# Out in the Field Versus Inside in the Lab: A Comparison of Design Professionals' Concept Screening Practices

**Xuan Zheng**
Mem. ASME
Department of Industrial and Manufacturing Engineering,
The Pennsylvania State University,
343 Leonhard Building,
University Park, PA 16802
e-mail: xxz159@psu.edu

**Scarlett R. Miller**[1]
Mem. ASME
School of Engineering Design,
Technology and Professional Programs,
The Pennsylvania State University,
213-P Hammond Building,
University Park, PA 16802-1401
e-mail: scarlettmiller@psu.edu

*Concept screening is one of the gatekeepers of innovation process and thus is considered a vital component of engineering design. Yet, we know very little about how decisions are made during concept screening or the factors that inform these decisions. This is due, in part, to the fact that most prior work on concept screening in engineering design has focused on student populations or on industry professionals in an experimental setting which is not indicative of the risks and consequences professionals face in their daily work—particularly when it comes to innovative design process. Thus, the current study was developed to identify how the environmental settings (i.e., experimental versus naturalistic) and the role of the professionals in the design process (i.e., idea generators versus executives) impacts the criteria used to screen design ideas. Two studies were conducted including a workshop study with 45 design professionals from two companies in an experimental setting and a participatory ethnographic study with seven design professionals from a small electromechanical company in a naturalistic setting. The results showed stark differences in the criteria used to screen ideas between naturalistic and experimental practices and between idea generators and company executives. In addition, the results showed differences in the factors considered during concept screening between naturalistic and experimental environments. These results are used to identify opportunities for tools and methods that encourage the consideration of creative ideas in the engineering design industry and encourage appropriate risk-taking in engineering design.* [DOI: 10.1115/1.4047904]

*Keywords: design theory, design evaluation, design theory and methodology, conceptual design, decision theory, design process*

## 1 Introduction

US companies spent nearly $310 billion on Research and Development (R&D) in 2018 alone, which accounts for nearly 37% of the R&D expenditures in the world [1]. Importantly, this US investment in R&D has continued to rise over time in an effort to boost US innovation capacities [2]; between the years of 2007 and 2016, US companies invested an additional 23% in R&D [3]. This is due, in part, to the fact that successful innovations often lead to significant payoffs [4] and market leadership [2]. For example, a longitudinal study conducted by McKinsey & Company with 300 publicly listed companies over a five-year period showed a clear link between good design performance, high revenues, and high total returns to shareholders (TRS) [5].

Because of the role of innovation (successful implementations of creative ideas) in economic success, companies have been focusing on creative idea generation by encouraging their employees to generate a large number of creative ideas and by incorporating design thinking into the workplace [6]. However, innovation is often said to be stunted not by the inability of companies to *come up with good ideas* but by the inability to *move these highly innovative concepts through the design process* [7]. This is mainly because imperfect human decision makers in the innovation process often block truly original concepts from moving forward [8–11]. This obstruction has been postulated to be attributed to the risk or uncertainty associated with creative ideas [12,13]. However, taking *sensible* risks is viewed as an imperative part of the creative process [14,15]; without taking these risks there is no potential for innovation.

While the importance of selecting creative ideas is an important facet of the design industry, there has been limited investigation of concept screening practices with design professionals. This has been attributed to the challenges of conducting research on professional designers in situ such as setting the boundaries between the researchers and the objects of investigation, and the difficulty of collecting and analyzing data [16]. For example, many companies are fearful of disclosing trade secrets (e.g., intellectual property) or additional expenses to the company in terms of the extra time needed for employees to partake in the research study.

Because of these challenges, researchers have relied largely on studies of professional designers in an experimental setting (see for example Refs. [10,11,17]) or utilized students to study concept screening behaviors [18–21]. Specifically, this previous research indicated that both internal factors (e.g., individual evaluation styles [22], cognitive biases [10,11,17]), and external factors (e.g., usability, looks, and feasibility) [17] could influence professional concept screening practices in the experimental setting. Although this prior work helps us begin to understand professional concept screening [10], the types of experimental challenges presented in these studies do not necessarily involve *real-life consequences or risks* that affect design decision-making or ensure the design challenge fits in the expertise of the participants [23]. In addition, recent research on engineering students has linked individual attributes, including the big-five personality traits, financial risk aversion and ambiguity aversion [18,20], and ownership bias [21,24] to engineering students' concept screening practices. However, we do not know if these same findings transfer to a professional setting or what differences exist. Moreover, professionals

[1]Corresponding author.

Contributed by the Design Theory and Methodology Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received January 10, 2020; final manuscript received July 7, 2020; published online October 12, 2020. Assoc. Editor: Daniel A. McAdams.

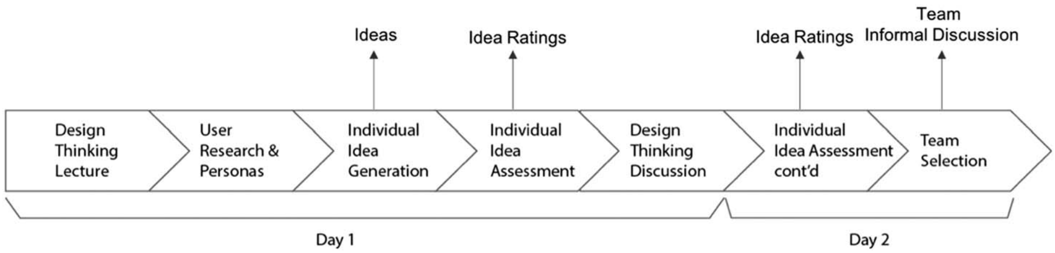



**Fig. 1 Timeline of study II and the corresponding data collected**

playing different roles in the design process often have different responsibilities [25,26], biases [9,27], and preferences for risk [28,29]. Nevertheless, limited research has explored the relationship between the role of the professionals in the design process and the discussion topics focused on when making screening decisions.

Based on previous work, the purpose of the current paper was to identify how environmental settings (i.e., experimental versus naturalistic) and the role of the professionals in the design process (i.e., idea generators versus executives) were related to the criteria used to screen design ideas. In addition, these criteria identified were also compared with those used by student teams. This was achieved through two studies, including an experimental workshop and a participatory ethnographic study that was conducted in a naturalistic setting. The results of this study can be used to identify opportunities for decision support tools that encourage appropriate risk-taking in the concept screening process.

## 2 Related Work

Before we can begin to discuss the similarities and differences in concept screening practices between professionals and students and between naturalistic and experimental settings, it is first important to understand the concept screening process and its role in design innovation. Concept screening often occurs at the beginning of the design process, when dozens of ideas are evaluated and the ideas with the most promise are moved forward in the design process [30]. This process forms the very first stage and gate in the Stage-Gate system developed by Cooper [26], see Fig. 1. While good decisions in this process help companies select the right projects to appropriately allocate their resources, expedite the speed of products to market, and maximize their return from R&D investments [26], poor decisions during this process result in great expense, including redesign costs, production postponement, and even product failure [31]. However, decisions made in the early phases of the design process are especially difficult to make due to the limited information about these early-stage ideas and their future potential for success [32].

One factor that greatly inhibits successful decision-making in the fuzzy front end of the design process, or the period between when an opportunity for a new product is first considered and when the product idea is judged ready to enter "formal" development,[2] is that there are no "ideal" concept screening practices. For example, a recent study that investigated the design process of eight companies from California supported this idea, indicating that there was no concept evaluation procedure found to choose the best concept [33]. This is not surprising since formal concept evaluation tools developed in academic research that provide frameworks or criteria to guide concept screening (see for example Refs. [34–37]) are usually not preferred in this fast-paced industry due to their time-consuming and not easy-to-use nature. Because of this, these formal concept evaluation tools are often criticized due to their disconnection with the requirements of the industry [38,39]. Instead, the concept evaluation and screening process in the engineering design industry often rely on informal methods, such as design review meetings [40], voting [41], and informal team discussions [39]. For example, Maurer and Widmann [33] found through interviews with and observations on professionals from eight companies in California that "creativity" and "gut feelings" were frequently mentioned features in "brainstorming" meetings. However, we know very little about how decisions are made in this process.

There has been some recent work that has begun to answer this question. For example, a recent study with engineering design students in an experimental design setting showed that students selected ideas primarily based on their technical feasibility, how they compared with other generated ideas, and how they compared with products already on the market [42]. In addition, research has also demonstrated that students often discarded novel ideas during concept screening for conventional alternatives, leading to a loss of creative capital [43]. Work with design professionals in an experimental setting has demonstrated that their decision-making may be influenced by factors such as individual evaluation styles [22], cognitive biases [10,11,17], and idea properties including usability, looks, and feasibility [17]. However, these experimental studies do not necessarily involve *real-life consequences or risks* that affect design decision-making. This is problematic because other factors such as social risks (e.g., the possibility of criticism, rejection, and failure [44]) and cognitive traps (e.g., halo effect [45] and social loafing [46,47]) are note captured in these environments.

While previous research has identified factors that are important to consider in the early conceptual design stages in order to cultivate successful products such as safety and durability [48–50], market and technological fit [51], aesthetics and ergonomics [52], reliability [53], compatibility [54], reusability [55], customer demand [56], and technical feasibility and cost [35,57], we do not know how these factors impact the screening process. Finally, since a company's sustainable growth is largely decided by its ability to innovate [58], creativity (or how feasible and original an idea is [22,59,60]) has been regarded a fundamental component in the new product development process [61].

In addition to comparing the criteria used to make screening decisions during informal concept screening in different settings (i.e., experimental versus naturalistic), it is also important to explore how the criteria used might differ by the role of the professionals in the design process. For example, members of R&D Teams often serve as idea generators and the first "gatekeepers" in the design process, as they must screen dozens of ideas generated through the process to a few key ideas to present to design executives [25]. In fact, the primary responsibilities of R&D units are undertaking research and developing new technological innovations for the market place [62], which require the employees' engagement of creative behaviors including both ideation and implementation [63]. However, the key decision makers need to make Go/No Go decisions based on the deliverables presented by the idea generators [26]. In addition, professionals could exhibit biases, such as ownership bias [10,11,17,27] and various levels of preference for risk [64–66] associated with their roles in the design process. For example, a previous study found that professionals showed biases toward their own ideas and exhibited a

[2]https://en.wikipedia.org/wiki/Front_end_innovation

preference toward highly feasible ideas, yet against highly original ideas when evaluating self-generated ideas in an experimental setting [27]. Similarly, managers who initiate a project are less likely to realize that it is failing and are more likely to continue funding it than managers who assume the leadership after the project has started [9]. What's more, a research study found that entrepreneurs who made the commitment to become business owners had a high degree of optimism on business success, which could be because they committed more hours on the company than employees [28].

## 3 Research Questions

The purpose of the current paper was to identify how environmental settings (i.e., experimental versus naturalistic) and the role of the professionals in the design process (i.e., idea generators versus executives) were related to the criteria used to screen design ideas. Specifically, our study was developed to answer the following questions:

*RQ1: What are the main discussion criteria used during professional concept screening? Does this differ between experimental and naturalistic settings?*

We hypothesized that user needs, technical feasibility, and cost would be the most frequently mentioned topics since these were typical design considerations adopted by companies that have been found in previous research [33,56,67]. In addition to the frequently mentioned topics identified in the experimental setting, we hypothesized that professionals would also consider factors related to technology, market, cost, and company resources in a naturalistic setting since these factors were found to be related to the Go/No Go decisions [51].

*RQ2: In a naturalistic setting, are idea generators and executives selecting ideas perceived to have high market potential?*

This research question was developed to understand if idea generators and executives made similar screening decisions on whether to further consider an idea or not regarding the market potential of the ideas. This question serves as a first step to understand the relationship between the role of the professionals and concept screening. We hypothesized that both idea generators and executives would be able to select ideas perceived with high future market potential in a naturalistic environment. This hypothesis was partially based on the previous research finding that professionals showed a preference for ideas perceived with high future value in a workshop setting [27].

*RQ3: In a naturalistic setting, what factors predict the screening decisions of the idea generators and the executives individually?*

This research question was developed to understand if idea generators and executives screened the ideas based on similar design criteria regarding the perceived feasibility, originality, overall creativity, and riskiness. We hypothesized that idea generators and executives would use different design criteria when making screening decisions due to their different responsibilities [25,26].

## 4 Methodology

In order to answer these research questions, two studies were conducted with design professionals. Specifically, study I was conducted in the form of design workshops in an experimental setting and study II was conducted in the form of a design workshop based in a naturalistic environment.

**4.1 Study I: An Experimental Design Challenge.** In study I, two workshops were conducted with design professionals from two companies. The remainder of Sec. 4.1 provides details of the methodology used in study I.

*4.1.1 Participants.* A total of 45 participants (eight females and 37 males) were recruited from two companies to participate in the workshops. Specifically, 15 professionals from a large (>30,000 employees) global automotive parts design and manufacturing company were recruited in the first workshop and 30 professionals from a small (<200 employees) hardware and software media support company were recruited to participate in the second workshop. These companies were selected because they have a mission to provide groundbreaking and innovative products within their respective industries. The average age of the participants was 38.53 years old (SD = 11.97 years) with an average of 11.05 years of design-related professional experience (SD = 9.74 years). The participants held a variety of educational degrees, including associates or high school ($N = 5$), bachelors ($N = 17$), masters ($N = 8$), and doctorate ($N = 5$). The participants' job titles included engineers, scientists, managers, and other positions, such as technical specialist and programmer.

*4.1.2 Procedure.* To recruit individuals for the study, an invitation email was sent to potential participants by a company representative. In the invitation email, individuals were encouraged (but not required) to participate in a company-wide design thinking workshop. Two workshops were then scheduled onsite at the two companies with individuals who were willing to participate based on their schedule, availability, and interest. Because the workshops were scheduled during work hours, no monetary compensation was provided. However, participation was completely voluntary; there was no consequence from the company for not attending the workshop, and participants were able to quit the workshop at any time. The procedure introduced next was used at both companies:

The workshop was held in a room provided by the company. At the beginning of the workshop, the demographic information of the participants was collected through a paper-based survey. Then, a 30 min lecture on design thinking and the importance of creative idea development was given to the participants. Next, the participants were grouped into three- or four-member teams, which were formed by the research team prior to the workshop. All managers were put on the same team in order to minimize the pressure that might exist if a manager was present in a team with subordinates, which could discourage other participants from expressing their genuine feelings during the activity. There were 12 teams studied throughout both workshops. Participants were provided with a design prompt, which was "to develop concepts for a new, innovative product that can froth milk in a short amount of time with minimal instruction." The full design prompt can be viewed in Appendix. This task was selected through a conversation with a representative from each company, where a variety of design prompts used within the literature were discussed. This task was selected with consultation with upper management at the organization in order to represent a task that the participants would be equally familiar with while also utilizing their expertise to develop solutions to the problem without violating company intellectual property. Once the design problem was understood, participants were given 20 min to generate ideas for this design problem individually by sketching ideas on index cards and providing a short-written description of each idea.

Following this brainstorming session, participants were asked to assess all ideas generated by their team without talking to their team members using an individual idea assessment sheet.[3] After the individual concept screening, the teams were instructed to categorize all the generated ideas into "Consider" and "Do not Consider" piles, rank the ideas, and conclude with a final idea through informal discussions. Specifically, the participants were asked to categorize each concept into "consider" if they felt the idea would most

[3] https://www.engr.psu.edu/britelab/resources/Concept%20Screening%20Sheets.pdf

likely satisfy the design goals or if they wanted to prototype and test these ideas immediately or "do not consider" if they felt the idea had little to no likelihood of satisfying the design goals or if they felt there was minimal value in the idea. The final idea could be one of the ideas, a combination of several ideas, or a completely new idea. This discussion was audio-recorded. Finally, the teams were asked to perform a 60 s "Elevator Pitch" to introduce and promote their final idea to the rest of the workshop participants. The workshop then concluded with a discussion on decision-making biases in the engineering design process.

*4.1.3 Qualitative Data Coding Procedure—Experimental Study.* During the concept screening practices, professional teams generated a total of 216 min of audio dialogue that was transcribed and coded by two independent coders. One of the raters was the first author and the other rater was an undergraduate research assistant. Then, the transcripts of the team dialogues were analyzed sentence-by-sentence following principles of deductive content analysis [68] in NVivo v.11 [69]. The deductive qualitative research approach was selected because prior research had identified discussion topics during student teams' informal concept screening discussion, therefore the coding schemes generated in previous research [42] were adopted in study I. Then, the two raters identified instances of discussions (defined as a block of dialogue between the team members on a particular topic). Both raters coded 20% of the team discussions (two teams) individually. The raters were allowed to add new topics to the coding scheme; however, no new topics were identified. Interrater reliability was composed for the two raters across all child nodes using Cohen's kappa and found to be 0.69. This is considered to be an acceptable level of agreement according to Landis and Koch [70]. Any disagreements were settled in a conference between the two raters. After that, the author coded the rest of the audio dialogue alone. The number of instances of discussion topics was computed. Similar categories were grouped together in order to reduce the number of categories and appropriately describe the types of topics professional teams focused on.

**4.2 Study II: A Naturalistic Design Problem.** In order to compare the discussion topics that professional teams focused on when working on an experimental design problem and a naturalistic design problem, a study was conducted with a four-person design team and three executives at a small electromechanical company with less than 50 employees. The remainder of Sec. 4.2 provides details of the procedure followed in study II.

*4.2.1 Participants.* A total of seven participants (one female and six male) participated in this study. Four of the participants were members of the design team (idea generators) working on redesigning existing products for a new target audience. The design team had been with the company around 2 years and had bachelor's degrees in Computer Science, Electrical Engineering, and Marketing. The three executives were the CEO, president and COO, and VP of operations at the company each with more than 25 years of experience in the industry. The executives were from an older generation than that of the target audience of the project.

*4.2.2 Procedure.* In order to recruit a design team for the study, an invitation email was sent to local electromechanical companies with less than 100 employees through email recruitment. While no compensation was provided to the company or the individual participants, the company and participants both benefitted from a design thinking workshop that was provided through the study free of cost. Once a company and the design team were identified, each participant individually signed the consent form. After that, the researchers investigated the conceptual design process of the design team working on redesigning existing products for a new target audience from a younger generation. The timeline of study II is shown in Fig. 1.

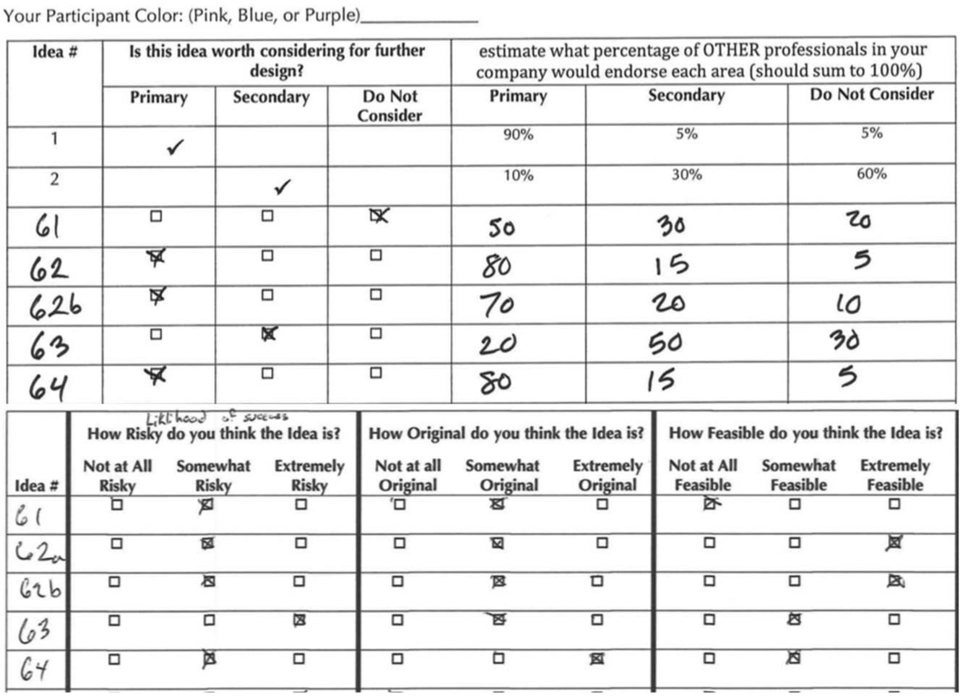
*Semi-structured interviews:* A semi-structured interview was completed individually with three of the design team members and a member of the executive team. The interview was conducted to understand the company culture and the background of the project, especially on whether the project was considered risky. Each interview was recorded and lasted about 30 min. The interview began by collecting basic demographic information about the participant and their role in the organization. Next, the participants were asked a series of questions related to the creative climate and risk-taking nature of the organization. In order to minimize response bias, questions were asked in a neutral form. Example questions include: "Do you feel your company values more creative or more traditional ideas? Why?", "How would you define risk-taking in the fuzzy front end of new product development?", and, "Within the NPD process at your current company, would you consider yourself a risk taker?" The interviews were transcribed and the responses to questions that were directly related to the design project were used to support the findings of the current study.

Next, a four-hour workshop was designed for the team in order to provide training on design thinking. Prior to the workshop, the design team members had conducted some research on their target customers and existing solutions on the market. The workshop then took place in two consecutive days in a conference room provided by the company.

*Design thinking workshop—day 1:* On the first day, a researcher gave a design thinking lecture to the design team. Then, the researchers and the design team spent half an hour together doing research online and discussing the current solutions being used in the target customer group. After that, the researchers instructed the design team to build personas to represent the characteristics and needs of their primary, secondary, and tertiary customers. Then, the participants were asked to sketch out their ideas on the idea generation sheets provided by the researchers following a few "How might we" questions in order to boost the idea generation process. An example of "How might we" questions was "how might we solve this problem if money and resources were not an issue". A total of 37 ideas were generated.

After idea generation, each participant was provided with a stack of ideas from one of their team members and then individually assessed the ideas using the categorization and evaluation sheets provided by the researchers, see Fig. 2. Specifically, participants were asked to categorize the ideas into three consideration categories: (1) *Primary*, if an idea has great future value and company funds should be expended to pursue the idea; (2) *Secondary*, if an idea may have some future value and should be funded only if there are extra resources available; (3) *Do Not Fund*, if an idea should not be financially supported. The participants were asked to pass their stacks of ideas until each member completed the categorization of all the ideas generated by the team. The completion of categorization marked the end of the first half of the workshop on the first day.

*Design thinking workshop—day 2:* On the second day, participants were guided to follow the same procedure with the categorization of ideas to individually evaluate the riskiness, originality, and feasibility of the ideas based on their perceptions using a three-point scale (not at all, somewhat, and extremely), see Fig. 2. Specifically, *riskiness* indicates the uncertainty surrounding whether or not positive outcomes will be realized if company resources are expended on developing the ideas; *originality* indicates how unique or surprising the idea is; *feasibility* indicates how logical, useful, valuable, or understandable the idea is. It needs to be stated that one of the team members did not participate in the second part of the workshop; therefore, there were three categorization and evaluations of the ideas collected from the design team. This sequence (first categorization and then evaluation) was designed to not prime the design professionals to think about the specific properties of the ideas when making categorizations. After that, the design team members were asked to discuss which ideas to recommend to management as a team. Specifically, the participants categorized the



Your Participant Color: (Pink, Blue, or Purple)____________

| Idea # | Is this idea worth considering for further design? | | | estimate what percentage of OTHER professionals in your company would endorse each area (should sum to 100%) | | |
|---|---|---|---|---|---|---|
| | Primary | Secondary | Do Not Consider | Primary | Secondary | Do Not Consider |
| 1 | ✓ | | | 90% | 5% | 5% |
| 2 | | ✓ | | 10% | 30% | 60% |
| 61 | ☐ | ☐ | ☒ | 50 | 30 | 20 |
| 62 | ☒ | ☐ | ☐ | 80 | 15 | 5 |
| 626 | ☒ | ☐ | ☐ | 70 | 20 | 10 |
| 63 | ☐ | ☒ | ☐ | 20 | 50 | 30 |
| 64 | ☒ | ☐ | ☐ | 80 | 15 | 5 |

| Idea # | How Risky (Likelihood of success) do you think the Idea is? | | | How Original do you think the Idea is? | | | How Feasible do you think the Idea is? | | |
|---|---|---|---|---|---|---|---|---|---|
| | Not at All Risky | Somewhat Risky | Extremely Risky | Not at all Original | Somewhat Original | Extremely Original | Not at All Feasible | Somewhat Feasible | Extremely Feasible |
| 61 | ☐ | ☒ | ☐ | ☐ | ☒ | ☐ | ☒ | ☐ | ☐ |
| 62a | ☐ | ☒ | ☐ | ☐ | ☒ | ☐ | ☐ | ☐ | ☒ |
| 62b | ☐ | ☒ | ☐ | ☐ | ☒ | ☐ | ☐ | ☐ | ☒ |
| 63 | ☐ | ☐ | ☒ | ☐ | ☒ | ☐ | ☐ | ☒ | ☐ |
| 64 | ☐ | ☒ | ☐ | ☐ | ☐ | ☒ | ☐ | ☒ | ☐ |

**Fig. 2 Categorization and idea evaluation sheets used by design team members in study II**

ideas into consider and do not consider and evaluated the ideas during the discussion, which was audio-recorded for further analysis.

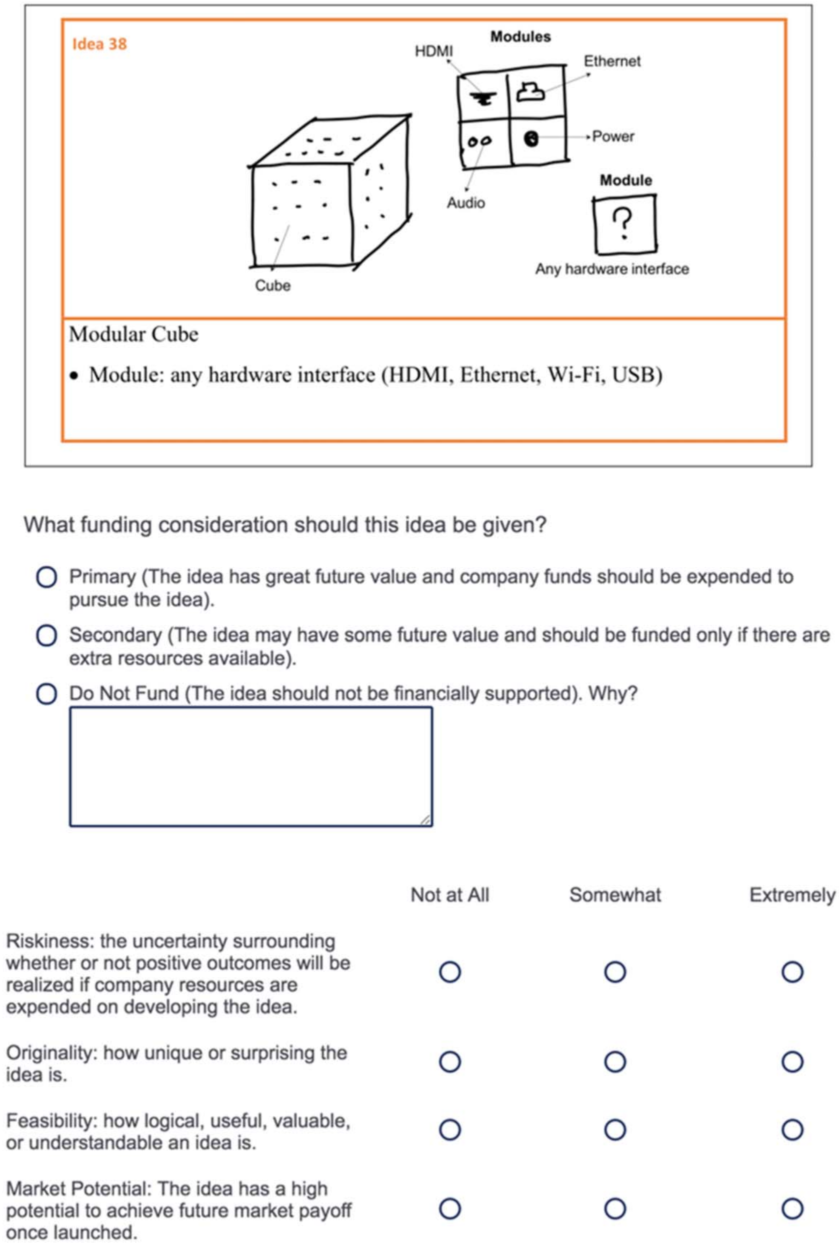
*Concept screening by executives:* After the workshop, the ideas were redrawn by a researcher (in order to anonymize the ideas, so that the executives could not identify the creator of the ideas based on the drawing skills or handwritings), and a Qualtrics survey was created for the executives to first categorize and then evaluate the ideas. In the survey, the ideas were presented in random order for both the categorization and the evaluation. If an idea was categorized as a Do Not Fund type, the executives were asked to provide comments on why an idea was not considered. In the evaluation process, in addition to riskiness, originality, and feasibility, the executives were also asked to estimate the *market potential* of the ideas, where a high market potential indicated a high possibility to achieve future market payoff once launched. See Fig. 3 for screenshots of the survey. The aggregated executives' ratings on the market potential of the ideas were used as a baseline measure of the value of the ideas. The market potential was included in the assessment in order to translate the value of the ideas into return on the investment, allowing executives to evaluate the ideas from a benefit point of view in addition to the costs. Only executives, but not the idea generators, were asked to evaluate the market potential. This was not assuming that the executives did a perfect job in evaluating the ideas. This was rather because, in many companies, the ideas selected by the executives are the ideas that actually will be considered further [10].

*4.2.3 Metrics—Naturalistic Study. Perceived properties:* Or the perceived riskiness, feasibility, originality, and market potential. The three-point Likert-scale ratings of the perceived properties were coded into numerical values. Specifically, 1 indicated that the idea was perceived as not at all risky/feasible/original/ with no market potential; 2 indicated that the idea was perceived as somewhat risky/feasible/original/with market potential; and 3 indicated that the idea was perceived as extremely risky/feasible/original/with market potential.

*Concept screening:* The three consideration categories were coded into numerical values as well. Specifically, 1, 2, and 3 indicated that the idea was categorized into the primary, secondary, and Do Not Fund consideration category, respectively.

*Aggregated market potential score*: Since mid-range levels of differentiation and integration within top management teams are associated with the improved performance [71], we did not expect a high level of agreement among the executives on the perceived future market potential of the ideas. Fleiss' kappa indicated that there was no significant agreement among the executives' ($N = 3$) evaluations on future market potential of the ideas, $\kappa = -0.114$, 95% CI [−0.260, 0.031], $p = 0.124$. Therefore, the executives' evaluations were aggregated to indicate the level of the market potential of an idea. Specifically, an aggregated market potential score rated by the executives was calculated for each idea as an indicator of the idea's possibility to achieve future market payoff. The coded values of different levels of perceived market potential by the executives were summed up to achieve the aggregated score for each idea. Six ideas had an aggregated score market potential of 5, 22 ideas had an aggregated market potential score of 6, and nine ideas had an aggregated market potential score of 7. The range of this aggregated market potential score was [2,8].

*4.2.4 Qualitative Data Coding Procedure—Naturalistic Study.* During concept screening, professional teams generated 90 min of audio dialogue that was transcribed and coded by two independent coders. One of the raters was the first author, who had conducted content analysis in previous research, and the other rater was an undergraduate research assistant, who completed training on how to conduct content analysis. The audio recording was first transcribed using Amazon Web Services (AWS) transcription service and later proofread by an undergraduate research assistant. The coding scheme developed by Toh and Miller [42] was found not suitable on the discussion data in study II since the discussion was developed around criteria in the naturalistic setting, while the discussion was developed around ideas in the experimental setting. Therefore, the transcripts of the team dialogue were analyzed sentence-by-sentence following principles of inductive content analysis [68] in NVivo v.12 [72]. Then, general themes regarding discussion topics were identified and similar categories were grouped together to appropriately describe the types of



Modular Cube

- Module: any hardware interface (HDMI, Ethernet, Wi-Fi, USB)

What funding consideration should this idea be given?

- Primary (The idea has great future value and company funds should be expended to pursue the idea).
- Secondary (The idea may have some future value and should be funded only if there are extra resources available).
- Do Not Fund (The idea should not be financially supported). Why?

| | Not at All | Somewhat | Extremely |
|---|---|---|---|
| Riskiness: the uncertainty surrounding whether or not positive outcomes will be realized if company resources are expended on developing the idea. | O | O | O |
| Originality: how unique or surprising the idea is. | O | O | O |
| Feasibility: how logical, useful, valuable, or understandable an idea is. | O | O | O |
| Market Potential: The idea has a high potential to achieve future market payoff once launched. | O | O | O |

**Fig. 3 Example questions in the qualtrics survey for executives in study II**

topics the professional teams discussed. The two raters coded the entire audio dialogue individually and achieved a moderate level of agreement with a Cohen's kappa of 0.57. Disagreements were settled in a conference between the two raters, then the word counts of discussion topics were computed.

## 5 Results

In order to answer the research questions, the data from the generated concepts and the coding of team discussions were analyzed. SPSS V. 25 [73] was used for the quantitative analysis with a significance level of 0.05, and NVivo 12 [72] was used for the qualitative analysis. The results of our analysis are presented in order of our research questions.

**5.1 What are the Main Discussion Topics Used During Professional Concept Screening? Does This Differ Between Experimental and Naturalistic Design Challenges?** The first research questions focused on identifying and comparing discussion topics used by teams during concept screening in an experimental and naturalistic setting. In order to answer this research question, the audio recordings from the informal team discussion ($N = 13$ teams) in study I and II were analyzed. Figure 4 provides an overview of the high-level discussion topics and relative frequency of these topics across these settings. Specifically, these results demonstrate that in an experimental setting, the teams focused on the *technical feasibility* (61%) of the idea, and comparing the ideas with *existing products* (18%) and *other generated ideas* (11%). However, teams in the naturalistic setting showed distinctive focuses on *user needs* (40%), followed by *technical feasibility* (22%), *marketing* (15%), and *company resources* (13%). The remainder of this section provides details on these discussion topics and discussed the similarities and differences in different conditions.

*5.1.1 Technical Feasibility.* *Technical feasibility*, which included discussions about "the ease of execution and effectiveness

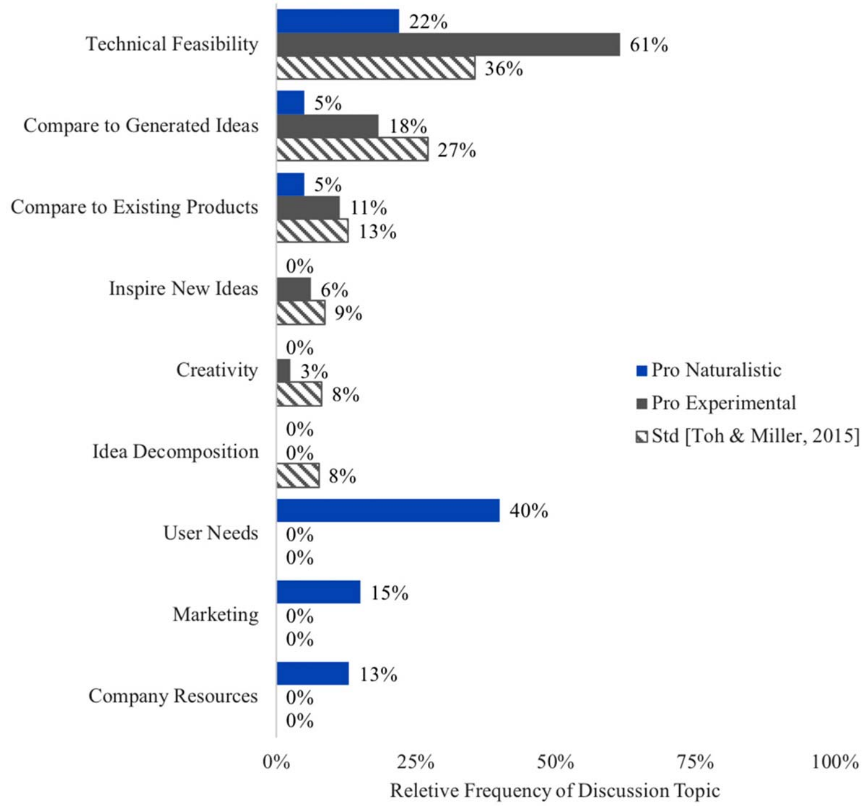



**Fig. 4 Relative frequency of high-level discussion topic**

of a concept in satisfying the design goal" [42] was an area of focus in both the naturalistic and experimental settings. However, the two settings emphasized different sub-topics of technical feasibility; see Fig. 5. For instance, in the experimental setting, the teams focused on the sub-topics of *mechanics* (50%), *ability to satisfy design goals* (39%), *pricing* (4%), and *maintenance* (3%). This indicated that in the experimental setting, the teams spent a large amount of their discussion on how the idea would be realized with specific mechanical structures or based on physics. For example, a participant in team T-1 explained the mechanics of an idea by introducing the mechanical parts, "So it could be connected to a steel shaft and as it shakes, the rubber gives it the ability to move." In addition, the teams also discussed the *ability to satisfy design goals,* including explicitly stated design goals in the description of the design problem (e.g., vigorous motion, heating element, efficiency, and easiness to use) and inexplicitly stated goals (e.g., safety). For example, a participant in team 4 recommended to reject ideas since they did not fulfill all the design requirements, "I will put my two here ("Do not Consider") because I did not consider the stirring element, just the heating element." In this experimental setting, the teams also briefly mentioned the *pricing* and *maintenance* of the future product.

However, in the naturalistic setting, the teams focused more on *pricing* (36%) and *manufacturing* factors (14%), in addition to the *ability to satisfy design goals* (52%). For instance, participant 4 in study II proposed a method to produce, or "*manufacture*" the product package while considering the *pricing*, "… for low quantities I mean within ten times of what we're doing right now, you could probably just do it with a [a lower cost system]." This finding suggested that in a naturalistic environment, teams shed more light on factors closely related to later design stages, such as the manufacturability regarding the production of the product and the market performance of the product.

*5.1.2 Comparing With Existing Products.* In addition to technical feasibility, *comparing ideas to existing products* was another discussion topic that was shared the experimental and naturalistic settings. In the experimental setting, the teams primarily compared generated ideas with existing products so as to *explain* the ideas (56%), to identify *similarities* (31%), or to use existing products as *proof of concept* (13%). For example, in study I, a participant in team T-1 explained an idea that used effervescent tablets to generate bubbles and heat by referring to "Coke and Mentos." In addition, a participant in Team V-2 proved the effectiveness of the heating effect of an idea by referring to her past experience of accidentally using warm milk to make butter in a butter churner. Though not heavily emphasized, the teams in the naturalistic setting compared new ideas with existing products for *competitor analysis* (30%) in addition to *explanation* (74%). For example, in study II, participant 2 provided an example from Apple to explain the unboxing experience, and participant 3 analyzed a potential competitor and identified an opportunity to develop a product dedicated to a specific market, "there are things out there that say, "works with [a website], but it's not dedicated for [the same design market]".

*5.1.3 Comparing With Other Generated Ideas.* In addition to comparing with existing products, the design teams also c*ompared new ideas with other generated ideas*. Specifically, the teams compared and contrasted ideas to identify *similarities* between generated ideas in the experimental setting. For example, a participant in team T-1 identified a similarity between two ideas and said, "So these two are kind of similar—injecting air through the aerator." In the naturalistic setting, the teams only occasionally compared generated ideas to identify the similarities among the ideas and used that information to group the ideas. This could be due to the fact that the team in the naturalistic environment followed a different discussion pattern than teams in the experimental environment. Instead of evaluating and making decisions on the ideas one by one in the experimental setting, the teams developed their discussions based on the high-level criteria in the naturalistic setting.

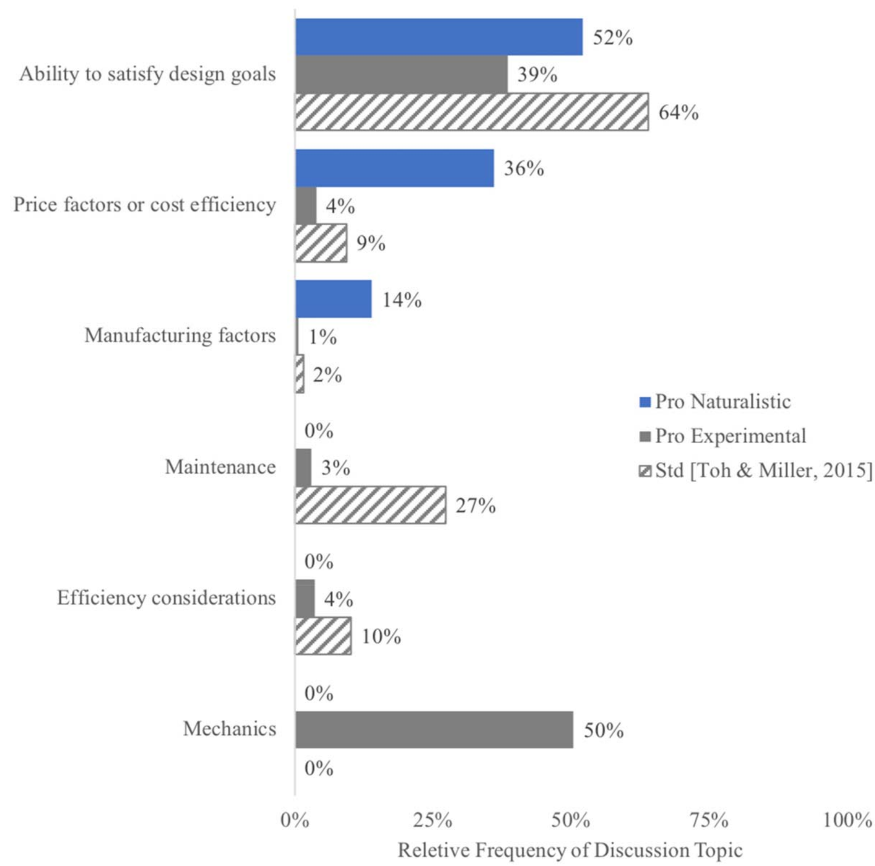



**Fig. 5 Relative frequency of sub-topics of technical feasibility**

*5.1.4 Inspire New Ideas and Creativity.* While teams in the experimental setting spent some of their attention on *inspiring new ideas* and judging the *creativity* of the ideas, loosely using adjectives such as "creative," "original," "novel," and "innovative," they did not focus on these topics in the naturalistic setting. In addition, in both environments, the designers rarely discussed the originality of the ideas during the discussions, even when the design task explicitly asked them to generate creative ideas.

*5.1.5 User Needs, Marketing, and Company Resources.* New topics emerged in the professional team discussion in the naturalistic setting. Specifically, these discussion topics included *user needs* (40%), *marketing* (15%), and *company resources* (13%). *User needs* referred to needs that were beyond the basic functionality requirements in the design hierarchy of needs [74] such as *ease of use* (16%), *personalization* (58%), and *customization* (42%). For example, participant 3 suggested the product should be *easy to use* and said, "It's like pretty much so that the user doesn't have to worry about all of the parts that are necessary to go into the solution." Participant 4 questioned an idea due to its lack of customization since, "… if people like customization then bundling might not be the best options."

In addition, in the naturalistic setting the teams extensively discussed *marketing* strategies, especially how to achieve certain ideas through *advertising* (23%) and building *sales channels* (81%). For example, participant 2 suggested "to use Google AdWords and SEO to drive awareness around the project" and participant 4 pointed out the lack of sales channels, "… right now, our channel is distributions or one-source and they're not going to be hitting the market at all. So, we really don't have a channel to sell this through right now."

On the other hand, in the naturalistic setting the professional team also focused on *company resources*. For example, participant 4 mentioned the limited budget that the company had for redesigning the hardware, "… we want to do these things, but our budget is not going to allow us to create a new device." When a participant proposed to sell the product on Amazon, participant 2 requested information on the development time, "What does it take to get another product on Amazon? What's the process? Timeline?" In addition, participant 2 brought out the necessity of a new hire in order to further develop an idea, "We'd have to bring someone to do the online store. I don't think anyone has the expertise."

As a summary, both the experimental setting and the naturalistic setting) shared discussion topics such as technical feasibility, idea and product comparison, while also showing a lack of attention on building on generated ideas and the creativity of the ideas. However, topics such as user needs, marketing, and company resources only emerged in the discussion in the naturalistic setting. These results confirm that the environmental settings were related to the discussion topics that the teams focused on.

**5.2 In a Naturalistic Setting, Are Idea Generators and Executives Selecting Ideas Perceived to Have High Market Potential?** The analysis in the first research question indicated that in a naturalistic setting, professional teams focused on marketing, a factor that was omitted by the teams in the experimental setting. This suggested that the market performance of the ideas is an important consideration in concept screening. Aligning with this finding, the second research question looked at the individual concept screening practices and was designed to determine whether idea generators and executives selected the "right" ideas with high future market potential. We hypothesized that both idea generators ($N=4$) and executives would ($N=3$) select ideas perceived with high future market potential in a naturalistic environment. This research question was based on the ideas generated and the evaluation of the ideas in the naturalistic design

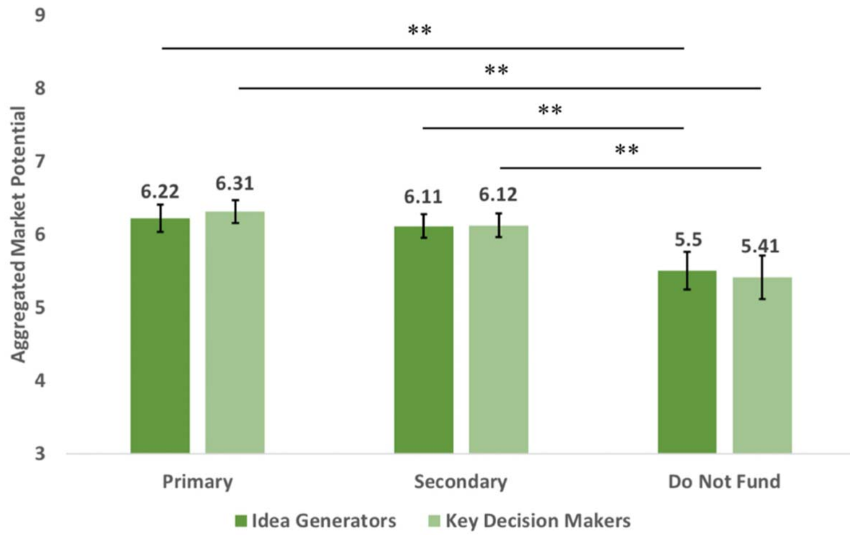



**Fig. 6 Aggregated market potential of ideas categorized by professionals. **Indicates a significant result with $p$<0.001. Error bars were calculated based on 95% CI.**

environment (study II), where the design team (idea generators) generated a total of 37 ideas.

In order to answer this question, a two-way ANOVA was calculated with the dependent variable being the aggregated market potential rated by the executives, and the independent variables being the role of the participants and the consideration types. As a reminder, the range of the aggregated market potential was 3 (low market potential) to 9 (high market potential). The results revealed a statistically significant main effect of the consideration type, $F(2, 215) = 24.19$, $p < 0.001$, partial $\eta^2 = 0.18$. Mean scores for the aggregated market potential of ideas categorized into the primary, secondary, and Do Not Fund categories were 6.26 (SD = 0.59), 6.11 (SD = 0.56), and 5.45 (SD = 0.56), respectively. Specifically, pairwise comparisons showed that ideas categorized in the Do Not Fund category had significantly lower aggregated market potential than the primary, −0.81, 95% CI [−1.09, −0.52], $p < 0.001$, and secondary categories, −0.66, 95% CI [−0.94, −0.38], $p < 0.001$. See Fig. 6 for the comparison of the means of the aggregated market potential in consideration categories determined by idea generators and the executives. There was no significant main effect of the role of the participants, $F(1, 215) < 0.001$, $p = 0.99$, partial $\eta^2 < 0.001$. The interaction effect between the role of the participant and the categorization of the ideas was also not statistically significant, $F(2, 215) = 0.32$, $p = 0.72$, partial $\eta^2 = 0.003$.

These results support our hypothesis that ideas with high market potential would be primarily considered by both idea generators and executives. These results indicated that both idea generators and executives were more likely to reject ideas with relatively low future market potential, yet there may be variability both within- and between-groups on the design criteria that informed these screening decisions.

**5.3 In a Naturalistic Setting, What Factors Predict the Screening Decisions of the Idea Generators and the Executives Individually?** While the results of the second research question indicated that both idea generators and executives filtered out ideas with low market potential, it was still unclear what other properties of the ideas were related to these screening decisions. Therefore, the third research question was developed to explore if and how perceived properties of ideas (i.e., riskiness, originality, feasibility) were related to the categorization of ideas and if this was different between idea generators ($N = 4$) and executives ($N = 3$). We hypothesized that the idea generators and the executives would use different design criteria since professionals with different roles have different responsibilities [25,26], biases [9,27], and preferences for risk in the design process [28,29].

In order to answer this, Cohen's kappa was first used to explore the level of agreement on the perceived properties of the ideas between the idea generators and the executives. The results indicated that there was fair agreement on the perceived riskiness of ideas, $\kappa = 0.295$, $p < 0.01$, and poor agreement on the perceived feasibility of ideas, $\kappa = 0.144$, $p = 0.042$, between idea generators and executives. However, no significant agreement was identified on the perceived originality between the two groups, $\kappa = 0.021$, $p = 0.76$. These results suggested that idea generators and executives had fair agreement on the perceived riskiness but not on the perceived feasibility and originality of the ideas.

Next, the dataset was split into two sub-datasets based on the role of the design professionals being idea generators or executives in order to investigate the design criteria used by each group separately. The multi-collinearity of the independent variables was tested and the requirements suggested in Kennedy [75] were met. Then, cumulative odds ordinal logistic regressions with proportional odds were run to determine the effect of perceived properties of ideas on the categorization of the ideas, respectively, for the idea generators and the executives. All independent variables in this analysis were treated as ordinal variables. For idea generators, the assumption of proportional odds was met, as assessed by a full likelihood ratio test comparing the fit of the proportional odds location model to a model with varying location parameters, $\chi^2(3) = 3.332$, $p = 0.343$. The deviance goodness-of-fit test indicated that the model was not a good fit to the observed data, $\chi^2(37) = 61.002$, $p = 0.008$, and many cells were sparse with zero frequencies in 38.1% of cells. However, the final model statistically significantly predicted the dependent variable over the intercept-only model, $\chi^2(3) = 12.106$, $p = 0.007$. Only the perceived *riskiness* had a significant effect on the prediction of the categorization of the ideas. An increase in the averaged perceived riskiness (expressed in three-point Likert scales) was associated with an increase in the odds of categorizing the idea into the Do Not Fund consideration category, with an odds ratio of 2.024, 95% CI [1.082, 3.781], Wald $\chi^2(1) = 4.875$, $p = 0.027$.

For executives, the assumption of proportional odds was met, as assessed by a full likelihood ratio test comparing the fit of the proportional odds location model to a model with varying location parameters, $\chi^2(3) = 1.593$, $p = 0.661$. The deviance goodness-of-fit test indicated that the model was a good fit to the observed data, $\chi^2(31) = 38.195$, $p = 0.175$, but most cells were sparse with zero frequencies in 37% of cells. However, the final model statistically

**Table 1 Parameter estimates of the ordinal logistic regression model**

| Role | | | Est. | Wald | df | Sig. | Exp_B | 95% CI Lower | 95% CI Upper |
|---|---|---|---|---|---|---|---|---|---|
| Idea generators | Threshold | (Type ≤ Primary) | 1.104 | 1.001 | 1 | 0.317 | 3.016 | 0.347 | 26.180 |
| | | (Type ≤ Secondary) | 3.198 | 7.787 | 1 | 0.005 | 24.484 | 2.591 | 231.366 |
| | Location | **Riskiness** | **0.705** | **4.875** | **1** | **0.027** | **2.024** | **1.082** | **3.781** |
| | | Originality | 0.231 | 0.586 | 1 | 0.444 | 1.260 | 0.697 | 2.280 |
| | | Feasibility | −0.317 | 1.259 | 1 | 0.262 | 0.728 | 0.419 | 1.266 |
| Key decision makers | Threshold | (Type ≤ Primary) | −0.859 | 0.346 | 1 | 0.556 | 0.424 | 0.024 | 7.411 |
| | | (Type ≤ Secondary) | 1.944 | 1.722 | 1 | 0.189 | 6.987 | 0.383 | 127.358 |
| | Location | **Riskiness** | 1.085 | 10.072 | 1 | 0.002 | 2.959 | 1.514 | 5.789 |
| | | Originality | −0.162 | 0.197 | 1 | 0.657 | 0.850 | 0.416 | 1.738 |
| | | **Feasibility** | **−0.799** | **4.123** | **1** | **0.042** | **0.450** | **0.208** | **0.972** |

Note: Link function: Logit.
The bold text indicates significant results.

significantly predicted the dependent variable over and above the intercept-only model, $\chi^2(3) = 32.199$, $p < 0.001$. The perceived *riskiness* and the perceived *feasibility* had significant effects on the odds of the categorization of the ideas. An increase in the perceived riskiness was associated with an increase in the odds of categorizing the idea into the Do Not Fund consideration category, with an odds ratio of 2.959, 95% CI [1.514, 5.789], Wald $\chi^2(1) = 10.072$, $p = 0.002$. An increase in the perceived feasibility was associated with a decrease in the odds of categorizing the idea into the Do Not Fund consideration category, with an odds ratio of 0.450, 95% CI [0.208, 0.972], Wald $\chi^2(1) = 4.123$, $p = 0.042$. See Table 1 for the full results.

These results support our hypothesis that idea generators and executives considered different criteria when making screening decisions. Specifically, both idea generators and executives agreed that perceived riskiness was a key factor when screening ideas, but the perceived feasibility only played a significant role only in the executives' concept screening practices.

## 6 Discussion

The purpose of the current paper was to identify how environmental settings (i.e., experimental versus naturalistic) and the role of the professionals in the design process (i.e., idea generators versus executives) related to the discussion topics and design criteria used to screen design ideas. The main findings of this paper are

- In the naturalistic setting, concept screening decisions were made primarily based on *user needs*, *technical feasibility*, *marketing*, and *company resources*. This differed from the experimental setting where the professional primarily focused on technical feasibility, comparing ideas with other generated ideas and existing products.
- Idea creativity was not a significant factor of discussion during concept screening in either of the environments explored.
- Idea generators screened ideas with low market potential and screened ideas based on their perceived risk.
- Executive screened ideas with low market potential and screened ideas based on their perceived risk and their perceived feasibility.

The following section summarizes the implications of our main findings for concept screening and identifies opportunities for supporting the concept screening practices using these findings.

**6.1 Importance of Design Environment in Concept Screening and Opportunities for Decision Support Tools.** The first contribution of this paper was the identification of similarities and differences in the discussion topics that were focused on by professional during informal concept screening in different environmental settings. First and foremost, the results showed that regardless of the design environment or task, creativity was rarely discussed in the concept screening process. This is problematic because while "mere novelty is not enough for creativity, and most agree that at least "utility" should also be evaluated (p51) [76]," the idea still needs to have originality in order to be creative. In other words, only fulfilling the basic requirement of functionality is not enough; it is the ability to differentiate from and outperform competitor products that leads to the payoff to stakeholders and the potential to gain market leadership [2,77].

In addition, these results showed that different environments may result in different research findings. Specifically, while teams in the experimental setting focused on technical feasibility, comparing ideas with other generated ideas, and existing products, the team in the naturalistic setting also discussed *user needs*, *technical feasibility*, *marketing*, and *company resources*. The attention on *marketing* related topics could have been due to the informational diversity or "differences in knowledge bases and perspectives" that members brought to the team [78], which may have promoted more diverse discussion topics. While the focus on company resources might be due to the fact that a naturalistic design problem provided the professionals with more realistic restrictions on company resources, such as investment money and development time. However, such restrictions were not provided in the design prompt in the experimental setting. These results highlight the need for tools that support the analysis of early-phase decisions based on user needs, marketing, and the use of company resources.

While the lack of focus on creativity during concept screening aligns with prior work conducted with engineering design students using the same empirical setup and design task [42], it lays in contrast to prior work with professionals. Specifically, this prior work found that professionals mentioned "creativity" and "gut feelings" during the conceptual design phase. This could be due to the fact that Maurer and Widmann [33] primarily studied companies in California, where there was an innovation culture. In addition, the conceptual design phase investigated in this study included not only the concept screening process but also the idea generation process, where generating a large number of creative ideas was the typical goal. This lack of attention to creativity is problematic, since creativity is a fundamental component in the new product development process [61].

Perhaps most importantly, these results highlight the need to explore how concept screening in naturalistic environments. However, considering the challenges of conducting in situ studies, the design community must be willing to accept research studies that have smaller sample sizes but greater depth of information, or those paired with experimental studies to triangulate findings. Such studies will ultimately lead to decision support that better translates to industry needs and thus have a larger chance of adoption and impact.

**6.2 Criteria Used in Individual Concept Screening and Opportunities for Design Tools.** The results of this study also highlight the relationship between the role of professionals and decision-making practices within an organization. Specifically, the results indicate that both idea generators and executives were able to filter out ideas with low future market potential in the naturalistic environment. In addition, the riskiness of ideas outweighed originality in the concept screening practices of both idea generators and the executives, which might lead to the filtering out of risky ideas regardless of their originality. The similarities of design criteria used by idea generators and executives could be related to the shared company culture and the realistic restrictions that were put on the project, such as limited company resources. Interestingly, this preference against risky ideas was not consistent with the interview responses of the idea generators where most of them indicated that the company provided a receptive culture where risk-taking was encouraged. For example, one of the idea generators, participant 3, mentioned that the company provided a "sandbox" for him so he did not "feel the hesitation that I do in my personal life with risks." This could be because the idea generators primarily referred to personal risks in the interview, while they considered company risks when rating the riskiness of the ideas. These company risks might regard investment money, development time, labor, and the uncertainty surrounding whether or not positive outcomes will be realized when company resources are expended. On the other hand, the executive indicated that part of the uncertainties in the fuzzy front end of the design process could be due to the market or if the customers would like the product [51], rather than the development of the technology or the embodiment of the technology.

In addition, the executive also showed a clear preference for highly feasible ideas, which the idea generators did not. This difference of the high-level design criteria used between idea generators and the executives could be due to their different roles in the design process and their different levels of familiarity to the new market. Even though the executives each had more than 25 years of experience, they were from an older generation, and the new market targeting the younger generation might seem foreign to them as mentioned by multiple idea generators. However, the idea generators could understand the new market better and have different perceptions of what ideas could be possible in the new market, since they belonged to the same generation. These results suggest that in order to keep the creative potential in the design process longer, more actions might be needed to be taken to encourage appropriate creative risk-taking at the individual level.

## 7 Conclusions, Limitations, and Future Work

The purpose of the current paper was to identify how environmental settings were related to the discussion topics that professional teams focus on when screening design ideas, and how the role of the professionals might be related to the design criteria used when making screening decisions in the naturalistic environment. Specifically, the results showed that regardless of the design environment, idea creativity (especially the originality component) was not a main topic discussed in the concept screening practice. In addition, the results highlight differences in concept screening discussion topics between naturalistic and experimental settings, and provide evidence on the types of tools needed to support decision-making in this process. Finally, the results highlight the relationship between the role of the professional within the organization and the concept screening practices. While both idea generators and executives rejected ideas with low market potential and screened ideas based on their perceived riskiness of the ideas, executives also screened ideas based on their perceived feasibility.

While the current study contributed to the understanding of the professional concept screening practices and provided future research opportunities, limitations still exist. First, the sample sizes of the two studies were small. Researchers should expand this work with a larger sample. Second, the current investigation only examined the early conceptual design stages of the project, since the project did not make it to market. It would be beneficial to investigate more R&D projects in the naturalistic environment in order to gain insights from all conditions to identify the "best practices." These conditions may include projects that are killed in the conceptual design stage, projects that are launched but fail, and projects that are launched and succeed.

Next, while the current study focused on the impact of technical, financial, and marketing considerations on the design process, these decisions may be driven by other hidden emotions and/or desires of decision makers. For example, executives closer to retirement may be less likely to embark on risky new ventures, the financial reward structure in place for executives may impact decision-making, or the timing of the decision within the fiscal quarter—e.g., decisions may impact decisions be made with reference to short-term gains like stock performance. In addition, the culture of the organization or the types of individuals within the organization may also impact decision-making. As such, further work is needed to identify how these types of factors impact decision-making.

In addition, while both studies explored small-electromechanical companies, the type of project being explored may also have impacted the discussion topics and criteria used to screen design concepts and the importance of those concepts. For example, the naturalistic study used a natural prompt for the company while the experimental study focused on a specific device and user group. This difference in task selection was due to intellectual property rights within the respective organizations. This represents some of the challenges of conducting research with engineering design companies.

There were also differences in the context of the design prompts/design tasks that may have led to differences in the data that should be further investigated, including the amount of time span of the workshop and the type of market research apparent in the design challenges. As such, while the results provide evidence of differences between these two environments, further work is needed to expand the depth and breadth of discussion topics in the screening process. This future work should be geared at studying a wider range of design companies over multiple projects within the same organization. However, do to the innate difficulties of studying naturalistic design teams, the methodologies used in these studies may need to be modified.

## Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 1351493. We also would like to thank Carolyn Reigel, Abigail O'Connell and our participants for their help in this project.

## Conflict of Interest

There are no conflicts of interest.

## Appendix: A Milk Frother Design Task

Upper management has put your team in charge of developing a concept for a *new innovative product that froths milk in a short amount of time*.

Frothed milk is the warm, pourable, virtually liquid foam that tastes rich and sweet. It is an ingredient in many coffee beverages, especially espresso-based coffee drinks (Lattes, Cappuccinos, Mochas). Frothed milk is made by incorporating very small air bubbles throughout the entire body of milk through some form of vigorous motion. As such, devices that froth milk can also be used in a number of other applications, such as for whipping cream, blending drinks, emulsifying salad dressing, and many others. Importantly frothed milk should be between 150 deg and 160 deg because above 160 deg the proteins in the milk start to breakdown causing sugars to be released which ultimately changes the flavor of the milk.

The design your team develops should be able to be used by the consumer with minimal instruction. It will be up to the board of directors to determine if your project will be carried on into production.

Once again, the goal is to develop concepts for a new, *innovative* product that can froth milk in a short amount of time. This product should be able to be used by the consumer with minimal instruction.

*Idea Generation Instructions.* Sketch your ideas in the space provided in the idea generation sheets. As the goal of this design task is not to produce a final solution to the design problem but to brainstorm ideas that could lead to a new solution, feel free to explore the solution space and focus on both the form and function of the design in order to develop innovative concepts. In other words, generate as many ideas as possible- do not focus on the feasibility or detail of your ideas. You may include words or phrases that help clarify your sketch so that your concept can be understood easily by anyone.

For clarity, please use the provided pen to generate your concepts (i.e., do not use pencil). Your participant number is included on each of the provided idea generation sheets. Generate one idea per sheet and label the idea number at the top of the sheet.

## References


[1] Hernández, H., Grassano, N., Tübke, A., Potters, L., Gkotsis, P., and Vezzani, A., 2018, "The 2018 EU Industrial R&D Investment Scoreboard," No. EUR 29450 EN, Luxembourg.
[2] Brands, R. F., and Kleinman, M. J., 2010, *Robert's Rules of Innovation: A 10-Step Program for Corporate Survival*, John Wiley & Sons, New York.
[3] Leubsdorf, B., 2017, "R&D Spending is a Rare Bright Spot in Business Investment," https://www.wsj.com/articles/r-d-spending-is-a-rare-bright-spot-in-business-investment-1477576800#livefyre-toggle-SB10660112401764113944404582394472151524232, Accessed August 1, 2019.
[4] Pahl, G., Beitz, W., Feldhusen, J., and Grote, K., 2007, *Engineering Design: A Systematic Approach*, 3rd ed., Springer Science+ Business Media Deutschland GmbH, Berlin, 2007, p. 632.
[5] Sheppard, B., Sarrazin, H., G., K., and Dore, F., 2018, "The Business Value of Design," https://www.mckinsey.com/business-functions/mckinsey-design/our-insights/the-business-value-of-design
[6] Liedtka, J., 2014, "Innovative Ways Companies are Using Design Thinking," Strategy Leadership, **42**(2), pp. 40–45.
[7] 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," J. Exp. Soc. Psychol., **42**(2), pp. 244–251.
[8] Cooper, R. G., 2008, "Perspective: The Stage-Gate® Idea-to-Launch Process—Update, What's New, and Nexgen Systems," J. Prod. Innovat. Manag., **25**(3), pp. 213–232.
[9] Schmidt, J. B., and Calantone, R. J., 2002, "Escalation of Commitment During New Product Development," J. Acad. Mark. Sci., **30**(2), pp. 103–118.
[10] Onarheim, B., and Christensen, B. T., 2012, "Distributed Idea Screening in Stage-Gate Development Processes," J. Eng. Des., **23**(9), pp. 660–673.
[11] Zheng, X., and Miller, S. R., 2018," Uncovering Ownership Bias: The Influence of Idea Goodness and Creativity on Design Professionals' Concept Selection Practices, IDETC Quebec City, Quebec, Canada, DETC2018-85964.
[12] Rietzschel, E., Nijstad, B. A., and Stroebe, W., 2010, "The Selection of Creative Ideas After Individual Idea Generation: Choosing Between Creativity and Impact," Br. J. Psychol., **101**(1), pp. 47–68.
[13] Mueller, J. S., Melwani, S., and Goncalo, J. A., 2012, "The Bias Against Creativity: Why People Desire but Reject Creative Ideas," Psychol. Sci., **23**(1), pp. 13–17.
[14] Amabile, T. M., 1996, *Creativity in Context: Update to "The Social Psychology of Creativity"*, Westview Press, New York.
[15] Sternberg, R. J., and Lubart, T. I., 1995, *Defying the Crowd: Simple Solutions to the Most Common Relationship Problems*, Free Press, New York.
[16] Jones, J., and Smith, J., 2017, "Ethnography: Challenges and Opportunities," Royal College of Nursing.
[17] Nikander, J. B., Liikkanen, L. A., and Laakso, M., 2014, "The Preference Effect in Design Concept Evaluation," Des. Stud., **35**(5), pp. 473–499.
[18] Toh, C., and Miller, S., 2014, "The Role of Individual Risk Attitudes on the Selection of Creative Concepts in Engineering Design," ASME Design Engineering Technical Conferences Buffalo, NY, DETC2014-35106.
[19] Toh, C. A., and Miller, S. R., 2016, "Creativity in Design Teams: The Influence of Personality Traits and Risk Attitudes on Creative Concept Selection," Res. Eng. Des., **27**(1), pp. 73–89.
[20] Toh, C. A., and Miller, S. R., 2016, "Choosing Creativity: The Role of Individual Risk and Ambiguity Aversion on Creative Concept Selection in Engineering Design," Res. Eng. Des., **27**(3), pp. 195–219.
[21] Toh, C., Strohmetz, A., and Miller, S., 2016, "The Effects of Gender and Idea Goodness on Ownership Bias in Engineering Design," ASME J. Mech. Des., **138**(10), p. 101105.
[22] Nikander, J. B., Liikkanen, L. A., and Laakso, M., 2013, "Naturally Emerging Decision Criteria in Product Concept Evaluation," Proc. DS 75-7: Proceedings of the 19th International Conference on Engineering Design (ICED13), Design for Harmonies, Vol. 7: Human Behaviour in Design, Seoul, Korea, 19-22.08.
[23] Ahmed, S., and Christensen, B. T., 2009, "An in Situ Study of Analogical Reasoning in Novice and Experienced Design Engineers," ASME J. Mech. Des., **131**(11), p. 111004.
[24] Toh, C., Patel, A., Strohmetz, A., and Miller, S., 2015, "My Idea is Best! Ownership Bias and Its Influence in Engineering Concept Selection," Design Engineering Technical Conferences, Boston, MA, DETC2015-46478.
[25] Kijkuit, B., and van der Ende, J., 2007, "The Organizational Life of an Idea: Integrating Social Network, Creativity and Decision-Making Perspectives," J. Manag. Stud., **44**(6), pp. 863–882.
[26] Cooper, R. G., 1990, "Stage-gate Systems: A New Tool for Managing New Products," Bus. Horiz., **33**(3), pp. 44–54.
[27] Zheng, X., and Miller, S. R., 2019, "Is Ownership Bias Bad? The Influence of Idea Goodness and Creativity on Design Professionals Concept Selection Practices," ASME J. Mech. Des., **141**(2), p. 021106.
[28] Cooper, A. C., Woo, C. Y., and Dunkelberg, W. C., 1988, "Entrepreneurs' Perceived Chances for Success," J. Bus. Ventur., **3**(2), pp. 97–108.
[29] Gilmore, A., Carson, D., and O'Donnell, A., 2004, "Small Business Owner-Managers and Their Attitude to Risk," Market. Intell. Plan., **22**(3), pp. 349–360.
[30] Kudrowitz, B. M., and Wallace, D., 2013, "Assessing the Quality of Ideas From Prolific, Early-Stage Product Ideation," J. Eng. Des., **24**(2), pp. 120–139.
[31] Huang, H.-Z., Liu, Y., Li, Y., Xue, L., and Wang, Z., 2013, "New Evaluation Methods for Conceptual Design Selection Using Computational Intelligence Techniques," J. Mech. Sci. Technol., **27**(3), pp. 733–746.
[32] Baucus, M. S., Norton, M. I., Baucus, D. A., and Human, S. E., 2007, "Fostering Creativity and Innovation Without Encouraging Unethical Behavior," J. Bus. Ethics, **81**(1), pp. 97–115.
[33] Maurer, C., and Widmann, J., 2012, "Conceptual Design Theory in Education Versus Practice in Industry: A Comparison Between Germany and the United States," Design Engineering and Technical Conferences, Chicago, IL, Aug. 12–15, pp. 277–283.
[34] Pugh, S., 1991, *Total Design: Integrated Methods for Successful Product Engineering*, Addison-Wesley, Workingham.
[35] Akao, Y., 1994, "Development History of Quality Function Deployment," The Customer Driven Approach to Quality Planning and Deployment, Asian Productivity Organization, Minato, Tokyo, p. 339.
[36] Marsh, E. R., Slocum, A. H., and Otto, K. N., 1993, *Hierarchical Decision Making in Machine Design*, MIT Precision Engineering Research Center, Cambridge, MA.
[37] Saaty, T. L., 1994, "How to Make a Decision: The Analytic Hierarchy Process," Interfaces, **24**(6), pp. 19–43.
[38] López-Mesa, B., and Bylund, N., 2011, "A Study of the Use of Concept Selection Methods From Inside a Company," Res. Eng. Design, **22**(1), pp. 7–27.
[39] Toh, C., Miele, L., and Miller, S., 2015, "Which One Should I Pick? Concept Selection in Engineering Design Industry," Design Engineering Technical Conferences, Boston, MA, DETC2015-46522.
[40] Salonen, M., and Perttula, M., 2005, "Utilization of Concept Selection Methods: A Survey of Finnish Industry," ASME Design Engineering Technical Conferences, Long Beach, CA, pp. 527–535.
[41] Dym, C. L., Wood, W. H., and Scott, M. J., 2002, "Rank Ordering Engineering Designs: Pairwise Comparison Charts and Borda Counts," Res. Eng. Des., **13**(1), pp. 236–242.
[42] Toh, C. A., and Miller, S. R., 2015, "How Engineering Teams Select Design Concepts: A View Through the Lens of Creativity," Des. Stud., **38**, pp. 111–138.

[43] Starkey, E., Toh, C. A., and Miller, S. R., 2016, "Abandoning Creativity: The Evolution of Creative Ideas in Engineering Design Course Projects," Des. Stud., **47**, pp. 47–72.
[44] Paulus, P. B., 2000, "Groups, Teams, and Creativity: The Creative Potential of Idea-Generating Groups," Appl. Psychol., **49**(2), pp. 237–262.
[45] Naquin, C. E., and Tynan, R. O., 2003, "The Team Halo Effect: Why Teams are Not Blamed for Their Failures," J. Appl. Psychol., **88**(2), pp. 332–340.
[46] Alnuaimi, O. A., Robert, L. P., and Maruping, L. M., 2010, "Team Size, Dispersion, and Social Loafing in Technology-Supported Teams: A Perspective on the Theory of Moral Disengagement," J. Manag. Inform. Syst., **27**(1), pp. 203–230.
[47] Larey, T. S., and Paulus, P. B., 1999, "Group Preference and Convergent Tendencies in Small Groups: A Content Analysis of Group Brainstorming Performance," Creat. Res. J., **12**(3), pp. 175–184.
[48] Cagan, J., and Vogel, C. M., 2002, *Creating Breakthrough Products: Innovation From Product Planning to Program Approval*, Ft Press.
[49] Cooper, R. G., and Kleinschmidt, E. J., 1996, "Winning Businesses in Product Development: The Critical Success Factors," Res. Technol. Manag., **39**(4), pp. 18–29.
[50] Cushman, W. H., and Rosenberg, D. J., 1991, *Advances in Human Factors/ Ergonomics*, Elsevier.
[51] Danneels, E., and Kleinschmidtb, E. J., 2001, "Product Innovativeness From the Firm's Perspective: Its Dimensions and Their Relation With Project Selection and Performance," J. Prod. Innov. Manage., **18**(6), pp. 357–373.
[52] Bylund, N., "Concept Selection in the Automotive Industry With Examples," Proc. DS 36: Proceedings DESIGN 2006, The 9th International Design Conference, Dubrovnik, Croatia.
[53] Zhang, P., Sadler, C. M., Lyon, S. A., and Martonosi, M., 2004, "Hardware Design Experiences in ZebraNet," Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems, Baltimore, MD, November, ACM, pp. 227–238.
[54] Ha, L., and Ganahl, R., 2004, "Webcasting Business Models of Clicks-and-Bricks and Pure-Play Media: A Comparative Study of Leading Webcasters in South Korea and the United States," Int. J. Media Manage., **6**(1–2), pp. 74–87.
[55] Batory, D., and O'malley, S., 1992, "The Design and Implementation of Hierarchical Software Systems With Reusable Components," ACM Trans. Soft. Eng. Methodol., **1**(4), pp. 355–398.
[56] Simonson, I., 1993, "Get Closer to Your Customers by Understanding How They Make," Calif. Manage. Rev., **35**(4), pp. 68–84.
[57] Kim, E., Chung, J., Beckman, S., and Agogino, A. M., 2016, "Design Roadmapping: A Framework and Case Study on Planning Development of High-Tech Products in Silicon Valley," ASME J. Mech. Des., **138**(10), p. 101106.
[58] Arnold, H., Erner, M., Mckel, P., and Schlf˜fer, C., 2010, *Applied Technology and Innovation Management*, Springer, New York.
[59] Fischer, G., 2013, "Learning, Social Creativity, and Cultures of Participation," *Learning and Collective Creativity: Activity-Theoretical and Sociocultural Studies*, A. Sannino, and V. Ellis, eds., Routledge, New York, NY, p. 198.
[60] Shah, J. J., Vargas-Hernandez, N., and Smith, S. M., 2003, "Metrics for Measuring Ideation Effectiveness," Des. Stud., **24**(1), pp. 111–134.
[61] Goel, P. S., and Singh, N., 1998, "Creativity and Innovation in Durable Product Development," Comput. Ind. Eng., **35**(1), pp. 5–8.
[62] Abbey, A., and Dickson, J. W., 1983, "R&D Work Climate and Innovation in Semiconductors," Acad. Manage. J., **26**(2), pp. 362–368.
[63] Unsworth, K. L., and Clegg, C. W., 2010, "Why do Employees Undertake Creative Action?," J. Occup. Organ. Psychol., **83**(1), pp. 77–99.
[64] Weber, E. U., and Milliman, R. A., 1997, "Perceived Risk Attitudes: Relating Risk Perception to Risky Choice," Manage. Sci., **43**(2), pp. 123–144.
[65] Mellers, B. A., Schwartz, A., and Weber, E. U., 1997, "Do Risk Attitudes Reflect in the Eye of the Beholder?" *Choice, Decision, and Measurement: Essays in Honor of R. Duncan Luce*, A. A. J. Marley, ed., Routledge, New York, pp. 57–71.
[66] Schwartz, S., and Griffin, T., 1986, "Evaluating Medical Information," *Medical Thinking*, Springer, New York, pp. 56–96.
[67] Pommer, A., 2015, "How do Firms Evaluate and Select Idea," Masters of Science, Master's thesis, University of Zurich.
[68] Elo, S., and Kyngäs, H., 2008, "The Qualitative Content Analysis Process," J. Adv. Nurs., **62**(1), pp. 107–115.
[69] QSR, I. P. L., 2015, "NVivo Qualitative Data Analysis Software Version 11.1.1".
[70] Landis, J. R., and Koch, G. G., 1977, "The Measurement of Observer Agreement for Categorical Data," Biometrics, **33**(1), pp. 159–174.
[71] West G. P., III, 2007, "Collective Cognition: When Entrepreneurial Teams, Not Individuals, Make Decisions," Entrep. Theory Pract., **31**(1), pp. 77–102.
[72] QSR, I. P. L., 2018, "NVivo Qualitative Data Analysis Software Version 12".
[73] IBM, Released 2017, "IBM SPSS Statistics for Mac, Version 25.0," IBM Corp., Armonk, NY.
[74] Bradley, S., 2010, "Designing for a Hierarchy of Needs", https://www.smashingmagazine.com/2010/04/designing-for-a-hierarchy-of-needs/
[75] Kennedy, P., 1998, *A Guide to Economics*, MIT Press, Cambridge, MA.
[76] Brown, D., 2013, "Developing Computational Design Creativity Systems," Int. J. Des. Creat. Innov., **1**(1), pp. 43–55.
[77] Dewar, R. D., and Dutton, J. E., 1986, "The Adoption of Radical and Incremental Innovations: An Empirical Analysis," Manage. Sci., **32**(11), pp. 1422–1433.
[78] Jehn, K. A., Northcraft, G. B., and Neale, M. A., 1999, "Why Differences Make a Difference: A Field Study of Diversity, Conflict and Performance in Workgroups," Adm. Sci. Q., **44**(4), pp. 741–763.