Prototyping for context: exploring stakeholder feedback based on prototype type, stakeholder group and question type

Michael Deininger, Shanna R. Daly, Jennifer C. Lee, Colleen M. Seifert & Kathleen H. Sienko

00163

Research in Engineering Design

Theory
Applications,
and Methodology

ONLINE FIRST

🌫 Springer

🌫 Springer

Springer

Prototyping for context: exploring stakeholder feedback based on prototype type, stakeholder group and question type

Michael Deininger[1] · Shanna R. Daly[2] · Jennifer C. Lee[1] · Colleen M. Seifert[3] · Kathleen H. Sienko[4]

Abstract

Engineering designers frequently use prototypes to gather input from stakeholders. Design guidelines recommend the use of quick and simple prototypes early and often in a design process. However, the type and quality of a prototype can influence how stakeholders perceive a new design concept and can, therefore, impact their responses. Additionally, different levels of experience, expertise, and preparedness for providing input to designers may lead stakeholders from different geographical or cultural settings to provide different responses, making the format of a prototype even more influential. Although design practitioners are known to intentionally align their prototyping approaches with the specific design question to be answered, it is unclear the extent to which prototyping approaches should vary based on the stakeholder, context, and setting of a design project. To investigate how the format and quality of prototypes influence stakeholders' responses, we conducted a field study with various medical professionals in Ghana. We presented prototypes for a medical device in different formats to stakeholders and collected responses to the design through semi-structured interviews. Our findings indicate that professional expertise, prototype format, and question type influenced the types of responses that stakeholders provided. These findings suggest that designers seeking input from stakeholders on new concepts should consider context-specific prototyping strategies, especially when designing at distance and across cultures.

Keywords Design decisions · Product design · Prototypes · Stakeholders · User behavior

1 Motivation

Various factors, including the visual appearance of a prototype, can influence how individuals perceive the objects or ideas to which they are introduced. Even though prototypes may not always reflect the actual quality or utility, they are critical factors that can impact the perception of a product and possibly motivate a decision to purchase or use a product

(Desmet and Hekkert 2007; Sauer and Sonderegger 2009). In design, when new ideas are shared with stakeholders—the individuals, groups or organizations that have direct or indirect interest in the product—through prototypes, several factors contribute to how these new ideas are perceived. Prototypes serve as vehicles for designers to communicate their thoughts to others, but the nature and level of refinement of a prototype often depends on the stage of a project (Atman et al. 2007; Christie et al. 2012; Menold et al. 2017). Prototypes used in the early stages of a design process might include conceptual sketches and crude mockups, while prototypes used during later stages might consist of more refined models with virtually indistinguishable properties from the production part (Crismond and Adams 2012; Hilton et al. 2015).

Stakeholders are important contributors to a design process, and designers need to consider the context of their stakeholders as well as the product itself. Consideration of stakeholders' backgrounds extends to the elicitation of feedback from stakeholders on prototypes. Stakeholders can vary in their backgrounds and experiences, which can influence

✉ Shanna R. Daly
     srdaly@umich.edu

✉ Kathleen H. Sienko
     sienko@umich.edu

1   1305 George G. Brown Laboratory, University of Michigan, 2350 Hayward Street, Ann Arbor, MI 48109, USA

2   3316 George G. Brown Laboratory, University of Michigan, 2350 Hayward Street, Ann Arbor, MI 48109, USA

3   3042 East Hall, University of Michigan, 530 Church Street, Ann Arbor, MI 48109, USA

4   3454 George G. Brown Laboratory, University of Michigan, 2350 Hayward Street, Ann Arbor, MI 48109, USA

the type and depth of design input they provide on prototypes. For example, some stakeholders may be focused on appearance, while others may be more focused on function or the underlying idea. As a result, promising design concepts might be overlooked by some stakeholders because of a less favorable presentation, while other, less-promising concepts might elicit false-positive responses motivated by a more refined form of presentation (Hekkert et al. 2003; Leder and Carbon 2005). The notion that stakeholders should therefore be presented only with highly refined prototypes is challenged by findings that prototypes that can be perceived as finished products might convey the impression that input is no longer needed or even possible since a great amount of time and energy has already been invested in the design (Viswanathan and Linsey 2011). Furthermore, products for global markets are typically designed at distance, and geographic, time and cultural differences may pose additional challenges to designers (Scrivener et al. 1993).

While factors such as project setting, stakeholder level of experience, motivation and investment in the project might be difficult for designers to influence, they *can* exercise control over the type and quality of the prototypes they share, as well as the questions they ask of those from whom they seek input. It is therefore critical that designers identify the presentation format and question type most appropriate for stakeholder interaction. In this study, we investigated if and how stakeholder input on a medical device concept was influenced by type of prototype, type of question, and group membership.

2 Background

2.1 Timing and fidelity of prototypes in a design process

Engineering designers often use prototypes as tools for testing and validation. However, multiple studies have shown that prototypes can be useful throughout an entire design process (De Beer et al. 2009; Moe et al. 2004; Viswanathan and Linsey 2011; Yang and Epstein 2005). For example, while functional prototypes such as 3D-printed models and CAD models are frequently used for functional testing later in the process (Baxter 1995), they can also be useful during front end design, the phases of design most commonly associated with problem identification and definition and concept development (Camburn et al. 2013; Christie et al. 2012; Kelley 2001). Even though designers might use simple prototypes such as sketches and mockups primarily early in the design process to quickly inspire, communicate, elicit input and select from new ideas (Brandt 2007; Campbell et al. 2007; Gerber 2009; Houde and Hill 1997; Kelley 2007), they too can be helpful later in the process.

Design experts frequently call for a minimalistic, or "quick and simple" approach to prototyping, constructing the quickest and cheapest prototype that still satisfies a particular requirement, e.g., the communication of an idea (Kelley 2001; Moogk 2012). Low-fidelity prototypes such as sketches and cardboard models are often intentionally simple, incomplete, and sometimes crude representations that convey some critical characteristics of the intended end product. They can be created quickly and inexpensively and allow designers to share and evaluate a large number of ideas. This quick and simple approach enables iteration and decision-making early in the design process and the selection of the most promising ideas before substantial "sunk costs," i.e., time and money, are incurred (Arkes and Blumer 1985). Higher fidelity prototypes, such as 3D-printed models and CAD models, that require additional resources such as time, skills and money to create are typically reserved for later stages in the design process, when functional and/or simulated testing is necessary (Dieter and Schmidt 2012; Rudd et al. 1996).

While collecting useful input from stakeholders can be challenging for designers (Castillo et al. 2012; Mohedas et al. 2015a), prototypes can serve to facilitate these interactions. For example, prototypes are increasingly being used during the earliest stages of a design process to support the elicitation of product requirements from stakeholders (Kelley 2007; Schrage 1999; Yock et al. 2015). Here, representations of diverse preliminary concept solutions might be shared with stakeholders to facilitate and support requirement elicitation interviews. Demonstrating ideas to stakeholders through the use of prototypes is preferable to providing a verbal description alone and is especially critical early in a project when designers are developing an understanding of stakeholder needs and wants—often across professional and geographical cultures (Jensen et al. 2017; Kelley and Littman 2006; Scott 2008). In these situations, prototypes can serve as shared objects that support communication, engage stakeholders in the design process, allow them to better express their opinions, and define requirements that designers might not otherwise discover or that can be difficult to elicit. The new insight into the problem as well as the solution space of a design project can introduce elements of surprise, and new and unexpected circumstances can lead to problem–solution co-evolution that often inspires creative solutions (Dorst and Cross 2001).

For example, in a study in sub-Saharan Africa, Sabet-Sarvestani and Sienko (2014) noted that the quantity and quality of responses to requirement elicitation interview questions dramatically increased when the team presented physical and functional prototypes to stakeholders compared to theoretically grounded elicitation questions that minimized bias by prompting stakeholders to provide responses about a hypothetical concept solution. Earlier conversations

with stakeholders had not provided critical insight into the cultural viewpoints and concerns about an adult male circumcision device, but when the research team introduced physical prototypes, participants started to interact with the models, compared concepts, discussed differences and provided input about both the concepts and culturally relevant information that would affect implementation if not fully captured in the product requirements. This degree of insight could not have been gathered through interviews alone; it only transpired through discussions and observations supported by stakeholders' interactions with physical prototypes.

Prototypes can also be invaluable tools for exploring design details and identifying potential issues early in a design process (Jensen et al. 2017). Often, the level of refinement, detail, and functionality of a prototype increases as designers develop a deeper understanding about the solution space and build on what they learned from earlier iterations (Ulrich and Eppinger 2015; Yang and Epstein 2005). Consequently, early prototypes do not always represent the quality and functionality of the intended end product, and stakeholders' perceptions of a new idea might potentially be negatively influenced by the nature and level of refinement of the prototype with which they are presented (Crilly et al. 2004; Hare et al. 2013; Lim et al. 2006).

Simply increasing presentation quality and functionality of a prototype, however, does not automatically lead to better input from stakeholders. Recent studies in the field of human–computer interaction concluded that a balance between quality and functionality of prototypes might be most beneficial for the collection of input from stakeholders (Hare et al. 2013; Lim et al. 2006). The authors further emphasized that the context surrounding the prototype feedback session, such as task scenarios, social and physical circumstances, as well as the participants themselves, can influence the type and quality of stakeholder input. For example, in a study by Sauer and Sonderegger (2009) examining the influence of prototype fidelity on user behavior, participants were presented with low-, medium-, and high-quality prototypes of cell phones and asked to perform tasks such as sending a text message and suppressing a phone number. The researchers found that the more attractive prototypes positively affected user emotions and consequently their judgment of usability of a concept. In a human–computer interface (HCI) study with simulated automatic teller machines (ATMs), perceived usability was strongly related to the perceived beauty of a design—the more beautiful participants rated a layout, the more usable they thought it was (Tractinsky et al. 2000). In another study, participants judged the creativity of ideas for new toaster concepts represented by sketches (Kudrowitz et al. 2012). The concepts represented by the highest quality sketches were most likely to be ranked as the most creative ideas.

2.2 Variation in stakeholder feedback in response to prototypes

Scholars in fields that leverage representations, such as the sciences, describe a variety of roles that representations can have in supporting processes and outcomes within the discipline. Models in science have been used for a variety of purposes including visualizing, forming hypotheses, critiquing ideas, examining theories, and deriving relationships (Daly and Bryan 2010; Giere 2004; Morgan and Morrison 1999; Seidewitz 2003). Similarly, in design domains, design professionals use prototypes in a variety of ways and follow many common best practice recommendations for using them to support design decision-making (Lauff et al. 2017). However, it is unclear the extent to which commonly accepted best practices (Deininger et al. 2017) are directly transferable across contexts, cultures, stakeholder characteristics, or environments of design projects.

Cultural norms are another factor that may direct a stakeholder's focus on particular aspects of a prototype. In a study evaluating cultural differences of consumer purchasing behavior, stakeholders from one cultural group (Singapore) focused more on the functionality of the product, while stakeholders from another cultural group (Philippines) valued aesthetics more when making purchasing decisions (Seva and Helander 2009).

Variation in experiences can also play a role in a stakeholder's ability to give feedback. In the art domain, naïve reviewers exhibit a tendency to stereotype based on personal taste (Parsons 1989), and while novices in any field tend to have more emotional reactions, experts tend toward cognitive responses that lead to a more analytical way of reviewing an unfamiliar object (Winston and Cupchik 1992). In physics, a study of novices and experts found that the understanding of examples differed based on expertise (Chi et al. 1981). Here, novices grouped physics problems together because they included "ramps," while experts defined a category as "work problems." This finding illustrates that differing levels of experience and expertise in a domain results in differences in how new examples are perceived. Translated to the design domain, when a stakeholder does not have experience or competency in a specific domain, he/she may not know how and what to look for when providing feedback about a design. For those with less domain experience, being asked for their feedback about a design may feel overwhelming, which can lead to frustration and put a stakeholder in a negative affective state about how they feel toward the object in question (Frijda 1989). This negative emotional response can then influence how a stakeholder processes and ultimately evaluates new information (Scherer 2003).

Stakeholders also have a variety of motivations and experiences not related to their expertise in a particular subject matter, which can impact how they respond to a prototype

(Chamorro-Koc et al. 2009). Different emphases in feedback given by stakeholders may be related to different interpretations of the affordances of a product as represented by the prototype. Affordances are properties of an object that suggest possible interactions of the user with the object. For example, the shape of a lever may imply that it should be pushed rather than pulled, or the shape of a knob might suggest turning rather than sliding (Gibson 1966). As affordances are perceivable actions, or actions that are considered possible by the user, affordances that users identify depend on prior experiences and knowledge (Norman 1990). Thus, stakeholder feedback may vary based on the affordances they interpret from the presented prototype, and the form and functionality represented in that prototype.

Finally, the questions asked of stakeholders can prompt variation in the types of feedback they provide (Creswell 2013; Patton 2014; Weiss 1995). Specifically, questions that are positioned outside of a stakeholder's expertise can negatively influence their response (Leder et al. 2004). Therefore, interview questions need to be carefully designed to extract unbiased information from stakeholders and should consider cultural context (Glesne and Peshkin 1991). Questions should aim to understand the underlying theory of a response and allow an interviewer the flexibility to adjust and avoid negative experiences such as boredom, annoyance, or even physical discomfort by the stakeholders (Silverman 2010). Semi-structured interview questions can provide a framework to guide a conversation while allowing for flexibility for both the interviewer and interviewee to explore and expand on interesting information (Patton 2014).

3 Methods

Our study focused on one product category (medical devices) and multiple stakeholder groups (nurses, medical students, and medical doctors) in one cultural context (Ghana) to provide initial insights into how prototype type, group membership (stakeholder characteristics) and question type can influence stakeholders' perceptions of a design concept and the resulting feedback they provide. The research questions that guided our work included:

- How does prototype format influence stakeholder feedback?
- How does group membership influence stakeholder feedback?
- How does question type influence stakeholder feedback?

The medical device product category was pursued for this study because there are known challenges in designing and implementing products for global health settings (Free 2004; Howitt et al. 2012; World Health Organization 2010). The study was performed in Ghana because of existing partnerships with multiple tertiary healthcare facilities; furthermore, performing the study in Ghana enabled us to conduct the interviews in English, one of Ghana's official languages.

3.1 Participants

Forty-five healthcare professionals from a teaching hospital in Ghana were recruited for participation in this study. They included 18 nurses or midwives, 10 medical students training to become medical doctors in obstetrics and gynecology, and 17 medical doctors. These participants represented a cross section of the target stakeholder groups, are likely the most easily accessible respondents to design teams working in similar settings, and would either be using, advising, or training others in the use of the proposed device. The participants were recruited by the family planning department of the hospital and received a small gift for their participation (pen, mini-flashlight, or USB memory stick). All participants were aware of long-term contraceptive implants, but none were familiar with the assistive insertion device concept used in this study or had seen it before.

3.2 Data collection

We introduced participants who were stakeholders in this cultural context, Ghana, to the design of a medical device concept that assists with the insertion of a long-term contraceptive implant. Long-term contraceptive implants are particularly appealing in resource-limited settings where patients have limited access to healthcare providers (Funk et al. 2005). A small polymer rod is implanted into the subcutaneous tissue on the inside of the upper arm of the patient. Properly inserted, the rod releases hormones into the woman's blood stream and, in contrast to oral contraceptives, does not require regular visits and monitoring by an obstetrician-gynecologist. The implants provide contraception for extended time periods, between 3 to 5 years, depending on the manufacturer. However, if not inserted properly, the rod can become embedded in the muscle tissue and complicate removal, sometimes even requiring a surgical procedure. Proper insertion is, therefore, critical and is typically performed by trained healthcare professionals such as doctors and nurses. The proposed concept represents a task-shifting device (McPake and Mensah 2008) that acts as a needle guide (Mohedas et al. 2015b). It allows lesser trained healthcare providers like community health workers (CHWs) to perform correct insertions in rural areas with limited access to healthcare. This simple, low-cost device was first conceived by mechanical engineering students during a capstone design course and is representative of projects in which designers might seek input from a variety of stakeholders, from government officials to rural healthcare workers.

Fig. 1 The types of prototypes used in this study: **a** paper sketch, **b** cardboard mockup, **c** CAD model, **d** 3D-printed model
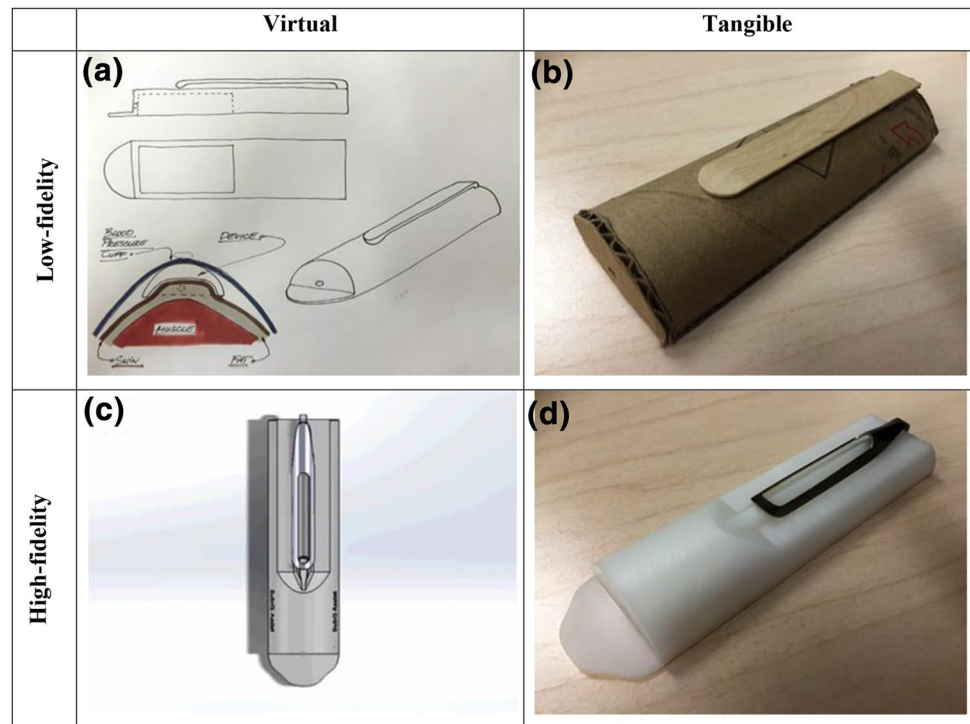
The device concept was presented through various prototypes that are commonly used during design. The four representations included a sketch, a cardboard mockup, an animated (rotating) CAD model and a 3D-printed, production-like representation of the device. The sketch and the CAD model were virtual, i.e., non-physical, representations that were shown either in low-fidelity, paper form (sketch) or on a laptop screen (high-fidelity animated CAD model). The cardboard mockup (low-fidelity) and the 3D-printed model (high fidelity) were physical objects that were given to the participants for examination. The prototypes are shown in Fig. 1.

Each participant was first shown one prototype—either a low-fidelity prototype (sketch or cardboard mockup) or a high-fidelity prototype (CAD or 3D-printed model) and then asked a series of questions to elicit feedback. These questions were developed to be consistent with questions designers typically ask when gathering input from stakeholders (Kelley and Littman 2006). The questions prompted participants to comment on several aspects of the device design and afforded the interviewer the opportunity to follow up when clarification was needed. Nine questions were asked, designed to elicit participants' impressions of the device and to encourage them to critique or add to the proposed design they reviewed. A full list of questions can be found in Appendix 1. Prior to collecting data, pilot interviews were conducted at a large, Midwestern University in the United States to test and refine the questions and the prototypes shown.

After the first prototype was shown and questions asked, each participant was then presented with a second prototype of the same device but from another fidelity group and asked the same nine questions again. Introducing both low- and high-fidelity prototypes to the participants helped to minimize answer biases caused by the nature of the prototypes. The order and type of prototype presented to participants were randomly assigned.

All interviews were conducted during a 1-week period and were carried out by the same researcher in English. All interviews except one were audio recorded and later transcribed for analysis. One participant did not agree to the use of an audio recorder and handwritten notes of the interview were taken instead.

3.3 Data analysis

After the interview data were collected, the audio files were transcribed for analysis. Three analytical methods were used to determine the usefulness of the answers that participants provided. These included (1) a deductive coding scheme that we developed to categorize the type of input elicited, (2) a modified version of the consensual assessment technique (CAT) to capture the quality of the input provided by individual participants (Baer and McKool 2009; Kaufman et al. 2008), and (3) a count of the number of words in participants' responses to each question.

Table 1 Rating rubric for deductive coding of response type

Category	Code	Definition	Example
A	Answer with design input	Provides input/suggestions beyond just answering the question	"So maybe it should be designed in [different] sizes"
B	Justified answer	Answers and explains why	"I like the ability of the device to isolate the skin and then the subcutaneous tissue from the muscle"
C	Unjustified answer	Answers affirmatively but provides no explanation	"Yes" or "No"
D	Non-useful answer	Provides no answer, is unsure, or answer was contradicting/made no sense	"I can't say" or "if you get it right then it will work" or "you're the designer"

3.3.1 Metrics

3.3.1.1 Deductive coding of response type We developed a coding rubric to categorize responses. This rubric consisted of four answer categories as shown in Table 1. Category A answers included design input as part of the response. Here participants provided design input by suggesting, for instance, that a change in color, size or material should be considered. Category B answers consisted of the participants' opinion backed by justification as to why they felt a certain way. Category C answers comprised unjustified answers that represented the opinion of the respondent only. Category D answers were non-useful and included statements by participants referring back to the design team's expertise rather than giving their own opinion.

Two researchers completed multiple rounds of coding and, once the individual answers were assigned a category, the results were normalized by adjusting the counts to a common scale to account for the different numbers of entries in each group. In a few cases (seven), a single interview question was not asked, and the missing data were removed, resulting in 99% valid answers across all interview questions. The inter-rater reliability for this coding activity calculated with Cronbach's alpha was 0.943, which is considered substantial agreement (Landis and Koch 1977). The power of the analysis for detecting a medium effect of $w = 0.3$ as defined by Cohen's effect size index (Cohen, 1977) was 80%, 53% and 78% for nurses, doctors and students, 59%, 58%, 58% and 60% for sketch, cardboard mockup, CAD model and 3D-printed model, respectively.

3.3.1.2 Consensual assessment technique to assess usefulness of responses We also evaluated stakeholder input by engaging subject matter experts in an assessment procedure called the Consensual Assessment Technique (CAT) (Amabile 1983; Amabile et al. 1996; Baer et al. 2004; Howard et al. 2008). This technique, commonly used to rate the creativity of products like paintings and poems, draws on a large number of experts who are presented with multiple artifacts. The experts are asked to rate the artifacts relative to one another on a Likert scale (Kaufman et al. 2008) based on a common criterion like creativity, composition, or use of color. Contrary to the deductive coding approach, the raters are not provided with detailed instructions. Instead, raters develop their own justification for why they think one artifact should be rated higher, lower, or the same as another. Research has shown that even with the lack of detailed instructions, or a request to justify their decisions, a large degree of agreement can commonly be found among subject matter experts (Amabile 1983). According to Landis and Koch (1977), a reliability between 0.61 and 0.80 is substantial, while agreements above 0.80 are considered almost perfect. Ideally, a large group of experts (for example 30) would judge a small sample of work (maybe two or three items). However, it is often not practical, and as a result, a small number of experts are often recruited to evaluate a large body of work (Kaufman et al. 2008).

To determine if the CAT would provide consistent results with our deductive coding approach, we recruited five subject matter experts (designers with several years of experience in product design and/or medical device development) to rate the usefulness of the input participants provided. A key aspect of the CAT is that the raters define their own criteria (Amabile 1983) and, therefore, the term "usefulness" was not defined beyond asking reviewers to rate how useful they thought an answer was for the design of the device. The answers to the individual questions from the first round of the interviews were printed and given to the reviewers in four sets to allow them to physically sort the responses. The four sets of data given to the raters consisted of:

- Individual responses to question 3,
- Individual responses to question 6,
- Individual responses to question 9, and
- Individual responses to all questions, i.e., the entire transcript of the first round of interviews.

Due to the significant amount of time required to complete this activity, the experts were only asked to rate answers for three individual questions (see 3.3.2) and all nine answers combined from the first round of interviews. The experts were instructed to rate all 45 answers for each of the four sets on a 1–5 Likert scale indicating how useful they thought the answers were for improving the design

(with 1 being the least useful and 5 being the most useful). Participants were asked to utilize the full scale (1–5) and perform their rating for all four sets of data independently. The raters performed three rounds of ordering for each set to ensure that they were satisfied with their final selection. Raters reported needing between 8 and 10 h to complete the rating activities. Once the experts rated the data, Cronbach's alpha was calculated to measure consistency among the five experts. In this study, the agreement was 0.914 across all interview questions and 0.957 for questions 3, 6 and 9, representing significant agreement for all four data sets.

3.3.1.3 Word count In addition to the previous techniques, a word count analysis was conducted to determine if the volume of words contained in an answer provided any indication of value, here defined as usefulness, of the responses. If so, this technique would require the least effort for analysis and would be easiest to perform. Thus, we investigated the relationship of word count to the other evaluation techniques.

3.3.2 Treatment of data

We analyzed metrics for all interview questions first, followed by three individual questions to investigate if the type of question affected stakeholder feedback. Three individual questions were identified during data analysis to explore the potential effects of question type on stakeholder responses. The chosen questions represented distinct areas of interest to inform design decisions: critique of the idea in general (question 3: Do you think this concept would work?), what the patient receiving the implant would think of the device (question 6: How do you think patients would feel about this device being used during the implant procedure?), and device-related design input (question 9: What would you suggest changing about this device?).

Due to the relatively low number of participants and high number of variables in this study, the results from the response type analysis were combined into category A or B answers and category C or D answers. Additionally, non-physical prototypes (sketch and CAD) were combined into "virtual" prototypes, and physical prototypes (mockup and 3D printed) were combined into "tangible" prototypes. Collapsing answer and prototyping categories for the response type analysis amplified the statistical power for the subsequent analyses. Both CAT and word count produced numerical values and did not require any collapsing. We focused on response type as the primary analytical output and referenced both CAT as well as word count when findings from these methods were significant.

Except for the CAT, where due to the volume of data only responses from the first prototype feedback session were included, answers from both prototype reviews were included in the statistical analyses.

For response type, we performed individual Chi squared analyses to determine any statistical significance among stakeholder groups and prototype types across all interview questions. Bonferroni corrections were applied to the typical p value of 0.05 for the individual tests when groups (stakeholders or prototypes) confounded the results. The detailed results of the Chi squared analysis can be found in Appendix 1.

For the results of the CAT, we performed ANOVAs to evaluate if significant differences existed among the categories (stakeholder groups and prototype types). The ANOVAs were followed by t tests with the same Bonferroni corrections. We also performed ANOVAs and t tests with the same Bonferroni corrections to determine if word count revealed any significant differences among stakeholder groups and prototype types.

4 Results

4.1 How does prototype format impact stakeholder feedback?

Considering response type based on prototype format for all interview questions, we observed significant differences among prototypes: Tangible prototypes (mockup and 3D-printed) provided more category A or B answers than virtual prototypes (sketch and CAD). Tangible prototypes resulted in 77% of category A or B answers, while virtual prototypes resulted in 65% of category A or B answers. Virtual prototypes provided more category C or D answers than tangible prototypes. We also observed that low-fidelity prototypes resulted in fewer category A or B answers than high-fidelity prototypes (sketch or CAD and mockup or 3D printed), but without statistical significance. A visual depiction of the results can be found in Fig. 2.

A Chi squared analysis of the combined answer categories (A or B, C or D) revealed statistical significance of the findings among individual prototypes ($p = 0.0011$) for all questions (Table 2).

When combining lower and higher fidelity prototypes, tangible prototypes resulted in significantly more category A or B answers ($p = 0.0001$) than virtual prototypes for all questions (Table 3).

We observed similar results to the response type analysis with the CAT metric for all questions, but CAT showed no statistical significance. However, we observed larger standard deviations across the two groups with this analysis. A final analysis with word count once again revealed similar

Fig. 2 Results for response category by prototype type for **a** individual answer categories and **b** combined answer categories
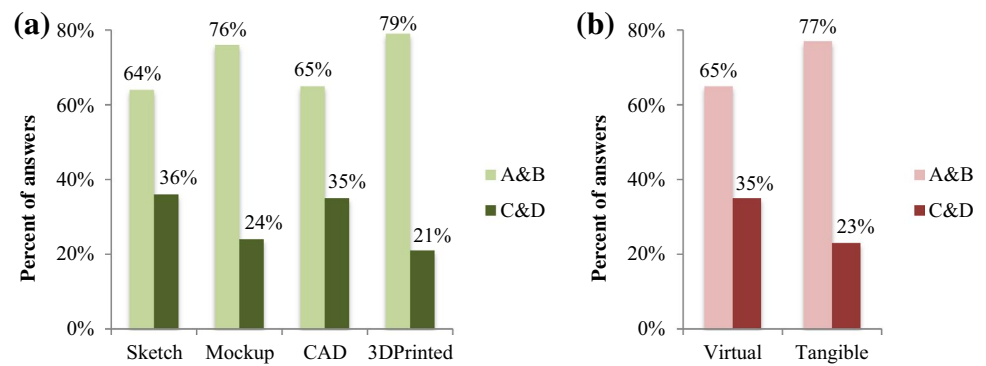

Table 2 Chi squared analysis of prototypes

	Observed values among participants					Expected values among participants				
Prototypes	SK	MO	CAD	3D	Total	SK	MO	CAD	3D	Total
A & B	130	151	129	161	571	144	141	140	146	571
C & D	73	47	68	44	232	59	57	57	59	232
Total	203	198	197	205	803	203	198	197	205	803
										$p = 0.0011$

Table 3 Chi squared analysis of virtual and tangible prototypes

	Observed values among participants				Expected values among participants			
Prototypes	Virtual	Tangible	Total	Prototypes	Virtual	Tangible	Total
A & B	259	312	571	A & B	284	287	571
C & D	141	91	232	C & D	116	116	232
Total	400	403	803	Total	400	403	803
							$p = 0.0001$

results, also with no statistical significance, and even larger standard deviations.

In the following paragraphs, we provide quotes that illustrate the range of answers received in response to the different prototype types.

For category A or B answers, Nurse 16 who commented on a 3D-printed prototype, suggested that the device should be made of a non-rigid material: *"I would love it if it had been more flexible than this."* The nurse also voiced concerns about the device being disposable, and that a way to prevent repeated use, and the inherent risk of cross-contamination, should be considered by the designers: *"Sometimes in our setting, we use it for different patients, so I am thinking that if it will be in such a way that we can use it once for a patient, and that is it. We don't use it for another patient to put the patient at a risk of infection."*

Student 17, whose response was also categorized as an A answer, suggested that thin patients might not have enough tissue to fill the cavity of the device: *"Maybe when someone is very slim, you may not have this space filled. You may not be able to take a great amount of tissue, just take the upper part of the skin and that will make your insertion partial."* The

student expanded on this concern and suggested that the device should be made available in different sizes to fit a variety of patients: *"Maybe there should be varying sizes for different weight measurement… so maybe from 60-70 kg you have this. Then from 50-60, you have a smaller one, so that everyone has his or her appropriate size."* The student also proposed a material change that would alter the procedure and give the provider more control: *"Maybe it should be a bit more transparent, so you can see through what you doing, because this, you can't see what you doing. It should be more transparent."*

Likewise, Doctor 28 commented on the material of the device but was concerned about potential allergic skin reactions: *"I think it should be gentle on the skin. If it's an inert material, then that's home and prime, then you don't expect people's skin to react to the material."* The same doctor suggested durable materials to prevent breakage of the device: *"It shouldn't be too much of a brittle material that can easily give way. Because if it's in a CHPS [Community-based Health Planning and Services] zone, then you expect this thing to be used, carried from place to place and all of that. If it is something that can easily give way or break, then it may be some minus to it."*

Fig. 3 Results for response category by stakeholder group for **a** individual answer categories and **b** combined answer categories
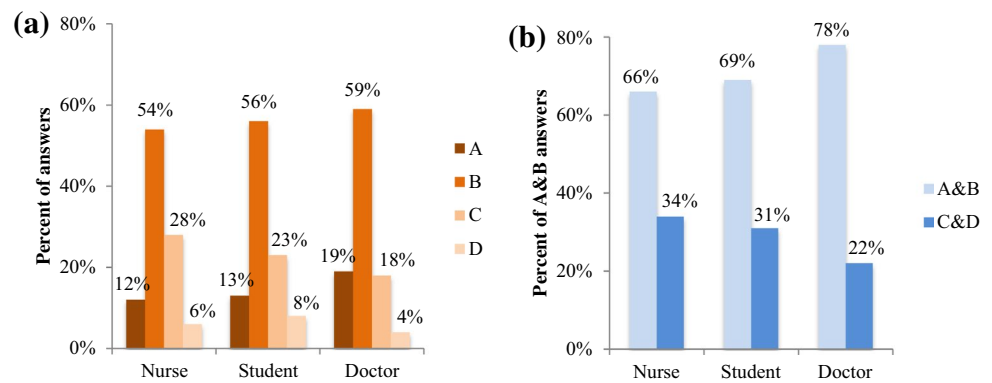

Table 4 Chi squared analysis results for stakeholder groups

Observed values among participants					Expected values among participants				
Groups	Nurse	Student	Doctor	Total	Groups	Nurse	Student	Doctor	Total
A & B	210	124	237	571	A & B	227	128	216	571
C & D	109	56	67	232	C & D	92	52	88	232
Total	319	180	304	803	Total	319	180	304	803
									$p = 0.0029$

In contrast, the virtual prototypes frequently led to confusion and conflicting information that resulted in C or D answers from the participants. For example, when asked if the concept would work as intended, Student 33, who saw a sketch, first expressed trust in the design: *"Yeah it will. Just because of the concept behind it, I think it will,"* but later added that it would have been beneficial to see the actual device, undermining the validity of the previous statement: *"I would have loved to see the device itself, but it's nice. It's really nice. I like the idea. I like everything."*

When participant 25, a nurse who saw a sketch, was asked to comment on the appearance of the device, the input given was: *"It would have been better if I have seen it in reality, this drawing; I can't say much about it."*

Nevertheless, not having enough information did not stop some participants from expressing their opinions. Participant 19, a student, mentioned that the CAD model was *"huge,"* even though the model shown on the screen afforded no size reference: *"I think you have to be very careful, when it goes under the skin and I feel it's huge so... I still prefer this [free hand insertion method]."* However, the student later concluded that additional information would have been required to recommend changes to the device: *"I don't really know the parts and everything well, so I can't make a comment on that."*

4.2 How does group membership impact stakeholder feedback?

Considering response type based on stakeholder group for all interview question responses, we found that doctors

provided the highest number of category A or B answers, followed by students, then nurses. The response type analysis showed that 78% of the responses provided by doctors were categorized as A or B answers, followed by 69% of A or B answers for students, and 66% of A or B answers for nurses. These findings were opposite for category C or D answers. Here, nurses provided the highest number of category C or D answers, followed by students, and then doctors. The results are shown in Fig. 3.

These findings were significant ($p = 0.0029$) for the collapsed answer categories (A or B, C or D) among the three stakeholder groups (Table 4).

Analysis of usefulness by CAT ratings revealed similar results across all interview questions, but none of them were statistically significant. We also observed larger standard deviations for all stakeholder groups with this technique. While word count did not show any statistical significance, it revealed that doctors had the highest average word count. Nurses had higher word counts than students with this analysis, but with a much larger standard deviation.

Next we provide quotes that illustrate the range of answers received by stakeholder groups.

For category A or B, for example, Doctor 38 thought that the presented concept was appropriate, but voiced concerns about patients' perceptions regarding the size of medical devices: *"Generally patients are scared when they see big things... So if things are portable, so just... like this. This is small, so I think it's ok."*

Doctor 11, who saw the cardboard mockup, expressed concerns that the tissue might actually not move into the cavity as intended and suggested that the designers might

investigate how the skin behaves during the procedure: *"You actually have to apply some, a little bit of counter traction on the skin so that the skin is actually not creased or folded. So how sure are we that we don't get that?"*

Doctor 36 stressed the importance of safety and the fact that the device should be disposable to avoid cross-contamination among patients: *"If it's going to be disposable, then I guess it will be safe to use. Because it's… invasive with the device… a little blood spillage, unless you plan on disinfecting and sterilizing after each patient."*

Students and nurses also provided category A or B answers, but fewer than doctors. For example, Student 20, who reviewed the cardboard mockup, was concerned about the size of the device, but focused more on how the size might influence the procedure: *"I kind of think it's too big… it's going to be like bulky in between the person's arm, so if you could have something smaller than this, but with the same concept, I think it's great."*

Student 23 compared the appearance of the 3D-printed device to an everyday object and posited that it would put a patient at ease: *"It looks… seriously, it doesn't look like something that is used to insert an implant; it's rather like an opener. Yeah, a bottle opener or something… It does not look like it's going to be used in the hospital."*

Similarly, Nurse 26 associated the 3D-printed prototype with a writing utensil and concluded that it would be non-threatening: *"It's just like a pen case. It looks like a pen case, so there is no problem with this."*

Nurse 14 thought about how the device would integrate into the implant procedure and stressed the need for training of the service provider to put the patient at ease: *"We should [have] adequate training on how the device would be used. Training of the facilitators, and then let the client know how it would be used on them. They would buy into the idea."*

Nurses provided the highest percentage of category C or D answers across all question and prototype types. For example, when asked if the concept would work while reviewing the sketch, Nurse 1 responded: *"I think it will be nice, but because I have not seen it, it will be very difficult for me to say. Maybe when it comes out and we're using it…"* When Nurse 5, who reviewed the cardboard mockup, was asked if the concept would work, the answer was referred back to the design team's earlier description of the device's intended use: *"You said it can do that."*

Members of other stakeholder groups also provided category C or D answers. For example, Doctor 12, who reviewed the cardboard mockup, offered the following insight: *"I don't know, I can't tell, but I hope it works."* The participant later added: *"Ok, well, I really can't tell, honestly, because I have no idea. I don't know, I really can't tell."* Similarly, Student 10, who reviewed a 3D printed model, was not prepared to assess the feasibility of the concept: *"Will it work? I have no idea!"*

4.3 How does question type impact stakeholder feedback?

To investigate if the feedback differed among individual questions, we examined the results for all stakeholder groups (doctors, nurses, and students) and prototype types (virtual and tangible) for questions 3 (Do you think this concept would work?), 6 (How do you think patients would feel about this device being used during the implant procedure?) and 9 (What would you suggest changing about this device?).

4.3.1 Questions and stakeholders

For question 3, 73% of the responses were categorized as A or B answers, for question 6, 89% of the responses were categorized as A or B answers, and for question 9, 54% of the responses were categorized as A or B answers. We found statistically significant differences based on the outcomes of response type analysis ($p = 0.0000$) between question 6 (most category A or B answers) and question 9 (least category A or B answers) for all stakeholder groups.

For individual stakeholder groups, we found that nurses provided significantly more category A or B answers ($p = 0.0000$) for question 6 (97%) than for questions 3 and 9 (64% and 47%), respectively. Nurses also provided more category A or B answers for question 6 than the other groups. Doctors provided the highest percentages of category A or B answers across the three questions compared to the other stakeholder groups, and offered significantly more category A or B answers ($p = 0.0011$) for both questions 3 and 6 (both 88%) than for question 9 (56%). An analysis of students' responses showed no significant differences among the three questions, and neither CAT nor word count analyses resulted in any significant findings for questions or stakeholders. The statistically significant differences are shown in Fig. 4.

All stakeholders were significantly more likely to provide category A or B answers for question 6: "How do you think patients would feel about this device being used during the implant procedure?" than for question 9. For example, Nurse 8 thought that not seeing the needle during the procedure would be an asset to the patient: *"Once she doesn't see the needle directly, it will rather make her more relaxed. You explain the procedure to her and how this thing is going to work on her, that will relax her…"* Nurse 16 mentioned concerns about the rigidity of the device: *"As I said, it's a bit hard so the patient will feel a bit uncomfortable."* Nurse 24 appreciated what the device would do for the medical provider, but was concerned about the patient: *"For us doing the insertion, it will be easy, but thinking about the patient, I think it will be a bit uncomfortable."* Nurse 25 suggested that after a brief explanation, the patients would be fine with the procedure: *"Just like you check their BP, you wrap a*

Fig. 4 Results for category A or B answers for questions 3, 6 and 9 for **a** all stakeholders, **b** nurses, **c** doctors
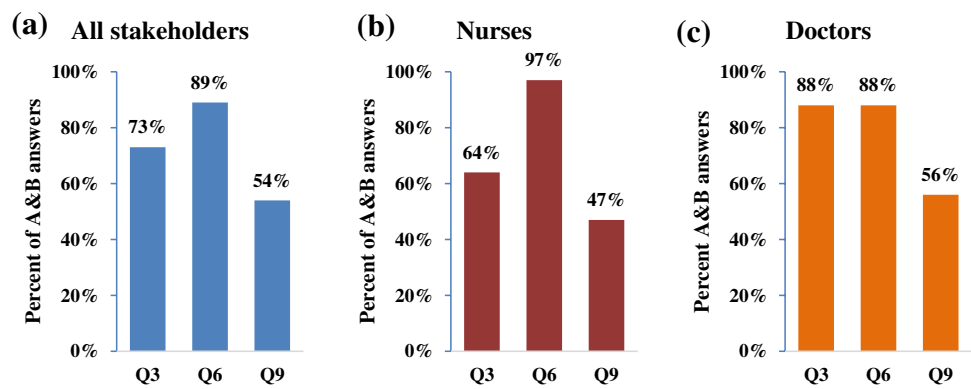
cuff around their arm, they will be comfortable once you've explained the procedure to them." Nurse 43 voiced concerns about patient comfort and asked if the designers had considered this already: *"I don't know whether there would be some discomfort when the tissues are going there, [do] you anticipate that?"*

In contrast, question 9 (What would you suggest changing about this device?) resulted in the lowest number of category A or B answers across all stakeholder groups. Category C or D answers for question 9 included examples from Doctor 9 who had only this to say: "*I think it's fine,*" or Doctor 11, who was uncomfortable commenting on technical details: "*Wow you're talking to me [about] engineering…*" Doctor 6 who saw a sketch of the device expressed a need more information to make any recommendations: "*I am wondering how it's going to lift the skin under the cavity, so until I see it, I can't comment.*"

4.3.2 Questions and prototypes

We found statistically significant differences in response type ($p = 0.0001$) between virtual and tangible prototypes for question 9, 78% of the responses to tangible prototypes were categorized as A or B answers, while only 24% of the responses to virtual prototypes were categorized as A or B answers (Fig. 5). We also observed statistically significant differences when assessing for the usefulness of the feedback using CAT between virtual and tangible prototypes ($p = 0.0002$), while the word count analysis showed similar trends, although not statistically significant ($p = 0.3507$) combined with greater variability.

Stakeholders who saw tangible prototypes when answering question 9 ("What would you suggest changing about this device?") responded with more category A or B answers, higher ratings of usefulness by experts, and lengthier answers than those who reviewed virtual prototypes. Examples of category A or B answers included quotes like the following by Nurse 4 who saw the cardboard mockup: "*I like it, but I think the size is a little big. Yeah, if it can be a little [more] portable, that will be fine.*" Doctor 10 was
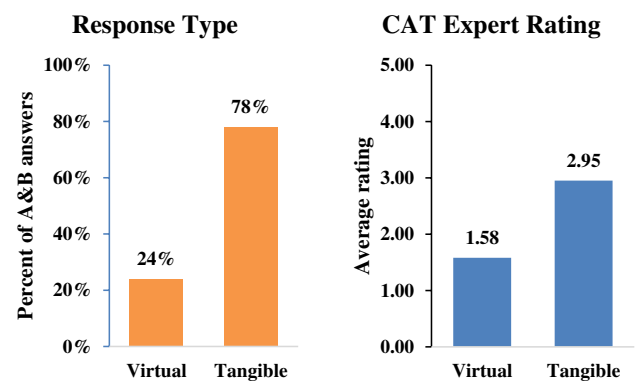

Fig. 5 Results for question 9 for prototype types and **a** response type and **b** expert rating of usefulness

even more specific about how the size of the device might be critical to a diverse patient population: "*Maybe there should be some form of adjustment to take care of thin people because it may accommodate more than the skin in the subcutaneous tissue. It may take some amount of muscle, so maybe some modifications should be made for thin people.*" Nurse 16 who commented on a 3D-printed prototype added concerns about the disposable nature of the device: "*Yes, it's enough to know it disposable, but unfortunately, for our setting, sometimes, due to inadequate consumables and all, we turn to reuse it. So if it can be done in such a way that you can't reuse it…*"

In contrast, virtual prototypes resulted in fewer category A or B answers, lower ratings of usefulness by experts, and shorter answers for question 9. Examples for category C or D included answers like "*I can't say much about it until I start using it or something,*" by Nurse 1 who did not think that the proposed concept was realistic enough on which to comment. The nurse later added: "*I think for now no because this is just on paper.*" This was echoed by Doctor 6 who questioned if the device would actually work and wanted to see the actual device perform: "*Ok I haven't seen it actually been done before, so I am wondering how it's going to lift the skin under the cavity. So until I see it, I can't comment.*"

5 Discussion

5.1 Influence of prototype format on stakeholder feedback

In our examination of how prototype format influenced the feedback stakeholders provided, we found that tangible prototypes provided more category A or B answers, higher ratings of usefulness by experts, and longer responses than virtual prototypes across all stakeholder groups and questions. These findings echo recommendations that call for tangible prototypes to be used for collecting stakeholder input on products and devices (De Beer et al. 2009; Kelley 2001; Otto and Wood 2000; Schrage 1999), but other researchers have found that, depending on the task, virtual prototypes can be equally beneficial during product development, as long as designers are aware of their benefits and limitations (Rudd et al. 1996; Ulrich and Eppinger 2015; Walker et al. 2016).

Using different types of prototypes might also affect variations of answers. The variations in the types of responses, ratings of usefulness by experts, and word count in our study were smaller for the cardboard mockup and 3D-printed model than for the than the virtual sketch and CAD model. This greater variation within the virtual prototyping categories might suggest greater diversity in participants' abilities to respond to these prototypes and likely makes the process of synthesizing input more difficult for designers. Several participants stated they were not able to obtain enough information from the virtual prototypes, which in some cases resulted in unjustified or non-useful feedback (category C or D responses). Limited experience with, and exposure to, design processes, medical device development, or the review and critique of virtual prototypes might have contributed to the perceived need for additional information. As a result, participants might have felt overwhelmed by the task, which could have led to frustration and emotional responses rather than analytical processing of information (Frijda 1989; Scherer 2003; Winston and Cupchik 1992).

Analysis of low versus high fidelity showed no statistical significance, but within each category, the more refined prototypes (CAD model for virtual, and 3D-printed for tangible prototypes) were related to more category A or B answers, higher ratings of usefulness by experts, and lengthier responses. These results align with Brandt's (2007) findings that greater levels of detail within a prototyping category led to smaller variations and more focused conversations between stakeholders and designers. Similarly, studies evaluating stakeholder feedback on new product concepts have found that the highest level of prototype quality correlated with higher ratings by the stakeholders regardless of the criteria, e.g., functionality or creativity of an idea (Häggman et al. 2015; Kudrowitz et al. 2012; Sauer and Sonderegger 2009).

Our findings align with Hannah's et al. (2012) conclusion that higher fidelity prototypes lead to more desirable results (more confidence), but also contradict the findings of Viswanathan and Linsey's (2011) study that investigated the effects of prototypes on designers' creative process. The researchers found that low-fidelity prototypes invited more contribution to the design, and that higher fidelity prototypes were sometimes seen as "too complete" to warrant more input by participants. However, the researchers also found that the physical models that required more time and effort to create led to more design fixation by the designers than the physical models that required less building effort. Thus, there are trade-offs in choice of prototype fidelity.

On the other hand, some studies found little difference between low- and high-fidelity prototypes, but these focused on non-tangible, two-dimensional products like user interfaces and websites only (Lim et al. 2006; Walker et al. 2016). In our study, the lesser refined prototype categories led to larger variations and more confusion in the stakeholder feedback. For example, one nurse asked which one of the views of the sketch to comment on, not realizing that all views of the sketch depicted the same product.

Similarly, even though participants were made aware they were looking at a prototype, some still voiced concerns about properties specific to a particular prototype, such as the fact that the blood of a patient might stain the cardboard material used for the mockup. This insight might indicate that some participants were not able to look beyond the prototype format and its inherent limitations when assessing lower-fidelity representations.

5.2 Influence of stakeholder group membership on stakeholder feedback

In our examination of group membership, we found differences among stakeholder groups. The feedback doctors provided included the most category A or B answers, the highest ratings of usefulness by experts, and the longest responses. The feedback students provided included more category A or B answers and higher ratings of usefulness by experts than nurses, but nurses provided longer responses than students. There might be several reasons for these differences. First, the introduction of "design thinking" to a clinical environment is a fairly recent development (Kalaichandran 2017; Roberts et al. 2016) that is often limited to physicians and medical students, and frequently exclude nurses (Rosen and Ku 2016). Therefore, many healthcare

professionals, including those in this study, likely had limited experience with the design and development of medical devices.

Further, nurses in African countries have traditionally been trained with a focus on physician order execution and task completion (Marks 1994). The mission-style training approach, adopted by many sub-Saharan African countries from the British colonial system (Edwards 1957), might have introduced a social desirability bias, where nurses are not necessarily accustomed to providing critique and voicing their opinions. More recently, efforts have been made to redefine nursing practices from a more task-oriented approach to one of caring for and caring about patients (Savage 1995), but without necessarily challenging the hierarchal structure within the healthcare system. These factors might explain why nurses performed better on the patient-centric question than on the design-specific question.

In addition, the training of Ghanaian medical doctors often includes fellowships in the United Kingdom and the United States (Klufio et al. 2003b), introducing them to cultures of critique. This international experience might explain why doctors provided the most category A or B answers to questions addressing the design of the medical device used in this study.

Nurses frequently provided less critical observations and instead compared the presented device to everyday objects. Leder defines these "looks like" or "feels like" responses as prototypicality, a cognitive way for a reviewer to associate a new object with another and more familiar object. The association of information content with their own situation and emotional state can lead a reviewer to be content with a simple recognition. Parsons (1989) posited that "a naïve perceiver might be satisfied with the recognition of the train station in Monet's La Gare Saint-Lazare because 'he likes trains because they remind him of a journey.'" This observation is not limited to art, as differing levels of expertise influence how new concepts are perceived in other domains as well. While experts tend to abstract principles when solving a problem, novices often focus on literal features (Chi et al. 1981). In our study, we saw indications that an emotional evaluation, association with a familiar product, and focus on features might have limited cognitive inquiry by a reviewer. Several participants compared the device concept to everyday objects like a bottle opener or pen case and concluded that since these objects are safe, non-threatening devices, a medical device concept that looks or feels similar must therefore share similar qualities.

Analogous to studies that showed that stakeholder input can be contradicting, making it difficult for designers to synthesize information (Mohedas et al. 2014; Scott 2008), we, too, found evidence of sometimes conflicting stakeholder input. Even when participants' feedback consisted of category A or B answers, high ratings of usefulness by experts,

and lengthy responses, their input was sometimes incompatible. For example, one stakeholder asked for the device to be transparent so that practitioners can see what they are doing, while another stakeholder appreciated the fact that an opaque device would hide the needle from the patient during the implant procedure. Both participants provided potentially useful input, yet suggested opposing product qualities (transparent or opaque). The fact that these arguments could both be valid underscores the fact that designers cannot simply take stakeholder input at face value. Instead, designers should expect contradicting feedback, especially early in the design process, when seeking a comprehensive understanding of the requirements. Here, prototypes provide a chance to interact with and evaluate, proposed solutions and can be used to help uncover "unknown unknowns" (Jensen et al. 2017). Designers need to embrace these findings and use them to inform prototyping strategies and design decisions.

Our results align with studies that have shown that physical prototypes that were more widely understood by and accessible to participants positively affected their emotions, and prompted participants to respond with a high degree of confidence (De Beer et al. 2009; Häggman et al. 2015; Sauer and Sonderegger 2009). Our results also reflect findings by Björklund (2013) and Simon (1973) that the mental representations of design experts (how design problems are transformed or structured into mental representations) were broader, more detailed, and more focused toward problem solving. Similar to studies that have shown the benefits of exposure to multinational and tangential experiences during training of medical doctors (Klufio et al. 2003a) and biomedical engineers (Sienko et al. 2014), we saw indications that stakeholders who might have been exposed to innovation and critique in addition to training in medical practice and patient care might have been better prepared to provide feedback on the design and development of new products, here medical devices.

5.3 Influence of question type on stakeholder feedback

In our examination of how question type influenced stakeholder feedback, we found that the feedback stakeholders provided depended on the question type as well as on stakeholder characteristics. Related to this finding, several studies have shown that the questions designers ask are contingent on the phase of their design process (Christie et al. 2012; Menold et al. 2017). Combined, these findings suggest that designers need to consider the questions they ask as well as whom they are asking in all stages of the design process. Not all stakeholders seem to be equally able to provide input to all questions, and designers may need to rephrase a question, or situate it in a different context, depending on whom they are engaging with. Through the question type, designers

can, and need to, enable stakeholders to relate to the design problem and feel comfortable enough to respond.

For example, we found that question 6 "How do you think patients would feel about this device being used during the implant procedure?" resulted in the most category A or B answers across all stakeholder groups. Particularly, nurses provided the most category A or B answers for this patient-centric question that was situated within this stakeholder group's knowledge domain of treating and caring for patients. By no longer asking stakeholders to critique the device directly, this question took them off of the "hot seat" and allowed them to assume the role of caregiver, associating with their patients. This new perspective enabled stakeholders to talk more freely and comment on the experiences both patients and caregivers might have when using the device during the contraceptive implant insertion procedure. Similar to Leder's example earlier (Leder et al. 2004), question 6 may have enabled stakeholders to pick up on the nuances only an expert could. Familiarity and experience may have allowed stakeholders to move through several stages of information processing for this question, evaluating the device on a much deeper level than before and thinking through the procedure from the patient and caregiver perspectives. For this particular question, stakeholders had become experts, and the highest number of category A or B answers we recorded for this question reflected this level of expertise.

We also found that question 9, "What would you suggest changing about this device?" resulted in the lowest number of category A or B answers for all stakeholders and all prototypes. Two reasons come to mind for why this question might have fared so poorly: first, it was asked last and stakeholders might have exhausted their input on the previous eight questions and simply became tired of repeating themselves. Second, this question asked stakeholders directly what they would change about the design. Since some stakeholders had little or no experience with medical device design, this might have caused them to feel uncomfortable and/or overwhelmed. As the findings from other questions indicated, stakeholders were more likely to provide input when they were asked about specific details rather than to give general input. This is another important finding, since novice designers tend to ask more general questions. Our findings suggest that not all stakeholders are equally prepared to do this; designers need to consider their stakeholders and carefully select and frame questions that enable stakeholders to provide feedback on the proposed design.

5.4 Influence of analytical methods on stakeholder feedback

When comparing the findings of the different analytical methods employed during this study, all three techniques identified similar results: Tangible prototypes led to more category A or B answers, higher ratings of usefulness by experts, and longer responses than virtual prototypes. All techniques also revealed that doctors provided more category A or B answers, higher ratings of usefulness by experts, and longer responses than students, who provided more category A or B answers and higher ratings of usefulness by experts than nurses. Nurses provided longer responses than students, but with large standard deviations.

In addition, the categorization of responses by type identified the highest number of statistically significant differences among prototypes and stakeholder groups. This finding is not surprising since this method relied on carefully developed codes to analyze the data. The codes provided specific criteria for the analysis and therefore revealed the most differences among the input categories. The iterative development of codes, in addition to several rounds of coding, were time-intensive tasks. Despite these efforts, the results suggest that this method led to the most insightful, significant, and reliable findings.

The CAT relied on individual rater established criteria (Amabile 1983), and we observed similar results for CAT and the categorization of responses by type. However, the larger standard deviations and less significant results of this analytical method make the findings less reliable. The small number of expert raters who participated in the analysis were likely a factor; a larger number of experts might improve the results.

Word count considered only the number of spoken words and showed less pronounced, and sometimes even conflicting results, with no statistical significance. This analytical method also resulted in the largest standard deviations. In one extreme case, when examining the influence of prototype type on question 9, the standard deviation of 54.75 words even exceeded the average count of 37.79 words for virtual prototypes, large enough to question the validity of this result.

We also found that for all stakeholders, nurses had the highest word count for question 9, but fewer than expected category A or B answers and the lowest average ratings of usefulness by experts for this question. In contrast, doctors had the lowest word count, but the most category A or B answers and second-highest ratings of usefulness by experts for the same question. In these cases, the word count results were in opposition to the findings of the other techniques. These observations contrast with Blumenstock's study (2008) that found a positive correlation among the length and the quality of articles published on Wikipedia. However, such articles are peer-reviewed and nominated, a process that is absent when collecting stakeholder input. In an earlier study, Weber (1983) concluded that content analysis may be the preferred way to generate

quantitative indicators, but our findings highlight the need to consider Morgan's argument (Morgan 1993) that it is critical to determine 'what' to code for in content analysis and that a range of techniques for analyzing qualitative data might be preferable. In summary, without developing codes for content analysis, how much a person says seems not to be a good indicator of quality of content, making word count the least reliable analytical technique used in this study.

6 Implications

The findings of this study are important for design practitioners planning to use prototypes, and in particular for projects designed at distances where access to stakeholders can be challenging. Specifically in global health design, where geographic distances and time-zone differences can limit and restrict conversations, interactions with stakeholders need to be carefully planned and executed. Here, a successful prototyping strategy is even more critical and should encompass the following elements:

First, designers must select appropriate prototype types. For example, when looking for procedural or in situ feedback, simple prototypes like sketches might not enable stakeholders to address issues that a functional prototype might reveal (Sauer et al. 2008). The commonly accepted prototyping best practices (quick and simple) used in the United States are not necessarily universally transferable and need to be adjusted to the unique context and background of the design project. It is not enough to consider that different stages in the design process call for different types of prototypes (Atman et al. 2007)—designers also need to select the most appropriate prototype types that allow stakeholders to best respond and provide useful input.

Second, designers need to recognize that not all stakeholders are equally prepared to respond to all prototypes—a sketch might work well for an engineer but not resonate with a social worker. When stakeholders have limited domain experience, or feel inadequately equipped to evaluate a new concept, they might not be able to move through the stages of information processing necessary for a comprehensive evaluation. Instead, they might feel overwhelmed and express an emotional response that can be misleading and even harmful, especially when designers do not have experience interpreting the feedback they receive. Designers need to recognize who their stakeholders are and select the types of prototypes that best support these individuals.

Third, the questions designers ask when using prototypes need to be carefully selected to enable dialog between stakeholders and designers. Designers need to consider the context of the question and develop questions that enable stakeholders to more effectively draw upon their own expertise. A stakeholder who is not well prepared to provide technical input might be an excellent candidate to offer insight into the social or psychological impact a new design concept might have on a community. Having experience with the use of a device is not the same as having experience with the design and development of a device. It is up to the designer to ask the "right" questions and take advantage of individual stakeholders' expertise.

Fourth, the findings can inform design pedagogy and curriculum development, since the application of the results is not limited to medical device design. The findings can be transferred to other contexts where designers use prototypes to gather stakeholder input. In this study, prototype type, stakeholder group, and question type all influenced stakeholder feedback. Educators can capitalize on this insight and guide students to carefully consider the unique circumstances of their design project. In particular, they can encourage students to develop prototyping strategies that optimize their interactions with stakeholders when looking for feedback on new designs.

7 Limitations and future work

There were several limitations to this study that could be addressed in future work. Only a subset of the answers participants provided was analyzed in detail, and the number of participants could be expanded. No information on participants' prior design experience was collected and some were likely inexperienced with providing design feedback. The study was limited to one unique setting and stakeholders with specific cultural, geographical and professional backgrounds. Future studies might explore the extent to which the findings can be transferred to different stakeholder groups, prototypes of products in other arenas, as well as systems and processes. The questions used during the interviews represent typical questions that designers might pose to stakeholders and were not explicitly designed or selected to specifically study the effects of question type on response type. We did not investigate how the individual features of a prototype or the order in which the prototypes were presented influenced the usefulness of the feedback that was elicited from stakeholders. A male researcher who was not a native of Ghana conducted all interviews and, although English is considered an official language, it was likely not the first language for some participants. These factors might have influenced the participants' responses, specifically the richness and explicitness of their feedback.

8 Conclusion

We found that tangible prototypes resulted in more category A or B answers and higher ratings of usefulness by experts than virtual prototypes, regardless of their fidelity. Designers need to be aware of this tendency and should proactively develop context-specific strategies that complement the "quick and simple" approach to prototyping, since prototype type matters. We also found that doctors provided the most category A or B answers and the highest ratings of usefulness by experts. However, nurses responded with more A or B answers and higher ratings of usefulness by experts for a particular question focused on how patients might feel. Questions positioned within a stakeholder's professional experience resulted in more category A or B answers and higher ratings of usefulness by experts than general and technical questions. It is therefore important for designers to carefully consider what questions they ask, and to whom they are asking them. Specific rather than general, or summative questions that are situated in a stakeholder's domain have the potential to empower stakeholders to comprehensively evaluate the prototypes with which they are presented.

Appendix 1

Interview questions

1. In general, what do you like about this concept?
2. What do you dislike about this concept?
3. Do you think this concept will work?
4. Do you think the device would be easy or hard to use?
5. Do you think the device would be safe to use?
6. How do you think patients would feel about this device being used during the implant insertion procedure?
7. What do you think about what the device looks like?
8. Do you think this device would be appropriate for CHPS workers (Community-based Health Planning and Services) to use in rural settings?
9. Based on what you know about this concept, what would you change about this device?

Appendix 2

Stakeholders

Nurse

	Observed Values				Expected Values			
	3	6	9	Total	3	6	9	Total
Round 1 — A&B	13	17	9	39	13	13	13	39
Round 1 — C&D	5	0	9	14	5	4	5	14
Round 1 — Total	18	17	18	53	18	17	18	53
p= 0,0036								
Round 2 — A&B	10	17	8	35	12	12	12	35
Round 2 — C&D	8	1	10	19	6	6	6	19
Round 2 — Total	18	18	18	54	18	18	18	54
p= 0,0043								
Rounds 1&2 — A&B	23	34	17	74	25	24	25	74
Rounds 1&2 — C&D	13	11	9	33	11	11	11	33
Rounds 1&2 — Total	36	35	36	107	36	35	36	107
p= 0,0000								

Student

	Observed Values				Expected Values			
	3	6	9	Total	3	6	9	Total
Round 1 — A&B	6	8	4	18	6	6	6	18
Round 1 — C&D	4	2	6	12	4	4	4	12
Round 1 — Total	10	10	10	30	10	10	10	30
p= 0,1889								
Round 2 — A&B	7	7	6	20	7	7	7	20
Round 2 — C&D	3	3	4	10	3	3	3	10
Round 2 — Total	10	10	10	30	10	10	10	30
p= 0,8607								
Rounds 1&2 — A&B	13	15	10	38	13	13	13	38
Rounds 1&2 — C&D	7	5	10	22	7	7	7	22
Rounds 1&2 — Total	20	20	20	60	20	20	20	60
p= 0,2557								

Doctor

	Observed Values				Expected Values			
	3	6	9	Total	3	6	9	Total
Round 1 — A&B	14	15	11	40	13	13	13	40
Round 1 — C&D	3	2	6	11	4	4	4	11
Round 1 — Total	17	17	17	51	17	17	17	51
p= 0,2216								
Round 2 — A&B	16	15	8	39	13	13	13	39
Round 2 — C&D	1	2	9	12	4	4	4	12
Round 2 — Total	17	17	17	51	17	17	17	51
p= 0,0020								
Rounds 1&2 — A&B	30	30	19	79	26	26	26	79
Rounds 1&2 — C&D	4	4	15	23	8	8	8	23
Rounds 1&2 — Total	34	34	34	102	34	34	34	102
p= 0,0011								

NSD

	Observed Values				Expected Values			
	3	6	9	Total	3	6	9	Total
Round 1 — A&B	33	40	24	97	33	32	33	97
Round 1 — C&D	12	4	21	37	12	12	12	37
Round 1 — Total	45	44	45	134	45	44	45	134
p= 0,0004								
Round 2 — A&B	33	39	22	94	31	31	31	94
Round 2 — C&D	12	6	23	41	14	14	14	41
Round 2 — Total	45	45	45	135	45	45	45	135
p= 0,0004								
Rounds 1&2 — A&B	66	79	46	191	64	63	64	191
Rounds 1&2 — C&D	24	10	44	78	26	26	26	78
Rounds 1&2 — Total	90	89	90	269	90	89	90	269
p= 0,0000								

Groups

	Observed Values				Expected Values			
	NUR	STU	DOC	Total	NUR	STU	DOC	Total
Round 1 — A&B	111	65	123	299	119	67	113	299
Round 1 — C&D	49	25	29	103	41	23	39	103
Round 1 — Total	160	90	152	402	160	90	152	402
p= 0,0568								
Round 2 — A&B	99	59	114	272	108	61	103	272
Round 2 — C&D	60	31	38	129	51	29	49	129
Round 2 — Total	159	90	152	401	159	90	152	401
p= 0,0485								
Rounds 1&2 — A&B	210	124	237	571	227	128	216	571
Rounds 1&2 — C&D	109	56	67	232	92	52	88	232
Rounds 1&2 — Total	319	180	304	803	319	180	304	803
p= 0,0029								

Prototypes

Sketch

	Observed Values				Expected Values			
	3	6	9	Total	3	6	9	Total
Round 1 — A&B	10	11	2	23	8	7	8	23
Round 1 — C&D	3	1	11	15	5	5	5	15
Round 1 — Total	13	12	13	38	13	12	13	38
p= 0,0002								
Round 2 — A&B	5	10	3	18	6	6	6	18
Round 2 — C&D	5	0	7	12	4	4	4	12
Round 2 — Total	10	10	10	30	10	10	10	30
p= 0,0044								
Rounds 1&2 — A&B	15	21	5	41	14	13	14	41
Rounds 1&2 — C&D	8	1	18	27	9	9	9	27
Rounds 1&2 — Total	23	22	23	68	23	22	23	68
p= 0,0000								

Mockup

	Observed Values				Expected Values			
	3	6	9	Total	3	6	9	Total
Round 1 — A&B	5	8	7	20	7	7	7	20
Round 1 — C&D	5	2	3	10	3	3	3	10
Round 1 — Total	10	10	10	30	10	10	10	30
p= 0,3499								
Round 2 — A&B	11	9	12	32	11	11	11	32
Round 2 — C&D	1	3	0	4	1	1	1	4
Round 2 — Total	12	12	12	36	12	12	12	36
p= 0,1396								
Rounds 1&2 — A&B	16	17	19	52	17	17	17	52
Rounds 1&2 — C&D	6	5	3	14	5	5	5	14
Rounds 1&2 — Total	22	22	22	66	22	22	22	66
p= 0,5301								

CAD

	Observed Values				Expected Values			
	3	6	9	Total	3	6	9	Total
Round 1 — A&B	9	11	6	26	9	9	9	26
Round 1 — C&D	2	0	5	7	2	2	2	7
Round 1 — Total	11	11	11	33	11	11	11	33
p= 0,0319								
Round 2 — A&B	7	9	0	16	5	5	5	16
Round 2 — C&D	4	2	11	17	6	6	6	17
Round 2 — Total	11	11	11	33	11	11	11	33
p= 0,0003								
Rounds 1&2 — A&B	16	20	6	42	14	14	14	42
Rounds 1&2 — C&D	6	2	16	24	8	8	8	24
Rounds 1&2 — Total	22	22	22	66	22	22	22	66
p= 0,0000								

3d Printed

	Observed Values				Expected Values			
	3	6	9	Total	3	6	9	Total
Round 1 — A&B	9	10	9	28	9	9	9	28
Round 1 — C&D	2	1	2	5	2	2	2	5
Round 1 — Total	11	11	11	33	11	11	11	33
p= 0,7900								
Round 2 — A&B	10	11	7	28	9	9	9	28
Round 2 — C&D	2	1	5	8	3	3	3	8
Round 2 — Total	12	12	12	36	12	12	12	36
p= 0,1238								
Rounds 1&2 — A&B	19	21	16	56	19	19	19	56
Rounds 1&2 — C&D	4	2	7	13	4	4	4	13
Rounds 1&2 — Total	23	23	23	69	23	23	23	69
p= 0,1652								

Prototypes (SK / MO / CAD / 3D)

	Observed Values					Expected Values				
	SK	MO	CAD	3D	Total	SK	MO	CAD	3D	Total
Round 1 — A&B	79	63	77	80	299	86	67	74	73	299
Round 1 — C&D	36	27	22	18	103	29	23	25	25	103
Round 1 — Total	115	90	99	98	402	115	90	99	98	402
p= 0,1041										
Round 2 — A&B	51	88	52	81	272	60	73	66	73	272
Round 2 — C&D	37	20	46	26	129	28	35	32	34	129
Round 2 — Total	88	108	98	107	401	88	108	98	107	401
p= 0,0000										
Rounds 1&2 — A&B	130	151	129	161	571	144	141	140	146	571
Rounds 1&2 — C&D	73	47	68	44	232	59	57	57	59	232
Rounds 1&2 — Total	203	198	197	205	803	203	198	197	205	803
p= 0,0011										

Prototypes (Virt / Tang)

	Observed Values			Expected Values		
	Virt	Tang	Total	Virt	Tang	Total
Round 1 — A&B	156	143	299	159	140	299
Round 1 — C&D	58	45	103	55	48	103
Round 1 — Total	214	188	402	214	188	402
p= 0,4680						
Round 2 — A&B	103	169	272	126	146	272
Round 2 — C&D	83	46	129	60	69	129
Round 2 — Total	186	215	401	186	215	401
p= 0,0000						
Rounds 1&2 — A&B	259	312	571	284	287	571
Rounds 1&2 — C&D	141	91	232	116	116	232
Rounds 1&2 — Total	400	403	803	400	403	803
p= 0,0001						

Prototypes (Low / High)

	Observed Values			Expected Values		
	Low	High	Total	Low	High	Total
Round 1 — A&B	142	157	299	152	147	299
Round 1 — C&D	63	40	103	53	50	103
Round 1 — Total	205	197	402	205	197	402
p= 0,0167						
Round 2 — A&B	139	133	272	133	139	272
Round 2 — C&D	57	72	129	63	66	129
Round 2 — Total	196	205	401	196	205	401
p= 0,1955						
Rounds 1&2 — A&B	281	290	571	285	286	571
Rounds 1&2 — C&D	120	112	232	116	116	232
Rounds 1&2 — Total	401	402	803	401	402	803
p= 0,5187						

References

Amabile TM (1983) The social psychology of creativity: a componential conceptualization. J Pers Soc Psychol 45(2):357–376. https://doi.org/10.1037/0022-3514.45.2.357

Amabile TM, Conti R, Coon H, Lazenby J, Herron M (1996) Assessing the work environment for creativity. Acad Manag J 39(5):1154–1184

Arkes HR, Blumer C (1985) The psychology of sunk cost. Organ Behav Hum Decis Process 35(1):124–140. https://doi.org/10.1016/0749-5978(85)90049-4

Atman CJ, Adams RS, Cardella ME, Turns J, Mosborg S, Saleem J (2007) Engineering design processes: a comparison of students and expert practitioners. J Eng Educ 96(4):359–379. https://doi.org/10.1002/j.2168-9830.2007.tb00945.x

Baer J, McKool SS (2009) Assessing creativity using the consensual assessment technique. In: Handbook of research on assessment technologies, methods, and applications in higher education, pp 65–77. IGI Global

Baer J, Kaufman JC, Gentile CA (2004) Extension of the consensual assessment technique to nonparallel creative products. Creativity Research Journal 16(1):113–117. https://doi.org/10.1207/s15326934crj1601_11

Baxter M (1995) Product design. CRC Press, Cambridge

Björklund TA (2013) Initial mental representations of design problems: differences between experts and novices. Des Stud 34(2):135–160. https://doi.org/10.1016/j.destud.2012.08.005

Blumenstock, J. E. (2008). Size matters: word count as a measure of quality on wikipedia. In: Proceedings of the 17th international conference on world wide web. ACM, pp 1095–1096

Brandt E (2007) How tangible mock-ups support design collaboration. Knowl Technol Policy 20(3):179–192. https://doi.org/10.1007/s12130-007-9021-9

Camburn BA, Dunlap BU, Kuhr R, Viswanathan VK, Linsey JS, Jensen DD, Wood KL (2013) Methods for prototyping strategies in conceptual phases of design: framework and experimental assessment. p V005T06A03. https://doi.org/10.1115/DETC2013-13072

Campbell RI, De Beer DJ, Barnard LJ, Booysen GJ, Truscott M, Cain R, Hague R (2007) Design evolution through customer interaction with functional prototypes. J Eng Des 18(6):617–635. https://doi.org/10.1080/09544820601178507

Castillo LG, Diehl JC, Brezet JC (2012) Design considerations for base of the pyramid (BoP) projects. In: Proceedings of the nothern world mandate: culumus helsinki conference, pp 24–26

Chamorro-Koc M, Popovic V, Emmison M (2009) Human experience and product usability: principles to assist the design of user-product interactions. Appl Ergon 40(4):648–656. https://doi.org/10.1016/j.apergo.2008.05.004

Chi M, Feltovich PJ, Glaser R (1981) Categorization and representation of physics problems by experts and novices. Cognitive Sci 5(2):121–152

Christie EJ, Jensen DD, Buckley RT, Menefee DA, Ziegler KK, Wood KL, Crawford RH (2012) Prototyping strategies: literature review and identification of critical variables. In: American Society for Engineering Education. pp. 011220–01154

Cohen J (1977) Statistical power analysis for the behavioral sciences. Routledge, New York. https://doi.org/10.1016/C2013-0-10517-X

Creswell JW (2013) Research design: qualitative, quantitative, and mixed methods approaches, 4th edn. SAGE Publications Inc., Thousand Oaks

Crilly N, Moultrie J, Clarkson PJ (2004) Seeing things: consumer response to the visual domain in product design. Des Stud 25(6):547–577. https://doi.org/10.1016/j.destud.2004.03.001

Daly S, Bryan LA (2010) Model use choices of secondary teachers in nanoscale science and engineering education. J Nano Educ 2(1–2):76–90

Crismond DP, Adams RS (2012) The informed design teaching and learning matrix. J Eng Educ 101(4):738–797. https://doi.org/10.1002/j.2168-9830.2012.tb01127.x

De Beer DJ, Campbell RI, Truscott M, Barnard LJ, Booysen GJ (2009) Client-centred design evolution via functional prototyping. Int J Prod Dev 8(1):22–41. https://doi.org/10.1504/IJPD.2009.023747

Deininger M, Daly SR, Sienko KH, Lee JC (2017) Novice designers' use of prototypes in engineering design. Des Stud 51:25–65

Desmet P, Hekkert P (2007) Framework of product experience. Int J Des 1(1):57–66

Dieter G, Schmidt L (2012) Engineering design, 5th edn. McGraw-Hill Education, New York

Dorst K, Cross N (2001) Creativity in the design process: co-evolution of problem–solution. Des Stud 22(5):425–437. https://doi.org/10.1016/S0142-694X(01)00009-6

Edwards AL (1957) The social desirability variable in personality assessment and research. Dryden Press, Ft Worth, TX

Free MJ (2004) Achieving appropriate design and widespread use of health care technologies in the developing world. Int J Gynecol Obstet 85(S1):S3–S13. https://doi.org/10.1016/j.ijgo.2004.01.009

Frijda NH (1989) Aesthetic emotions and reality. Am Psychol 44(12):1546–1547. https://doi.org/10.1037/0003-066X.44.12.1546

Funk S, Miller MM, Mishell DR, Archer DF, Poindexter A, Schmidt J, Zampaglione E (2005) Safety and efficacy of Implanon™, a single-rod implantable contraceptive containing etonogestrel. Contraception 71(5):319–326. https://doi.org/10.1016/j.contraception.2004.11.007

Gerber E (2009) Prototyping: facing uncertainty through small wins. In: DS 58-9: Proceedings of ICED 09, the 17th international conference on engineering design, vol. 9, Human Behavior in Design, Palo Alto, CA, USA, p 24. 27.08.2009

Gibson JJ (1966) The senses considered as perceptual systems. Houghton Mifflin, Boston

Giere RN (2004) How models are used to represent reality. Philos Sci 71(5):742–752. https://doi.org/10.1086/425063

Glesne C, Peshkin A (1991) Becoming qualitative researchers: an introduction. Longman Pub Group, White Plains

Häggman A, Tsai G, Elsen C, Honda T, Yang MC (2015) Connections between the design tool, design attributes, and user preferences in early stage design. J Mech Des 137(7):071408

Hannah R, Joshi S, Summers JD (2012) A user study of interpretability of engineering design representations. J Eng Des 23(6):443–468. https://doi.org/10.1080/09544828.2011.615302

Hare J, Gill S, Loudon G, Lewis A (2013) The effect of physicality on low fidelity interactive prototyping for design practice. In: IFIP conference on human–computer interaction. Springer, pp 495–510

Hekkert P, Snelders D, Wieringen PC (2003) 'Most advanced, yet acceptable': typicality and novelty as joint predictors of aesthetic preference in industrial design. Br J Psychol 94(1):111–124

Hilton E, Linsey J, Goodman J (2015) Understanding the prototyping strategies of experienced designers. In: IEEE Frontiers in Education Conference (FIE), 2015. 32614 2015, pp 1–8. https://doi.org/10.1109/FIE.2015.7344060

Houde S, Hill C (1997) What do prototypes prototype? In: Helander MG, Landauer TK, Prabhu PV (eds) Handbook of human-computer interaction, 2nd edn. North-Holland, Amsterdam, pp 367–381

Howard TJ, Culley SJ, Dekoninck E (2008) Describing the creative design process by the integration of engineering design and cognitive psychology literature. Des Stud 29(2):160–180. https://doi.org/10.1016/j.destud.2008.01.001

Howitt P, Darzi A, Yang G-Z, Ashrafian H, Atun R, Barlow J, Conteh L (2012) Technologies for global health. The Lancet 380(9840):507–535

Jensen MB, Elverum CW, Steinert M (2017) Eliciting unknown unknowns with prototypes: introducing prototrials and prototrial-driven cultures. Des Stud 49:1–31. https://doi.org/10.1016/j.destud.2016.12.002

Kalaichandran A (2017) Design thinking for doctors and nurses. New York Times. https://www.nytimes.com/2017/08/03/well/live/design-thinking-for-doctors-and-nurses.html. Accessed 25 Jan 2018

Kaufman JC, Baer J, Cole JC, Sexton JD (2008) A comparison of expert and nonexpert raters using the consensual assessment technique. Creat Res J 20(2):171–178. https://doi.org/10.1080/10400410802059929

Kelley T (2001) Prototyping is the shorthand of innovation. Des Manag J Former Ser 12(3):35–42. https://doi.org/10.1111/j.1948-7169.2001.tb00551.x

Kelley T (2007) The art of innovation: lessons in creativity from IDEO, America's Leading Design Firm. Crown Publishing Group, Nat Wartelsc

Kelley T, Littman J (2006) The ten faces of innovation: IDEO's strategies for defeating the devil's advocate and driving creativity throughout your organization. Crown Publishing Group, Nat Wartels

Klufio CA, Kwawukume E, Danso K, Sciarra JJ, Johnson T (2003a) Ghana postgraduate obstetrics/gynecology collaborative residency training program: success story and model for Africa. Am J Obstet Gynecol 189(3):692–696. https://doi.org/10.1067/S0002-9378(03)00882-2

Klufio CA, Kwawukume EY, Danso KA, Sciarra JJ, Johnson T (2003b) Ghana postgraduate obstetrics/gynecology collaborative residency training program: success story and model for Africa. Am J Obstet Gynecol 189(3):692–696. https://doi.org/10.1067/S0002-9378(03)00882-2

Kudrowitz B, Te P, Wallace D (2012) The influence of sketch quality on perception of product-idea creativity. Artif Intell Eng Des Anal Manuf AI EDAM 26(3):267–279. https://doi.org/10.1017/S0890060412000145

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33(1):159–174. https://doi.org/10.2307/2529310

Lauff C, Kotys-Schwartz D, Rentschler ME (2017) What is a prototype? Emergent roles of prototypes from empirical work in three diverse companies. In: ASME 2017 international design engineering technical conferences and computers and information in

engineering conference. American Society of Mechanical Engineers, pp V007T06A033–V007T06A033

Leder H, Carbon C-C (2005) Dimensions in appreciation of car interior design. Appl Cognitive Psychol 19(5):603–618. https://doi.org/10.1002/acp.1088

Leder H, Belke B, Oeberst A, Augustin D (2004) A model of aesthetic appreciation and aesthetic judgments. Br J Psychol 95(4):489–508

Lim YK, Pangam A, Periyasami S, Aneja S (2006) Comparative analysis of high-and low-fidelity prototypes for more valid usability evaluations of mobile devices. In: Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles. ACM, pp 291–300

Marks S (1994) Divided sisterhood: race, class and gender in the South African nursing profession, 1994th edn. Palgrave Macmillan, New York

McPake B, Mensah K (2008) Task shifting in health care in resource-poor countries. The Lancet 372(9642):870–871

Menold J, Jablokow K, Simpson T (2017) Prototype for X (PFX): a holistic framework for structuring prototyping methods to support engineering design. Des Stud 50:70–112. https://doi.org/10.1016/j.destud.2017.03.001

Moe RE, Jensen DD, Wood KL (2004) Prototype partitioning based on requirement flexibility 2004:65–77. https://doi.org/10.1115/DETC2004-57221

Mohedas I, Daly SR, Sienko KH (2014) Student use of design ethnography techniques during front-end phases of design. In: Paperpresented at 2014 ASEE annual conference & exposition, Indianapolis, Indiana. https://peer.asee.org/23059

Mohedas I, Daly SR, Sienko KH (2015a) Requirements development: approaches and behaviors of novice designers. J Mech Des 137(7):071407

Mohedas I, Sabet Sarvestani A, Daly SR, Sienko KH (2015) Applying design ethnography to product evaluation: a case example of a medical device in a low-resource setting. In: DS 80-1 Proceedings of the 20th international conference on engineering design (ICED 15), Vol 1: Design for Life, Milan, Italy

Moogk DR (2012) Minimum viable product and the importance of experimentation in technology startups. Technol Innov Manag Rev 2(3):23

Morgan DL (1993) Qualitative content analysis: a guide to paths not taken. Qual Health Res 3(1):112–121. https://doi.org/10.1177/104973239300300107

Morgan MS, Morrison M (1999) Models as mediators: perspectives on natural and social science. Cambridge University Press, Cambridge

Norman D (1990) The design of everyday things. Doubleday Business, New York

Otto K, Wood K (2000) Product design: techniques in reverse engineering and new product development, 1st edn. Pearson, Upper Saddle River

Parsons MJ (1989) How we understand art: a cognitive developmental account of aesthetic experience. Cambridge University Press, Cambridge

Patton MQ (2014) Qualitative research and evaluation methods: integrating theory and practice. SAGE Publications, London

Roberts JP, Fisher TR, Trowbridge MJ, Bent C (2016) A design thinking framework for healthcare management and innovation. Healthcare 4(1):11–14. https://doi.org/10.1016/j.hjdsi.2015.12.002

Rosen P, Ku B (2016) Making design thinking a part of medical education. NEJM Catalyst website: https://catalyst.nejm.org/making-design-thinking-part-medical-education/. Accessed 25 Jan 2018

Rudd J, Stern K, Isensee S (1996) Low vs. high-fidelity prototyping debate. Interactions 3(1):76–85

Sabet Sarvestani A, Sienko KH (2014) Design ethnography as an engineering tool. Demand ASME Glob Dev Rev 2:2–7

Sauer J, Sonderegger A (2009) The influence of prototype fidelity and aesthetics of design in usability tests: effects on user behaviour, subjective evaluation and emotion. Appl Ergon 40(4):670–677. https://doi.org/10.1016/j.apergo.2008.06.006

Sauer J, Franke H, Ruettinger B (2008) Designing interactive consumer products: utility of paper prototypes and effectiveness of enhanced control labelling. Appl Ergon 39(1):71–85. https://doi.org/10.1016/j.apergo.2007.03.001

Savage J (1995) Nursing intimacy: an ethnographic approach to nurse-patient interaction. Scutari Press, London

Scherer KR (2003) Introduction: cognitive components of emotion. In: Handbook of affective sciences. Oxford University Press

Schrage M (1999) Serious play: how the world's best companies simulate to innovate, 1st edn. Harvard Business Review Press, Boston

Scott JB (2008) The practice of usability: teaching user engagement through service-learning. Tech Commun Quart 17(4):381–412. https://doi.org/10.1080/10572250802324929

Scrivener SA, Harris D, Clark SM, Rockoff T, Smyth M (1993) Designing at a distance via real-time designer-to-designer interaction. Des Stud 14(3):261–282

Seidewitz E (2003) What models mean. IEEE Softw 20(5):26–32

Seva RR, Helander MG (2009) The influence of cellular phone attributes on users' affective experiences: a cultural comparison. Int J Ind Ergon 39(2):341–346. https://doi.org/10.1016/j.ergon.2008.12.001

Sienko KH, Kaufmann EE, Musaazi ME, Sarvestani AS, Obed S (2014) Obstetrics-based clinical immersion of a multinational team of biomedical engineering students in Ghana. Int J Gynecol Obstet 127(2):218–220. https://doi.org/10.1016/j.ijgo.2014.06.012

Silverman D (2010) Doing qualitative research. SAGE, Newcastle upon Tyne

Simon HA (1973) The structure of ill structured problems. Artif Intell 21:181–201

Tractinsky N, Katz AS, Ikar D (2000) What is beautiful is usable. Interact Comput 13(2):127–145. https://doi.org/10.1016/S0953-5438(00)00031-X

Ulrich K, Eppinger S (2015) Product design and development, 6th edn. McGraw-Hill Education, New York

Viswanathan V and Linsey J (2011) Design fixation in physical modeling: an investigation on the role of sunk cost. In: Volume 9: 23rd international conference on design theory and methodology; 16th design for manufacturing and the life cycle conference, pp 119–130. https://doi.org/10.1115/DETC2011-47862

Walker M, Takayama L, Landay JA (2016) High-fidelity or low-fidelity, paper or computer? Choosing attributes when testing web prototypes. In: Proceedings of the human factors and ergonomics society annual meeting. https://doi.org/10.1177/154193120204600513

Weber RP (1983) Measurement models for content analysis. Qual Quant 17(2):127–149. https://doi.org/10.1007/BF00143616

Weiss RS (1995) Learning from strangers: the art and method of qualitative interview studies. Simon and Schuster, New York

Winston AS, Cupchik GC (1992) The evaluation of high art and popular art by naive and experienced viewers. Visual Arts Res 18(1):1–14

World Health Organization (2010) Landscape analysis of barriers to developing or adapting technologies for global health purposes. World Health Organization. http://www.who.int/iris/handle/10665/70543

Yang MC, Epstein DJ (2005) A study of prototypes, design activity, and design outcome. Des Stud 26(6):649–669. https://doi.org/10.1016/j.destud.2005.04.005

Yock PG, Zenios S, Makower J, Brinton TJ, Kumar UN, Watkins FTJ, Kurihara C (2015) Biodesign: the process of innovating medical technologies, 2nd edn. Cambridge University Press, Cambridge