predict a greater likelihood for ideas to be rated as paradigm-breaking. As KAI scores are representative of cognitive style [31, 45], we would expect to see a similar relationship based on the idea ratings. Statistical assumptions were checked prior to the analysis.

A multinomial logistic regression was performed to model the relationship between an individual's KAI score and the classification of their ideas by the MTurk workers into the three paradigms (paradigm-consistent, paradigm-challenging, paradigm-breaking). Addition of the predictor (KAI score) to the model that contained only the intercept significantly improved the fit between model and data, χ^2 (2, N = 12,455) = 17.454, Nagelkerke R^2 = .002, p < .001. For this analysis, the reference group was paradigm-consistent. Accordingly, each predictor had two parameters, one for predicting membership in the paradigm-challenging group rather than the paradigm-consistent group, and one for predicting membership in the paradigm-breaking group. The parameter estimates are shown in Table 2. The predictors only had significant parameters for comparing the paradigm-consistent and paradigm-breaking groups; for each unit of increase in KAI score, the odds of being in the paradigm-breaking group compared to the paradigm-consistent group increased by 1.007.

Ryan-Einot-Gabriel-Welsch tests were used to make univariate pairwise comparisons between groups for each predictor that had a significant unique effect in the logistic regression. As shown in Table 2, the paradigm-breaking group can be characterized as being significantly higher in KAI scores (93.32) than the other two groups.

Table 2. Predictors' Unique Contributions in the Multinomial Logistic Regression for Individual KAI. Bold indicates significance at p < 0.05.

Predictor	Consistent Vs.	β	Exp(β)	*p*
KAI	Challenging	0.002	1.002	0.122
	Breaking	.007	1.007	**< 0.001**

While statistically significant, our findings do not necessarily support our hypothesis that individuals with higher KAI scores were more likely to produce ideas that were rated as paradigm-breaking. Specifically, the coefficient (.007) and effect size ($R^2 = .002$) are too small to suggest that there is an impact of individuals' KAI scores on the likelihood for ideas being categorized as paradigm-breaking for higher KAI. Similarly, the same can be said for suggesting that lower KAI scores were more likely to predict a likelihood for ideas being rated as paradigm-consistent. While our intentions were to find a relationship for using cognitive style as defined by KAI score [14, 15] to predict the likelihood of paradigm shifts occurring within ideation sessions for a given design problem, the small effect size conveys that without falling victim to the large effect size fallacy [90], our analysis oversimplifies any other mechanics occurring within the data. Particularly, the self-consistency in the MTurk ratings indicates that we should take a more nuanced approach to the data collected. Additionally, factors such as coping behavior can impact how individuals perform with respect to their cognitive style, depending on the task [15]. This supports further analysis to understand the data.

RQ3: Can the elevation or diversity of team cognitive style be used to predict the paradigm-relatedness of a design team's solutions?

The objective of our second research question was to examine whether KAI scores could be used to predict how ideas would be binned by paradigm-relatedness at the team level. Specifically, we hypothesized that higher team cognitive style *elevation* (average) [35] would predict greater likelihood for paradigm-breaking ideas. Conversely, we hypothesized that lower cognitive style elevation would lead to a greater likelihood of an idea being rated as paradigm-consistent, following the same logic stated before. Next, we hypothesized that team cognitive style *diversity* (standard deviation) [35] would not predict paradigm-relatedness, because the ideas would be more likely to be spread evenly across the paradigm-relatedness bins. In other words, teams with higher diversity would be more likely to produce ideas that contain both incremental and radical elements [15, 39]. Finally, we hypothesized that heterogeneity in a team would also predict a greater likelihood for an even spread of ideas across the bins, whereas homogeneous teams would not, based on prior literature about cognitive gaps [15]. Prior to the analysis, statistical assumptions were checked.

A multinomial logistic regression was performed to model the relationship between the team's KAI elevation (mean), KAI diversity (standard deviation), KAI homogeneity/heterogeneity, and the classification of their ideas into the three paradigms (paradigm-consistent, paradigm-challenging, paradigm-breaking). Addition of the predictor (KAI diversity, elevation, and KAI homogeneity/heterogeneity) to the model that contained only the intercept significantly improved the fit between model and data, χ^2 (4, N = 11,435) = 23.955, Nagelkerke $R^2 = .002$, $p < .001$. Significant unique contributions were made by the diversity of KAI scores (χ^2 (2) = 14.790, $p < .001$) and the KAI homogeneity/heterogeneity of the team (χ^2 (2) = 6.444, $p < .040$).

Table 3. Predictors' Unique Contributions in the Multinomial Logistic Regression for Team Measuring KAI. Bold indicates significance at p < 0.05.

Predictor	Consistent Vs.	β	Exp(β)	*p*
KAI Elevation (mean)	Challenging	0.005	1.005	0.031
	Breaking	0.003	1.003	0.216
KAI Diversity (SD)	**Challenging**	**0.015**	**1.015**	**< 0.001**
	Breaking	0.002	1.002	0.532
KAI homogeneity	Challenging	0.104	1.109	0.076
	Breaking	**0.162**	**0.176**	**0.019**
KAI heterogeneity (1)	Challenging	-	-	-
	Breaking	-	-	-

Similar to RQ2, the reference group was paradigm-consistent in this analysis. Accordingly, each predictor had two parameters, one for predicting membership in the paradigm-challenging group rather than the paradigm-consistent group, and one for predicting membership in the paradigm-breaking group. The parameter estimates are shown in Table 3. One of the predictors had a significant parameter for comparing the paradigm-consistent and paradigm challenging groups – the diversity of the team's KAI scores. For each unit of increase in the standard deviation of the team, the odds of being in the paradigm-challenging group compared to the paradigm-consistent group increased by 1.002. When teams were homogenous, the odds of being in the paradigm-breaking group compared to the paradigm consistent group increased by 1.176.

Ryan-Einot-Gabriel-Welsch tests were used to make univariate pairwise comparisons between groups for each predictor that had a significant unique effect in the logistic

Table 4: A Posteriori Pairwise Comparisons Between Group Means. Within each column, means sharing a superscript are not significantly different from each other.

Paradigm	Team KAI diversity	
Consistent	9.31^a	
Challenging		9.85^b
Breaking	9.40^a	

regression. As shown in Table 4, the paradigm-challenging group can be characterized as being significantly higher in the team's KAI elevation (9.85) than the other two groups.

These results refute our hypothesis such that team KAI elevation was not found to predict the likelihood for a team's ideas being categorized as any one of the three paradigm-relatedness types. Additionally, similar to RQ2, while statistically significant, our findings do not necessarily support our hypothesis that there is a greater likelihood for a team's ideas to be placed in the paradigm-challenging group when team KAI diversity increases. Specifically, the coefficient (.015) and effect size ($R^2 = .002$) are too small to suggest that there is practical significance. While prior literature shows that as a team's cognitive gap increases, as defined by diversity (standard deviation) [22], that team tends to develop ideas that span the solution space [15, 39], the lack of practical significance signals that these results are inconclusive. Typically, paradigm-challenging ideas can be seen as ideas that incorporate both incremental and radical elements [27], thus we would expect that teams with greater team KAI diversity are more likely to include both paradigm-consistent and paradigm-breaking elements in their ideas as a way of exploring the solution space. However, the small effect size conveys that our analysis oversimplifies any other mechanics occurring within the data, indicating a need for taking a more nuanced approach to analyzing data from a team perspective.

DISCUSSION

The main objective of this paper was to explore the role of cognitive style as defined by KAI on the paradigm-relatedness of design outputs from concept generation at the individual and team levels. In addition, we sought to identify the utility of crowdsourcing for measuring the paradigm-relatedness of design solutions. The main findings of this study were as follows:

- MTurk workers were able to provide consistent ratings of the paradigm-relatedness of design solutions
- Increases in individual KAI scores were related to an increased likelihood of MTurk workers rating the ideas as *paradigm-challenging*; consisting of ideas that contain incremental and radical elements [27], but results are limited due to the low pseudo R^2.
- Increases in team KAI diversity were related to an increased likelihood of MTurk workers rating the ideas as *paradigm-challenging*, but low pseudo R^2 limits our interpretation here as well.
- If a team had homogeneous KAI scores (i.e., within 10 points of each other), they had an increased likelihood of MTurk workers rating the ideas as *paradigm-breaking,* which includes ideas that shift the focus of the problem to a larger problem while violating some or all relevant constraints. Similarly, issues with low pseudo R^2 are present as well.

It is important to discuss more about these results before we draw significant conclusions. Specifically, while these results were statistically significant, the pseudo R^2 values highlight an important question: what can we meaningfully conclude from the results? To answer this question, we must discuss factors that contribute to lower pseudo R^2 values.

The Consistency of the Crowd

One potential cause of lower levels of pseudo R^2 may be caused by the raters used in the design rating process. Particularly, while crowdsourcing comes with many benefits, certain aspects, such as lack of worker experience [84] and difficulty in understanding how to rate certain criteria [46] can harm the reliability of crowdsourced data. Other factors such expertise or motivation may play a role in the data's reliability. Thus, it is important to ensure raters are being consistent with themselves [49]. While we attempted to control for this in our study by adding an additional question that was a replicate of one of the ideas in the set, we found that removing the 188 inconsistent raters did not change the conclusion of our interrater reliability. In other words, the inter-rater agreement (ICC(2)) only changed from 0.699 to 0.702 when removing these raters indicating that their lack of self-consistency was not negatively impacting the overall reliability of the crowd.

These results indicate that crowdsourcing can be used to consistently rating the paradigm-relatedness of ideas. However, we do not know if these ratings are *accurate*. As such future work is needed to explore the accuracy of the ratings by comparing them to expert reviewers. Future work is needed to explore further impact of ICC(2) on individual rater performance. For example, applying more exclusion criteria for detecting inconsistent responses could be applied to "clean" the data further, as hinted at in [49]. Additionally, some other techniques could be implemented, such as gathering ratings from more raters. Success was seen in studies that employed more workers to evaluate ideas [46, 48], leading us to believe that hiring more workers would be necessary for more robust results.

Individuals and Coping Behavior

Another potential cause of the pseudo R^2 may be because there is something else going on between the KAI scores and paradigm relatedness. Specifically, coping behavior can be used by individuals to overcome issues presented by cognitive gap by making an individual modify their behavior to solve problems in a way that is not consistent with their preferred style [15]. For example, although higher KAI scores (innovative trend) predicted more paradigm-breaking solutions, whereas lower KAI scores (adaptive trend) predicted more paradigm-consistent solutions, the parallel effects were not found at the team level. Team KAI elevation may not have been significantly associated with paradigm-consistent or paradigm-breaking design solutions because additive aggregation models assume that all team members' scores should be equally represented (e.g., [38]).

Yet, the reality is that some members may exert a disproportion influence on team creativity. Specifically, some individuals may have a greater impact on team outcomes through dominating the conversation, persuasively advocating for their idea, or demonstrating charismatic leadership (e.g., [67]). In contrast, some individuals may have a lower impact on team outcomes through failing to speak up, quickly acquiescing to others' ideas, or conforming to the majority decision even if they hold an alternative opinion [67]. Therefore, one high KAI scorer in the team may have persuaded three low KAI scorers to produce more paradigm-breaking solutions. Additionally, more investigation needs to be conducted to understand the skew of cognitive style within the homogeneous teams, as we would expect more innovative styles, when grouped together, to product more radical ideas as a team [15]. Therefore, individuals that are more innovative may be able to encourage each other to perform within their preferred cognitive style. The effects of how coping behavior leads individuals to change their behavior depending on the context have been examined in prior literature [15, 18, 39]. Thus, further investigation into the data is needed to understand the pseudo R^2.

LIMITATIONS, CONCLUSION, AND FUTURE WORK

While this study presents some interesting results to further broaden our view of how cognitive style as defined by KAI scores plays a role in predicting the paradigm-relatedness of design solutions during concept generation, such results do not come without limitations. First, many factors can influence the quality and types of ideas an individual proposes during concept generation; these might include their understanding of the design problem or reluctance in sharing ideas that come to mind. Additionally, prior work on creativity shows that other individual qualities can promote or prevent creative outputs [53], in our case, *paradigm-relatedness*. Particularly, individuals with more tacit knowledge about the design challenge tend to produce complex ideas more quickly, which can contribute to a cognitive gap within the team due to differences in cognitive level [15]. Because other individual characteristics may influence the paradigm-relatedness of design outputs during concept generation, more investigation of other variables is needed to understand what impacts the paradigm-relatedness of design outputs.

While crowdsourcing has been reliable for rating ideas' novelty, crowdsourced workers could not rate feasibility or marketability accurately [46]. This implies that while rating some aspects of ideas can be obtained accurately through crowdsourcing, other aspects may be too difficult to obtain. First, depending on each worker's experience, it is likely that some workers give less accurate and consistent ratings [84]. This may have affected the workers in this study, as experience was not a requirement to complete the HIT in this study. Another issue that could impact ratings is the complexity of the design problem. While other studies presented idea rating tasks that asked workers to rate the creativity of common household items [46] or judge whether ideas were similar to one another [49], the problem statements in this study more complicated. Although crowdsourced workers represent the U.S. population relatively well [81], this does not mean that every person is going to understand how to judge ideas that fit a certain niche [83]. Our results are further complicated by spreading the workload among the workers, where each worker rated one team only. However, this is necessary, as it would take too much time for each worker to rate all 1,395 ideas. Future work will include further cleaning of the MTurk data.

Although the current study sheds some light on how crowdsourcing impacts the accuracy of rating the paradigm-relatedness of design solutions from concept generation and how these ratings relate to KAI scores, further investigation must be done to understand the pseudo R^2 values. Specifically, are there other factors beyond cognitive style that impact the paradigm-relatedness of ideas, especially at the team level? While prior studies show that coping behavior can impact the overall behavior of individuals within a team or organization [64, 66], this study does not investigate those factors. Additionally, the structure of teams based on KAI was also limited by the students within each class, where we tried to construct homogeneous teams with relatively low or high KAI, but this could not be guaranteed. This also impacted how the 3- and 4-person teams were constructed, as some participants had unreliable KAI results or did not consent to the study, therefore 3-person teams were unavoidable at some points.

While our results provide a first look at applying paradigm-relatedness as a rating technique, specifically in crowdsourcing, these results are not exhaustive. Although MTurk workers' ratings were consistent with each other (precision), we currently do not have baseline for comparing whether ratings were in line with what should be expected as paradigm-consistent, paradigm-challenging, and paradigm-breaking. Future work will include investigating how in-person rating, using either expert raters or trained quasi-experts, compares to MTurk ratings for paradigm-relatedness. Additionally, as coping behavior can impact whether individuals work in their preferred cognitive style in a team [15, 18], future studies should look at how coping behavior can impact the outcome.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1825830. Special thanks are given to Abby O'Connell and Ava Drum for setting up Qualtrics surveys for Amazon Mechanical Turk, Katie Heininger for collecting data for Summer 2018, and Randall Doles for assisting with the MATLAB codes used to quantify the raw data. We also would like to thank our participants for their help in this project.

REFERENCES

[1] Kuhn, T. S., 1962, The structure of scientific revolutions, Chicago, University of Chicago Press.
[2] Merriam-Webster, "Paradigm shift," Merriam-Webster.com dictionary.
[3] Norman, D. A., and Verganti, R., 2014, "Incremental and radical innovation: Design research vs. technology and meaning change," Design Issues, 30(1), pp. 78-96.
[4] Sternberg, R. J., 1999, "A Propulsion Model of Types of Creative Contributions," Review of General Psychology, 3(2), pp. 83-100.
[5] Sternberg, R. J., 2005, "Creativity or creativities?," International Journal of Human-Computer Studies, 63(4), pp. 370-382.
[6] Benner, M. J., and Tushman, M. L., 2003, "Exploitation, Exploration, and Process Management: The Productivity Dilemma Revisited," Academy of Management Review, 28(2), pp. 238-256.
[7] Crilly, N., 2010, "The structure of design revolutions: Kuhnian paradigm shifts in creative problem solving," Design issues, 26(1), pp. 54-66.
[8] Farrell, R., and Hooker, C., 2013, "Design, science and wicked problems," Design Studies, 34(6), pp. 681-705.
[9] Rittel, H. W. J., and Webber, M. M., 1973, "Dilemmas in a general theory of planning," Policy Sciences, 4(2), pp. 155-169.
[10] Brush, S. G., 2000, "Thomas Kuhn as a Historian of Science," Science & Education, 9(1), pp. 39-58.
[11] Bird, A., 2014, Thomas Kuhn, Routledge.
[12] Hughes, T. P., 1990, American genesis: a history of the American genius for invention, Penguin Books, New York, N.Y.
London.
[13] Carmine, G., 2011, Innovation Secrets of Steve Jobs: Insanely Different Principles for Breakthrough Success, McGraw-Hill Education, New York.
[14] Kirton, M., 1976, "Adaptors and innovators: A description and measure," Journal of Applied Psychology, 61(5), pp. 622-629.
[15] Kirton, M. J., 2011, Adaption-Innovation in the Context of Diversity and Change, Routledge, London, UK.
[16] Kirton, M. J., 1984, "Adaptors and innovators—Why new initiatives get blocked," Long Range Planning, 17(2), pp. 137-143.
[17] Jablokow, K. W., 2008, "Developing problem solving leadership: A cognitive approach," International Journal of Engineering Education, 24(5), pp. 936-954.
[18] Jablokow, K. W., Teerlink, W., Yilmaz, S., Daly, S. R., and Silk, E. M., 2015, "The Impact of Teaming and Cognitive Style on Student Perceptions of Design Ideation Outcomes," ASEE Conferences, Seattle, Washington.
[19] Jablokow, K., Teerlink, W., Yilmaz, S., Daly, S., Silk, E., and Wehr, C., "Ideation Variety in Mechanical Design: Examining the Effects of Cognitive Style and Design Heuristics," Proc. ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference.
[20] Menold, J., Jablokow, K. W., Kisenwether, E. C., and Zappe, S. E., "Exploring the Impact of Cognitive Preferences on student receptivity to design thinking," Proc. ASEE Annual Conference and Exposition.
[21] Jablokow, K. W., Sonalkar, N., Edelman, J., Mabogunje, A., and Leifer, L., 2019, "Investigating the Influence of Designers' Cognitive Characteristics and Interaction Behaviors in Design Concept Generation," Journal of Mechanical Design, 141(9).
[22] Menold, J., and Jablokow, K., 2019, "Exploring the effects of cognitive style diversity and self-efficacy beliefs on final design attributes in student design teams," Design Studies, 60, pp. 71-102.
[23] Lapp, S., Jablokow, K., and McComb, C., 2019, "KABOOM: an agent-based model for simulating cognitive style in team problem solving," Design Science, 5, p. e13.
[24] Lapp, S., Jablokow, K., and McComb, C., "Collaborating With Style: Using an Agent-Based Model to Simulate Cognitive Style Diversity in Problem Solving Teams," Proc. ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering ConferenceV007T06A029.
[25] Heininger, K., Chen, H.-E., Jablokow, K., and Miller, S. R., "How Engineering Design Students' Creative Preferences and Cognitive Styles Impact Their Concept Generation and Screening," Proc. ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference.
[26] Silk, E. M., Daly, S. R., Jablokow, K. W., Yilmaz, S., Rechkemmer, A., and Wenger, J. M., "Using paradigm-relatedness to measure design ideation shifts," Proc. American Society for Engineering Education (ASEE) Annual Conference.
[27] Silk, E. M., Daly, S. R., Jablokow, K. W., and McKilligan, S., 2019, "Incremental to radical ideas: paradigm-relatedness metrics for investigating ideation creativity and diversity," International Journal of Design Creativity and Innovation, 7(1-2), pp. 30-49.
[28] Nagasundaram, M., and Bostrom, R. P., 1994, "The Structuring of Creative Processes Using GSS: A Framework for Research," Journal of Management Information Systems, 11(3), pp. 87-114.
[29] Dean, D. L., Hender, J., Rodgers, T., and Santanen, E., 2006, "Identifying good ideas: constructs and scales for idea evaluation," Journal of Association for Information Systems, 7(10), pp. 646-699.
[30] Paulus, P. B., and Brown, V. R., 2007, "Toward More Creative and Innovative Group Idea Generation: A Cognitive-Social-Motivational Perspective of Brainstorming," Social and Personality Psychology Compass, 1(1), pp. 248-265.
[31] Nijstad, B. A., and De Dreu, C. K. W., 2002, "Creativity and Group Innovation," Applied Psychology, 51(3), pp. 400-406.

[32] Fairchild, J., and Hunter, S. T., 2014, "“We've Got Creative Differences”: The Effects of Task Conflict and Participative Safety on Team Creative Performance," The Journal of Creative Behavior, 48(1), pp. 64-87.
[33] Paulus, P. B., and Kenworthy, J. B., 2018, "Overview of team creativity and innovation," Team creativity and innovation., Oxford University Press, New York, NY, US, pp. 11-38.
[34] Boon, A., and Dochy, F., 2016, "“The worst enemy to creativity is team-doubt”. The power of team creative efficacy to foster team processes of learning and creativity, and team effectiveness," Journal of Adult Learning, Knowledge and Innovation JALKI, 1(1), p. 1.
[35] Neuman, G. A., Wagner, S. H., and Christiansen, N. D., 1999, "The Relationship between Work-Team Personality Composition and the Job Performance of Teams," Group & Organization Management, 24(1), pp. 28-45.
[36] Mathieu, J., Maynard, M. T., Rapp, T., and Gilson, L., 2008, "Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future," Journal of management, 34(3), pp. 410-476.
[37] Mathieu, J. E., Gallagher, P. T., Domingo, M. A., and Klock, E. A., 2019, "Embracing Complexity: Reviewing the Past Decade of Team Effectiveness Research," Annual Review of Organizational Psychology and Organizational Behavior, 6(1), pp. 17-46.
[38] Van der Vegt, G. S., de Jong, S. B., Bunderson, J. S., and Molleman, E., 2010, "Power Asymmetry and Learning in Teams: The Moderating Role of Performance Feedback," Organization Science, 21(2), pp. 347-361.
[39] Jablokow, K. W., and Booth, D. E., 2006, "The impact and management of cognitive gap in high performance product development organizations," Journal of Engineering and Technology Management, 23(4), pp. 313-336.
[40] Gosnell, C. A., and Miller, S. R., "A Novel Method for Providing Global Assessments of Design Concepts Using Single-Word Adjectives and Semantic Similarity," Proc. ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering ConferenceV007T07A044.
[41] Miller, S. R., Hunter, S. T., Starkey, E., Ramachandran, S. K., Ahmed, F., and Fuge, M., 2020, "How should we measure creativity in engineering design? A Comparison of social science and engineering approaches (DETC2020-22446)," Journal of Mechanical Design, pp. 1-16.
[42] Aitamurto, T., Leiponen, A., and Tee, R., 2011, "The Promise of Idea Crowdsourcing - Benefits, Contexts, Limitations."
[43] Aguinis, H., Villamor, I., and Ramani, R. S., 2021, "MTurk Research: Review and Recommendations," Journal of Management, 47(4), pp. 823-837.
[44] Cheung, J. H., Burns, D. K., Sinclair, R. R., and Sliter, M., 2017, "Amazon Mechanical Turk in Organizational Psychology: An Evaluation and Practical Recommendations," Journal of Business and Psychology, 32(4), pp. 347-361.
[45] Burnap, A., Ren, Y., Gerth, R., Papazoglou, G., Gonzalez, R., and Papalambros, P. Y., 2015, "When Crowdsourcing Fails: A Study of Expertise on Crowdsourced Design Evaluation," Journal of Mechanical Design, 137(3).
[46] Kudrowitz, B. M., and Wallace, D., 2013, "Assessing the quality of ideas from prolific, early-stage product ideation," Journal of Engineering Design, 24(2), pp. 120-139.
[47] Grace, K., Maher, M., Fisher, D., and Brady, K., 2014, "A data-intensive approach to predicting creative designs based on novelty, value and surprise," International Journal of Design, Creativity, and Innovation.
[48] Lykourentzou, I., Ahmed, F., Papastathis, C., Sadien, I., and Papangelis, K., 2018, "When Crowds Give You Lemons: Filtering Innovative Ideas using a Diverse-Bag-of-Lemons Strategy," Proc. ACM Hum.-Comput. Interact., 2(CSCW), p. Article 115.
[49] Ahmed, F., Ramachandran, S. K., Fuge, M., Hunter, S., and Miller, S., 2018, "Interpreting Idea Maps: Pairwise Comparisons Reveal What Makes Ideas Novel," Journal of Mechanical Design, 141(2).
[50] Pugh, S., 1991, Total design: integrated methods for successful product engineering, Addison-Wesley.
[51] Plan, E. T., and Khandani, S., 2005, "Engineering design process."
[52] Dym, C. L., and Little, P., 2014, Engineering design: A project-based introduction, John Wiley and sons.
[53] Amabile, T. M., 1988, "A model of creativity and innovation in organizations," Research in organizational behavior, 10(1), pp. 123-167.
[54] Rietzschel, E. F., Nijstad, B. A., and Stroebe, W., 2006, "Productivity is not enough: A comparison of interactive and nominal brainstorming groups on idea generation and selection," Journal of Experimental Social Psychology, 42(2), pp. 244-251.
[55] Nickerson, R. S., 1999, "Enhancing Creativity," Handbook of Creativity, R. J. Sternberg, ed., Cambridge University Press, Cambridge, pp. 392-430.
[56] Rietzschel, E. F., Nijstad, B. A., and Stroebe, W., 2007, "Relative accessibility of domain knowledge and creativity: The effects of knowledge activation on the quantity and originality of generated ideas," Journal of Experimental Social Psychology, 43(6), pp. 933-946.
[57] Garfield, M., Satzinger, J., Taylor, N., and Dennis, A., "The Creative Road: The Impact of the Person, Process and Feedback on Idea Generation," Proc. AMCIS 1997.
[58] Garfield, M. J., Taylor, N. J., Dennis, A. R., and Satzinger, J. W., 2001, "Research Report: Modifying Paradigms—Individual Differences, Creativity Techniques, and Exposure to Ideas in Group Idea Generation," Information Systems Research, 12(3), pp. 322-333.
[59] Vercellone-Smith, P., Jablokow, K., and Friedel, C., 2012, "Characterizing communication networks in a web-based classroom: Cognitive styles and linguistic behavior of self-organizing groups in online discussions," Computers & Education, 59(2), pp. 222-235.

[60] Puccio, G. J., Treffinger, D. J., and Talbot, R. J., 1995, "Exploratory Examination of Relationships Between Creativity Styles and Creative Products," Creativity Research Journal, 8(2), pp. 157-172.
[61] Jablokow, K. W., and Kirton, M. J., 2009, Problem solving, creativity, and the level-style distinction, Springer, New York.
[62] Samuel, P., and Jablokow, K. W., "Toward an adaption-innovation strategy for engineering design," Proc. International Conference on Engineering Design (ICED11).
[63] Kirton, M. J., and McCarthy, R. M., 1985, "Personal and Group Estimates of the Kirton Inventory Scores," Psychological Reports, 57(3_suppl), pp. 1067-1070.
[64] Clapp, R. G., and De Ciantis, S. M., 1989, "Adaptors and Innovators in Large Organizations: Does Cognitive Style Characterize Actual Behavior of Employees at Work? An Exploratory Study," Psychological Reports, 65(2), pp. 503-513.
[65] Buffinton, K. W., Jablokow, K. W., and Martin, K. A., 2002, "Project Team Dynamics and Cognitive Style," Engineering Management Journal, 14(3), pp. 25-33.
[66] McCarthy, R., 1993, "The relationship of individual characteristics of women managers to the pressures experienced at work and choice of coping strategy," University of Hertfordshire.
[67] Mathieu, J. E., Tannenbaum, S. I., Donsbach, J. S., and Alliger, G. M., 2014, "A Review and Integration of Team Composition Models:Moving Toward a Dynamic and Temporal Framework," Journal of Management, 40(1), pp. 130-160.
[68] Amabile, T. M., 1982, "Social psychology of creativity: A consensual assessment technique," Journal of Personality and Social Psychology, 43(5), pp. 997-1013.
[69] Amabile, T. M., 1983, "The social psychology of creativity: A componential conceptualization," Journal of Personality and Social Psychology, 45(2), pp. 357-376.
[70] Amabile, T. M., 1996, Creativity in Context: Update to the Social Psychology of Creativity, Westview Press, Boulder, CO.
[71] Cropley, D. H., Kaufman, J. C., and Cropley, A. J., 2011, "Measuring creativity for innovation management," Journal of Technology Management & Innovation, 6(3), pp. 13-30.
[72] Horn, D., and Salvendy, G., 2009, "Measuring consumer perception of product creativity: Impact on satisfaction and purchasability," Human Factors and Ergonomics in Manufacturing & Service Industries, 19(3), pp. 223-240.
[73] Amabile, T. M., 1983, "Brilliant but cruel: Perceptions of negative evaluators," Journal of Experimental Social Psychology, 19(2), pp. 146-156.
[74] Kaufman, J. C., Baer, J., Cropley, D. H., Reiter-Palmon, R., and Sinnett, S., 2013, "Furious activity vs. understanding: How much expertise is needed to evaluate creative work?," Psychology of Aesthetics, Creativity, and the Arts, 7(4), pp. 332-340.
[75] Puccio, G. J., and Chimento, M. D., 2001, "Implicit Theories of Creativity: Laypersons' Perceptions of the Creativity of Adaptors and Innovators," Perceptual and Motor Skills, 92(3), pp. 675-681.
[76] Talbot, R. J., 1997, "Taking Style On Board (or how to get used to the idea of creative adaptors and uncreative innovators)," Creativity and Innovation Management, 6(3), pp. 177-184.
[77] Grace, K., and Maher, M. L., "Surprise and reformulation as meta-cognitive processes in creative design," Proc. Proceedings of the Third Annual Conference on Advances in Cognitive Systems, Cognitive Systems Foundation, p. 8.
[78] Sarkar, P., and Chakrabarti, A., 2014, "Ideas generated in conceptual design and their effects on creativity," Research in Engineering Design, 25(3), pp. 185-201.
[79] Estellés-Arolas, E., and González-Ladrón-de-Guevara, F., 2012, "Towards an integrated crowdsourcing definition," Journal of Information Science, 38(2), pp. 189-200.
[80] Highhouse, S., and Zhang, D., 2015, "The New Fruit Fly for Applied Psychological Research," Industrial and Organizational Psychology, 8(2), pp. 179-183.
[81] Paolacci, G., Chandler, J., and Ipeirotis, P. G., 2010, "Running experiments on amazon mechanical turk," Judgment and Decision Making, 5(5), pp. 411-419.
[82] Chen, L., Xu, P., and Liu, D., 2015, "Experts versus the Crowd: A Comparison of Selection Mechanisms in Crowdsourcing Contests," SSRN Electronic Journal.
[83] O'Quin, K., and Besemer, S. P., 1999, "Creative products," Encyclopedia of creativity, M. A. Runco, and S. R. Pritzker, eds., Academic Press, San Diego, CA, pp. 267-278.
[84] Liu, S., Xia, F., Zhang, J., and Wang, L., 2016, "How crowdsourcing risks affect performance: an exploratory model," Management Decision, 54(9), pp. 2235-2255.
[85] LeBreton, J. M., and Senter, J. L., 2008, "Answers to 20 questions about interrater reliability and interrater agreement," Organizational research methods, 11(4), pp. 815-852.
[86] Cole, C., Marhefka, J., Jablokow, K., Mohammed, S., Ritter, S., and Miller, S., "How Engineering Design Students' Psychological Safety Impacts Team Concept Generation and Screening Practices," Proc. ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering ConferenceV008T08A026.
[87] Fisher, R. A., 1925, "Theory of Statistical Estimation," Mathematical Proceedings of the Cambridge Philosophical Society, 22(5), pp. 700-725.
[88] Lance, C. E., Butts, M. M., and Michels, L. C., 2006, "The sources of four commonly reported cutoff criteria: What did they really say?," Organizational research methods, 9(2), pp. 202-220.
[89] James, L. R., Demaree, R. G., and Wolf, G., 1984, "Estimating within-group interrater reliability with and without response bias," Journal of applied psychology, 69(1), p. 85.
[90] Lantz, B., 2013, "The large sample size fallacy," Scandinavian Journal of Caring Sciences, 27(2), pp. 487-492.

