Doctoral theses at NTNU, 2023:217

Marius Auflem

Development of Medical Training Equipment through Prototyping, Wayfaring, and Triple Loop Learning

Industrial Research at Laerdal Medical

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of Philosophiae Doctor
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU
Norwegian University of Science and Technology

Marius Auflem

Development of Medical Training Equipment through Prototyping, Wayfaring, and Triple Loop Learning

Industrial Research at Laerdal Medical

Thesis for the Degree of Philosophiae Doctor

Trondheim, June 2023

Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU
Norwegian University of Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Engineering
Department of Mechanical and Industrial Engineering

ISBN 978-82-326-7136-6 (printed ver.)
ISBN 978-82-326-7135-9 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2023:217

Printed by NTNU Grafisk senter

Abstract

Prototyping has become widely adopted across engineering design practices and is suggested practical and reflective advantages in new product development. Although this has yielded increased research focus, holistic and empirically grounded insight into —what prototyping entails in projects, how prototyping is facilitated, and why prototyping should be encouraged is nevertheless limited. This thesis involves an industrial research project joint between Laerdal Medical and TrollLABS NTNU, regarding prototype-driven development in the fuzzy front end. The project considers the context of simulation-based medical training equipment and investigates how prototyping contributes to their development.

This thesis includes 11 academic contributions and 10 case projects and initiatives presented and utilized in a binding article. The research project has pursued and addressed three goals: 1.) Explore new product opportunities within the healthcare training domain. 2.) Develop and utilize new technology for healthcare training applications. 3.) Enhance process knowledge and methodological foundation on which Laerdal Medical, as the industry partner and stakeholder, develop new concepts and technology.

The collective research output of case projects, activities, and academic contributions address the two first goals, while the binding article of this thesis addresses the third. Chapters 4, 5, and 6, each viewed from an organizational level, —describe prototyping to develop new medical training equipment, facilitate prototype-driven development by wayfaring, and expand organizational knowledge by prototyping. The chapters address *what*, *how*, and *why* we prototype in industrial and research settings, and they account for three individual learning loops. Respectively, the chapters present empirical insights for operational, tactical, and strategic use of prototyping in the fuzzy front end of new product development.

Sammendrag

Prototyping har blitt utbredt I design- og produktutviklings-praksis, og er sett på som et nyttig verktøy for å skape og lære under utviklingen av nye produkter. Selv om det har blitt forsket på prototyping, mangler man en helhetlig forståelse av hvordan prototyping brukes I prosjekter, hvordan man kan legge til rette for prototyping, og hvorfor prototyping bør fokuseres på. Denne oppgaven beskriver et samarbeidsprosjekt mellom Laerdal Medical og TrollLABS NTNU, som omhandler prototypedrevet utviklingen i *Fuzzy Front End*. Prosjektet har involvert utviklingen av medisinsk simulering- og trenings-utstyr, og hvordan prototyping bidrar til denne utviklingen.

Oppgaven inneholder 11 vitenskapelige artikler og 10 case -prosjekter og -initiativer som blir brukt I en sammenfattende avhandling. Forskningsprosjektet har følgende mål: 1.) Utforske nye produktmuligheter innen medisinsk opplæring og trening. 2.) Utvikle og nyttiggjøre seg av ny teknologi og applikasjoner for medisinsk opplæring og trening. 3.) Styrke kunnskapsfundamentet som samarbeidspartner og prosjekteier, Laerdal Medical, bruker til å utvikle nye konsepter og teknologi.

De samlede forskingsresultatene, i form av case-prosjekter, aktiviteter og akademiske bidrag, tar for seg de to første målene, mens denne avhandlingen vil ta for seg det tredje. Kapittel 4, 5 og 6, alle sett fra ett organisasjonsnivå, beskriver prototyping for å utvikle nytt medisinsk treningsutstyr, tilrettelegge for prototypedrevet utvikling ved hjelp av wayfaring, og forbedre organisasjonskunnskap og læring ved prototyping. Kapitlene tar for seg *hva*, *hvordan* og *hvorfor* vi prototyper i industrielle- og forsknings -miljøer, og de står for tre individuelle læringssløyfer. Henholdsvis presenterer kapitlene empirisk innsikt for operasjonell, taktisk og strategisk bruk av prototyping i *Fuzzy Front End* av produktutvikling.

Preface

This thesis has been submitted to the Norwegian University of Science and Technology (NTNU) for the degree of Philosophiae Doctor (PhD). The PhD project has been conducted at TrollLABS within the Department of Mechanical and Industrial Engineering (MTP), Faculty of Engineering Science (IV), NTNU and at Laerdal Medical in Stavanger. The research has been supported by the Research Council of Norway (RCN) through its industrial PhD funding scheme, project number 290404. The research project has been supervised by Professor Martin Steinert (NTNU), and co-supervised by Arild Eikefjord (Laerdal Medical), Hans Gundersen (Laerdal Medical), Jørgen Falck Erichsen (NTNU), and Magnus Ove (Laerdal Medical).

Trondheim, January 2023

Acknowledgements

Without TrollLABS, the lab environment and the people it brings together, I would not have started nor finished a PhD. For this, I want to thank my supervisor and head of lab Professor Martin Steinert.

I want to thank my fellow Trolls and great colleagues, Håvard, Sampsa, Henrikke, Torjus, Daniel, Benjamin, Øystein, Sindre, Kjetil, Pasi, Kim, Ole, Jørgen F.E, Carlo, Heikki, Achim, Andreas, Kristoffer, Jørgen B., Evangelos, and Christer. To all master's students, fuzzy students, and Laerdal summer interns, thank you so much for being a part of my research and contributing greatly by being eager to learn, explore, and go beyond what is expected.

Enhet for helsefaglig simulering (EHS), DRIV, Trams, Medisinsk simulatorsenter (MSS), Safer, and healthcare personnel from St. Olavs, SUS, and Norsk Luftambulanse. To all that have contributed, thank you so much.

To Laerdal Medical, thank you for giving me this opportunity. Especially Arild Eikefjord and Hans Gundersen, from taking me in as a summer intern, providing guidance, encouragement, and contributing to the project, thank you. Magnus Ove, the SimMan team, and the HTP organization, thank you so much for your interest and belief in me and the project. This is truly an amazing place to work and too many great colleagues to fit onto this page. Production, engineering, designers, sports- and football -team, and the rest of the organization, thank you.

To my family and friends. This would never have been possible without your support and love. I am truly grateful. To Aleksander, thank you for being an amazing brother and an even better friend. I am proud of you. Til Mamma og Pappa, thank you for supporting and taking care of me, without your love, encouragement, and unconditional support I would never have managed. I promise my surfboards, and I will move out (eventually). To Ninni, you are truly fantastic, and I cannot sufficiently express my gratitude for your support, love, and belief in me. Thank you. Ps. Can me and my surfboards move in with you?

Table of contents

Figures

Tables

Abbreviations

CPR	Cardiopulmonary Resuscitation
IV	Intravenous
OSH	Open-Source Hardware
EHP	Experienced Healthcare Practitioner(s)
SBMTE	Simulation-Based Medical Training Equipment
NPD	New Product Development
FFE	Fuzzy Front End
CAD	Computer Aided Design
HMW	How Might We
GDQ	Generative Design Question(s)
DRQ	Deep Reasoning Question(s)
OSH	Open-Source Hardware
AI	Artificial Intelligence
ML	Machine Learning

1

Industrial Research Project

What, How, and Why?

This thesis describes an industrial PhD project with Laerdal Medical – a research, development, and manufacturing company and worldwide provider of medical training solutions, as the industry partner and stakeholder. It considers applied research within engineering design, focusing on prototyping in the fuzzy front end – the earliest stage of new product development where uncertainty is high. The project has entailed research on the pre-requirement development of simulation-based medical training equipment, and it has involved gaining insights into how prototyping facilitates exploring new concepts, technologies, and applications within healthcare training. Empirical cases, participatory, and applied research methods through development projects are at the core of this thesis.

Figure 1.1 Laerdal headquarters and a Laerdal human patient simulator. Source: (Laerdal stock images, 2022).

This research spans healthcare, development, and technology domains. It involves projects and activities within industrial development and academic research contexts. Hence, the thesis will investigate how synergies and knowledge transfer could bridge the gap between these domains and research contexts. This chapter introduces the collaboration on which the research project is based, the actors involved, and the motivation for research within the domain of Laerdal Medical. Furthermore, the research project's goal will be presented. Aim and scope, methods, readers guide, and contributions of this thesis will follow.

1.1 Research Collaboration between Laerdal Medical and TrollLABS

This PhD project is part of a mutually beneficial collaboration between Laerdal Medical and TrollLABS (NTNU). In 2015 this relationship was initiated to explore opportunities and synergies between the development and research areas of the two actors. This collaboration and project have evolved and formalized across three core research areas of product development. Firstly, exploring *tools and technology*, how we develop new products and what functionality they embed. Secondly, *problems and applications*, exploring real-world challenges and product opportunities within healthcare training. And thirdly, applied *methods* and *process* knowledge for research, development, and accelerated learning in new product development.

Core Research Areas:

Tools and Technology

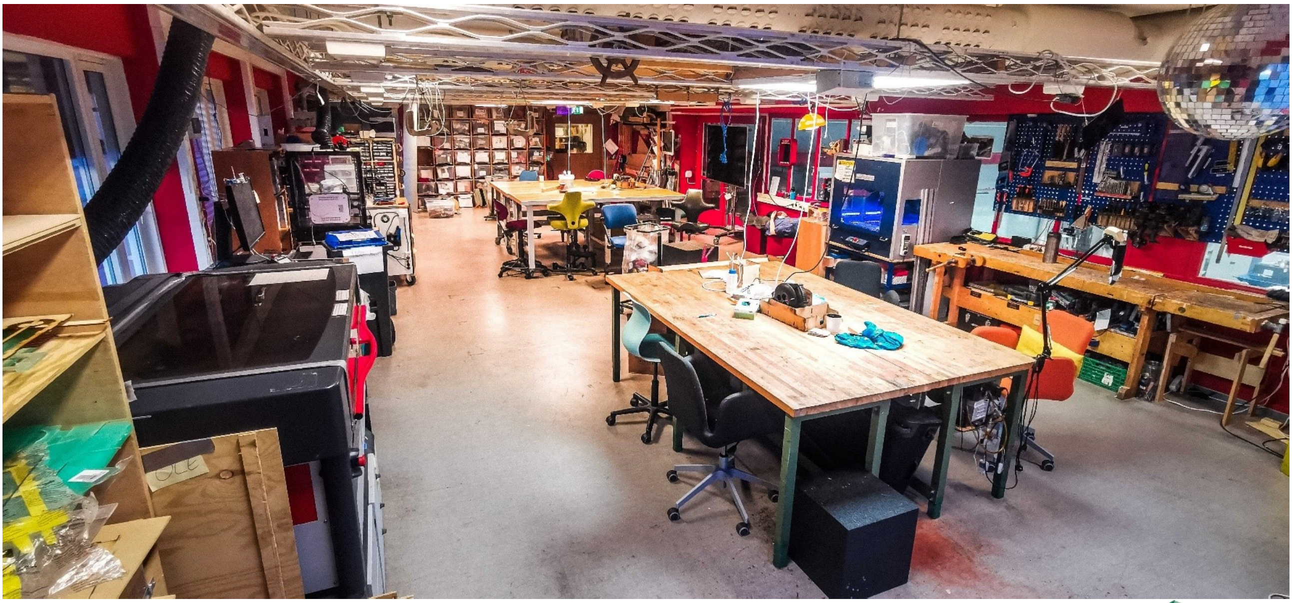
Problems and Applications

Methods and Process

What is **Laerdal Medical?**

A publisher and toy manufacturer from Stavanger (Norway), Laerdal, designed and produced the first mannequin for resuscitation training, the *Resusci Anne* in the 1960's. Following the impact and importance of *Resusci Anne*, Laerdal pivoted their development and production to focus entirely on medical training and clinical equipment. Ever since, Laerdal Medical has been on the forefront of training equipment and human patient simulators, delivering a wide range of products world-wide. Laerdal has more than 2000 employees in 25 countries, with the headquarters located in Stavanger. Laerdal has set the ambitious goal of helping save one million more lives each year by 2030, a large portion of which from preventable deaths due to insufficient, or lack of, medical treatment (*One Million Lives by Laerdal Medical*, 2020).

The collaborative efforts have been aimed at exchange of experiences, projects, and people between industry and research. This has involved Laerdal Medical providing real world challenges from the context of medical training to be explored and solved in an educational and research setting at TrollLABS. This has included providing product development challenges in courses such as *TMM4245 Fuzzy Front End*, as master's projects, or as research projects. The exchange of experiences has been facilitated by joint activities where students



What is **TrollLABS?**

TrollLABS is a research lab (and community) at the department of mechanical and industrial engineering at the Norwegian University of Science and Technology, (NTNU). At TrollLABS we are trying to understand the mechanisms of radical innovations and prototyping in early stages of product development. TrollLABS is a workshop and makerspace that is equipped with various machines and tools to enable rapid prototyping. It is furthermore a multidisciplinary group, and the space is therefore equipped to serve various purposes ranging from prototyping mechanisms, mechatronic systems, test setups and collaborative research activities. TrollLABS is dynamic, by being easily restructured given the needs of the people working there.

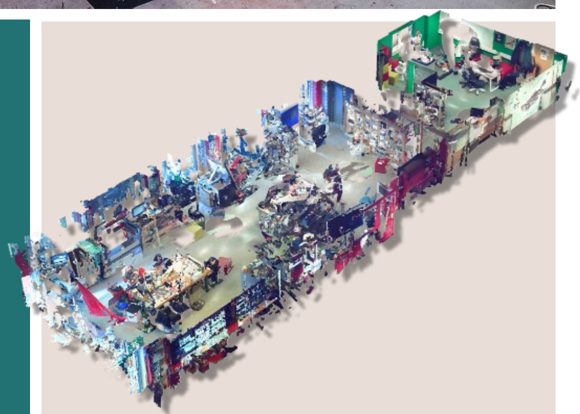


Figure 1.2 Picture of the TrollLABS workshop facilities and a 3D scan of the space. Used with permission (Aalto, 2022).

and researchers have visited the Laerdal headquarters in Stavanger, Norway. Furthermore, Laerdal employees have partaken in workshops, presentations, and discussions at the university in Trondheim, Norway. Students from TrollLABS have had internships or been employed permanently by Laerdal Medical because of this relationship.

The role of the PhD project has been a formal integration in Laerdal Medical. It has been allowed autonomy for independent and unconfined research and development projects to be performed interchangeably between TrollLABS and Laerdal. As a PhD student, the author has had the opportunity to coach students and projects at TrollLABS, collaborate with other PhD projects, and do independent development and research related to Laerdal Medical's challenges. Ideas and methods originating from the research setting, have been applied and tested in an industrial context with different parts of the Laerdal organization. By being employed by Laerdal Medical throughout this project, the author has had a 25% position as a product developer. This has ensured allocated time to coach and facilitate intern projects, contribute to ongoing development activities, and work on integrating technology, working modes, and novel product opportunities within the company.

1.2 Motivation for Researching Product Development in Laerdal

By the sustainable development goals provided by the UN in 2016 at its core, Laerdal has a goal and formalized strategy *to help save one million more lives by 2030* (*One Million Lives by Laerdal Medical*, 2020). The underlying initiatives for reaching this goal are diverse but enhancing the life-saving capabilities and competency of healthcare workers worldwide is imperative (Medicine & America, 2000). From a product development perspective, this requires innovation. Hence, we ought to develop opportunities into products, services, and solutions that have a positive impact on user's lives. This means products that can enable improvement and retention of performance, skillset, and knowledge of healthcare workers by products being widely adopted and used. With the backdrop of technological advancements, global connection, and sharing of knowledge, the way we think, act, and reflect in development of new products are evolving. This sparks the question if ideas are getting harder to find (Bloom et al., 2020), and subsequently, if we possess the right toolbox and approaches for hunting and developing disruptive ideas to make an impact (Steinert & Leifer, 2012)?

By a wide product portfolio, Laerdal Medical covers a broad product space within healthcare. This ranges from clinical equipment, such as bag valve masks and suction units, to advanced human patient simulators and various task and skill trainers. Subsequently, they account for considerable in-house production and prototype capabilities, assembly lines and testing facilities, which Laerdal Medical has established in Stavanger, Norway. This ensures strong links between the core functions from development and delivery of finished goods. While mechanical and manufacturing capabilities are maintained, cyber-physical systems, electronics, fluidic, and electromechanical systems are key elements of their modern simulators. Adding to their physical products Laerdal also offers a range of digital solutions, covering a wide span of the training and practice environment within healthcare. E.g., feedback devices for clinical interventions, debrief solutions for simulations, and control systems for patient simulators. Having this span of products, the link with users and stakeholders within healthcare is critical. Moreover, this bridges several domains spanning from clinical practice, technical staff, and educational environments. These preconditions describe product context and development domain that inherently introduce great complexity. Arguably, this also entails many and hitherto unknown opportunities, which motivates the goals of this research project.

1.3 Goals, Aim, and Scope

1.3.1 Goals for the Industrial Research Project

The goals for the research project have been threefold. On a fundamental level, it has been to explore and exploit new product opportunities within the domain of healthcare training. This involves finding needs for training applications, establish product opportunities, developing concepts, and exploring methods to enhance current training. Secondly, this has yielded needs and opportunities to explore, develop, and test new technology within this product context.

Finally, a goal has been to synthesize learnings and insights, to inform and enhance the Laerdal product development model seen in Figure 1.3.

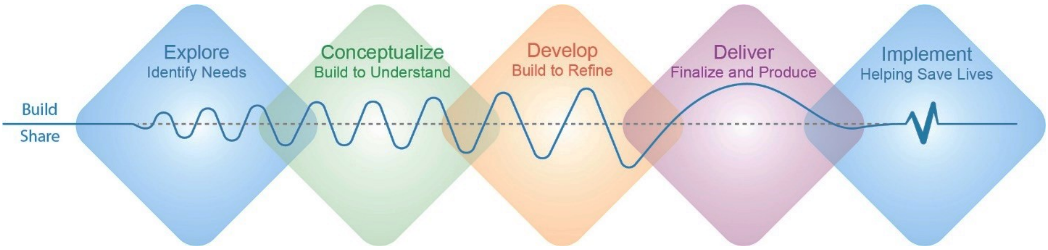


Figure 1.3 The Laerdal Medical product development model.

1.3.2 Aim and Scope of this Thesis

This thesis addresses the project goals through three distinct, yet interconnected, research and project outputs. Project contributions are product development projects and activities carried out throughout this PhD project. These are presented as cases to provide insights into product development opportunities and resulting project outcomes. Academic contributions present novel concepts, technology, and methodological findings, to support the development of medical training equipment. This thesis accounts for the third contribution and aims to synthesize findings and insights from the cases and academic contributions to support —development practitioners, development managers, and development organizations to leverage prototype-driven development and wayfaring. Hence, the following questions are to be addressed by this thesis:

- What we do in prototype-driven development projects and how prototyping methods facilitate project insights for development of medical training equipment?
- How we think in product development projects, and how could we utilize questions and a wayfaring framework to facilitate prototype-driven development?
- Why do we prototype in development organizations, and what are the strategic and organizational benefits of prototyping, wayfaring, and industrial research projects?

The thesis's scope is bound by three abstraction levels and the contexts of the projects it involves. The product context is the development of simulation-based training equipment. The development context is prototyping in the fuzzy front end of medical training equipment, while the research context is applied industrial research using a triple loop learning framework.

1.4 Methods

Innovation projects, novel outcomes, and creative processes add to a complex, dynamic and contextual scene for doing research. Thus, insights and empirical examples will be used to suggest improvements of current development practice using experimental case studies (Yin, 2012). The thesis relies on grounded theory by gathering empirical observations from cases for iterative and continuous theory building (Glaser & Strauss, 1967). Moreover, the author has assumed observatory and acting roles within all the cases, suggesting biases or causal effects to be unavoidable. Yet, participatory research has enabled deep and contextual insights, allowing theoretical sampling and replication logic to support the findings and resulting arguments (Eisenhardt, 1989). However, this also proposes that replicating the findings for ascertaining the theories might not be possible, even though the cases show replicability under comparable conditions and process methods.

With the three abstracted levels of product development in organizations mentioned, the thesis aims to provide a holistic image by building from multiple cases viewed across different levels and viewports through triple loop learning. This enables the author to assume different roles for building the arguments and presenting the cases, as indicated in Figure 1.4. Triple loop learning suggests a higher modality of learning by addressing what we apply, how we apply it, and why we apply it (Argyris & Schön, 1997; Tosey et al., 2012). Contextualized for this thesis, triple loop learning is viewed by organizational learning and design practice through the operational, tactical, and strategic layers, as suggested by (O. Eris & Leifer, 2004; Leifer & Steinert, 2011).

Triple Loop Learning

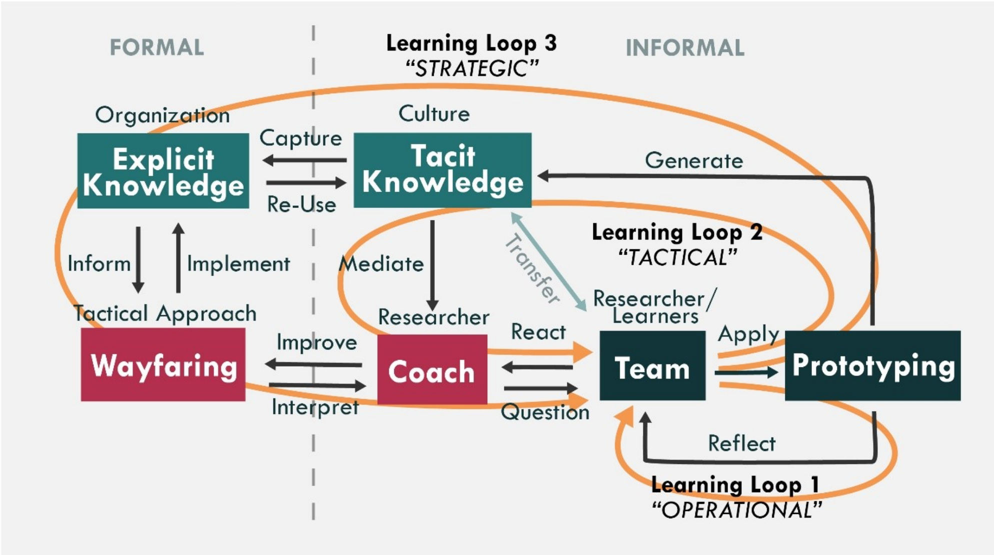


Figure 1.4 Triple loop learning illustration for this project. Coloured boxes indicate the relation to chapters, i.e., loop 1 – chapter 4, loop 2 – chapter 5, loop 3 – chapter 6. The figure is adapted from (O. Eris & Leifer, 2004; Leifer & Steinert, 2011).

1.5 Readers Guide

The remaining sections of this chapter will introduce the project contributions, by the role the author have assumed in the projects, and how they are related to the academic contributions. The academic contributions will be presented and briefly described. **Chapter 2** provide the required context and definitions for the framing and grounding of this thesis. This considers product context, i.e., the development of simulation-based training equipment. Development context, i.e., prototyping in the fuzzy front end. Methodological context, i.e., what development and thinking modes are encouraged and practiced at TrollLABS, wayfaring, organizational knowledge, and triple loop learning. **Chapter 3** describes the case projects and activities considered by this thesis. Furthermore, the following chapters, 4, 5, and 6, describe the three learning loops mentioned and, thus, are considered to address different readers.

- **Chapter 4**: *What we do* – Are addressed to the design practitioner and product developer. This considers insights and considerations for deploying prototype-driven development of medical training equipment. It accounts for the first learning loop.

- **Chapter 5**: *How we think* – Are addressed to the project coach or manager. This synthesizes findings and insights from development projects with the aim of operationalizing prototype-driven development in a wayfaring framework. It accounts for the second learning loop.
- **Chapter 6**: *Why we Prototype* – Presents organizational learning and organizational strategy through the lens of prototype-driven development. It drives a discussion on the limitations and further research opportunities concluding this research project. It accounts for the third learning loop.

Chapter 7, final reflections, presents the authors reflections on opportunities for continuation of similar initiatives, and implications for doing industrial research projects within the scope and contexts described in this thesis.

1.6 Project Contributions

Project contributions considered in this thesis are development projects where the author have participated either as main contributor, design team member, or project coach. These are used as cases to provide insights and empirical observations on development of medical simulation and training equipment. Development activities within Laerdal Medical are furthermore contributing to the application of these insights. These abstraction levels of performing projects, gathering insights, and attempt at incorporating changes to current practice builds on the foundation of the *triple loop learning* approach utilized in this thesis.

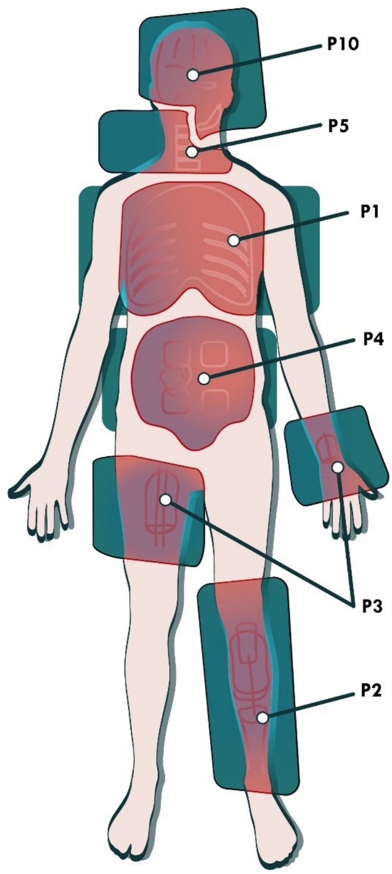


Figure 1.5 Overview of the various development projects that are used as cases.

Table 1.1 Overview of case -projects and -activities. Connection to academic contributions and authors role are presented.

Case	Explanation	Used in Publication(s)	Role
P1	Chest for CPR Training	C1	Main developer
P2	Break-A-Leg	C1	Coach
P3	Vascular Phantom(s)	C8, C10	Contributor
P4	Palpation Trainer(s)	C4, C5, C6, C7	Coach
P5	A Pain in the Neck	C11	Coach
P6	TrollLABS Medical		Contributor
P7	Joint Workshops and Sharing Sessions		Coach
P8	Intern Technology Projects		Coach
P9	Product Development Sprints	i1	Coach
P10	Expressive Simulator Face	C3	Main developer

The author, who served as the main developer in projects P1 and P10, had the primary responsibility and made the main contribution to their development. In projects P3 and P6, the author played a team role as a contributor and was responsible for distinct parts of the development. As a coach, the author supervised and followed up on student course projects (P2), master's projects (P4 and P5), and summer intern projects (P8). This included having regular scrum meetings and sparring sessions with the teams, either daily or weekly. Furthermore, this role also included providing technical support and contributing to prototyping and thus parts of the projects. Further explanation of how the roles were assumed is made in the case descriptions in chapter 3.

1.7 Academic Contributions

This section introduces the academic contributions made throughout this PhD project. They contribute to the research on prototyping, design methodology, prototyping, and testing simulation-based medical training equipment, and novel and enabling technologies for healthcare training applications. The contributions will be briefly introduced and provided a storyline highlighting achievements and findings. They consist of 10 published research articles, one patent application, and a paper from an internal development initiative within Laerdal Medical. The articles are appended at the end of this thesis.

C1 *Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators*

Marius Auflem, Jørgen Falck Erichsen, Martin Steinert (2019),

Procedia CIRP 84, 572-578: CIRP Design 2019 Conference

This article exemplifies prototype-driven development, as opposed to specification-driven. It emphasizes using prototypes as knowledge artefacts for sharing and communicate ideas across disciplinary and organizational boundaries. Insights revealed from prototyping establish requirements for future products. Thus, prototypes are the main drivers of the development. The paper contributes to the knowledge gap on how to do prototype-driven development in projects. This is aimed at engineering design practitioners in the fuzzy front end and in the context of developing medical training equipment.

C2 *On Prototyping Methods to Leverage Non-Rigid Materials in the Early Stages of Engineering Design*

Marius Auflem, Hans Hagenes Bøe, Jørgen Falck Erichsen, Martin Steinert (2020),

Proceedings of the Design Society: DESIGN2020 Conference

Using serval case examples, this paper presents various methods for creating prototypes with non-rigid characteristics. It describes their applicability in various contexts and how they are used, performed, and contribute to design practice. Methods presented are soft prototyping, living hinges, flexible fused filament fabrication, and moulding. Furthermore, fabrication methods are evaluated based on their effect on design activity, highlighting trade-offs designers encounter and should acknowledge when prototyping.

C3 ***Facing the FACS-Using AI to Evaluate and Control Facial Action Units in Humanoid Robot Face Development***

Marius Auflem, Sampsa Kohtala, Malte F Jung, Martin Steinert (2022),

Frontiers in Robotics and AI

This paper presents a new approach for evaluating and controlling expressive humanoid robotic faces using open-source computer vision and machine learning methods. The approach is not dependent on specific facial attributes nor actuation capabilities, making it applicable for different robots at different stages of development to inform design decisions. The results from the presented case example demonstrate the flexibility and efficiency of the presented AI-based solutions to support the development of humanoid facial robots.

C4 ***Dealing with Ecological Validity and User Needs when Developing Simulation Based Training Equipment – Case Study of a Medical Palpation Task Trainer***

Daniel Nygård Ege, Oscar Lilleløkken, Marius Auflem, Martin Steinert (2020),

Procedia CIRP 91, 722-727: CIRP Design 2020 Conference

Complex interactions with patient simulators make it difficult to elicit and fixate requirements in early-stage development. This paper showcase how testing prototypes with both experienced and novice users in healthcare, could help navigate complexity and determine requirements for ecological valid solutions and addressing user needs. This is enabled by prototyping in different domains and different abstraction levels. From coaching the project this article is based on, the author has contributed to the design of the study as well as conceptualizing and writing of the manuscript.

C5 ***Proposing a novel approach for testing complex medical task trainer prototypes***

Oscar Lilleløkken, Daniel Nygård Ege, Marius Auflem, Martin Steinert (2020),

DS 101: Proceedings of NordDesign

This paper presents a standardized testing framework for medical task trainer prototypes to support designers in the early design stages. The framework is suggested to accommodate the ambiguous challenge of developing and testing prototypes combining subsystems and functionality to address user needs and ensure training validity. The framework is discussed as a potential tool in the development of medical task trainers. From coaching and supporting the project on which this contribution is made, the author also contributed to framing, writing, and reviewing of the manuscript.

C6 ***User Involvement in Early-Stage Design of Medical Training Devices – Case of a Palpation Task Trainer Prototype***

Daniel Nygård Ege, Marius Auflem, Oscar Lilleløkken, Martin Steinert (2021),

Design for Health

This paper uses a novel concept for abdominal palpation training to exemplify challenges and approaches for designing new medical training equipment. A Likert scale questionnaire, clinical assessment experiment, and recorded sensor data are used to evaluate prototype functionalities. The paper discusses observations from user interactions and considerations regarding the fidelity of medical training equipment prototypes. Multifaceted evaluation of conceptual prototypes in early design stages can ensure that user needs, product requirements, and sufficient fidelity are collectively addressed in new medical training equipment. For this publication the author also contributed to design of the study as well as supporting the main author during the experiment, analysis, and presentation of the results. The author also contributed to both the writing and review of the manuscript.

C7 *Perception by Palpation: Development and Testing of a Haptic Ferrogranular Jamming Surface*

Sigurd Bjarne Rørvik, Marius Auflem, Henrikke Dybvik, Martin Steinert (2021),

Frontiers in Robotics and AI

This paper presents a novel haptic surface concept for palpation training using ferrogranular jamming. The concept can generate a variety of geometric shapes (output) by manipulating and arranging granules with permanent magnets. The tactile hardness of the palpable output is altered via pneumatics. A psychophysical experiment (N = 28) investigates users' tactile perception of objects and evaluates the haptic interface concept. Analysed results on palpation skills and concept performance show potential for the technology to enhance medical palpation applications. The author contributed to the development of the ferrogranular jamming principle and the design of the experiment. Furthermore, writing, review, graphical work, and being corresponding author for the publication.

C8 *Embedded Soft Inductive Sensors to Measure Arterial Expansion of Tubular Diameters in Vascular Phantoms*

Torjus Steffensen, Marius Auflem, Håvard Vestad, Martin Steinert (2022),

IEEE Sensors Journal

In this article, we present a soft braided coil embedded in an elastomer tube as a method to continuously measure change in diameter. By the change in inductance in the braided coil, we estimate the expansion and contraction through an empirical calibration model. Applying this method, we demonstrate that diameter waves in a vascular phantom, a model of a radial artery embedded in a viscoelastic wrist structure, can be recorded continuously. This opens opportunities for investigating physiologically interesting fluid-structure interactions. Furthermore, this technology can provide new ability to measure diameter changes in tubular systems where access is obstructed, or space is restricted. For this article the author was responsible for development of the sensors and prototyping the technology, as well as contributing to the design of the study and carrying out the experiments.

C9 ***Lost in Transit: Implications and Insights for Making Medical Task Trainer Prototypes with an Open-Source Hardware Paradigm***

Daniel Nygård Ege, Marius Auflem, Martin Steinert (2022),

Proceedings of the Design Society: DESIGN2022 Conference

This paper presents an open-source novel intravenous cannulation (IV) task trainer developed during the Covid-19 pandemic for unsupervised clinical skill practice. When observing 13 registered nurses using the task trainer during a two-hour unsupervised skill training session, multiple user errors were uncovered. These insights raise the question of how open-source hardware (OSH) needs to share more than just device descriptions and assembly instructions, as designs are being shared only in the current state of an ongoing project. Sharing insights, user errors and test results should be encouraged and prioritized in development of OSH designs. While being a coach for this project the main contribution of the author for this paper was in framing, writing, and revising the manuscript.

C10 ***Ultrasound Phantom***

Torjus Lines Steffensen, Carlo Kriesi, Martin Steinert,

Thomas Lafrenz, Jostein Rødseth Brede, Marius Auflem (2022),

US Patent App. 17/641,651

Patent application of ultrasound phantom. Technology and material for realistic tactile feel and reflected ultrasound image. The author is in this contribution listed as one of the inventors based on the involvement in the earliest stages of the project. The publication is not appended in the thesis but can be accessed by the patent publication number: US20220304922A1.

C11 A Pain in the Neck: *Prototyping and Testing of a Patient Simulator Neck for Spinal Immobilization Training*

Emil Matias Henriksen, Marius Auflem, Martin Steinert (2023),

Design for Health

This paper presents a novel concept for neck immobilization training using patient simulators. Insights from testing conceptual prototypes with experienced healthcare practitioners revealed a need for more sufficient resources and evidence-based training procedures to support and inform spinal immobilization training. Consequently, we present a realistic and compliant neck concept with sensor feedback. An experiment (n=12) provided results that are used to present and discuss performance metrics and whether these can be used as quality performance indicators for both formative and summative feedback in training. The results suggest that the developed prototype is a good representation of an unconscious patient's neck regarding the range of motion and spinal compliance. Coaching this project, the author contributed to the design of the study, support during experiments and analysis of motion data utilizing machine vision tools. Further contributions involved writing, reviewing, illustrations, and being the corresponding author for the finished manuscript.

i1 *On Iterative Prototyping, Testing, and Learning in New Product Development Projects - Accelerate Concept Development through Internal Prototyping Workshops*

Marius Auflem, Sindre Hegglund-Dalseth, Arild Eikefjord (2020)

Internal Project Paper for Laerdal Medical

This paper was written and shared internally within the company to showcase working methods and facilitate adoption of rapid and prototype-driven sprints. Further described by project case P9.

2

Contexts and Definitions

for Medical Training Equipment and New Product Development

This chapter describes the product and development context of this thesis, namely the development of medical training equipment and prototyping in the fuzzy front end respectively. Firstly, describing the product context, the background and grounding for use of simulation-based medical training equipment will be provided. Furthermore, definitions and considerations for developing novel products within this domain will be introduced and highlighted. The development context describes the use of prototyping to deal with complexity and ambiguity of the fuzzy front end. Lastly definitions and contextualizing this thesis by organizational and triple loop learning will contribute to clarify the aim and scope of the presented research.

2.1 Medical Training Equipment and Patient Simulators

Simulation-based training has become significant in current medical curricula and been implemented in various healthcare educational paradigms. This training offers opportunities for repeated practice, under safe conditions, without endangering or causing discomfort for real patients. Moreover, realistic, rare, and specific scenarios can be created on demand, and training can be tailored to specific needs and competency levels of trainees (Aggarwal et al., 2010). However, simulated training requires enabling and adequate products that allow for clinical interactions and practice of relevant cognitive, psychomotor, and interpersonal skills (Alinier, 2007; Issenberg et al., 2005). This constitutes the needs for the product context considered by this thesis, namely the development of simulation-based medical training equipment.

Simulation-based medical training equipment (SBMTE), or simulators, is a broad range of products, services, and solutions that facilitate safe and repeatable training for healthcare personnel. Products within this domain span a continuum of fidelity levels, i.e., closeness to real patients, ranging from simplified and abstracted task trainers (low fidelity) to realistic human patient simulators or mannequins (high fidelity) (Aggarwal et al., 2010). These products provide significant training benefits over traditional mentor and apprentice arrangements that relies on real patient-interactions as described by Maran & Glavin, (2003). These are summarized by the following:

- Avoiding risks to patients and learners.
- Reducing undesired interference.
- Tasks and training can be created and tailored on demand.
- Repeatability.
- Increased accuracy and retention.
- Enhanced transfer of training to clinical situations.
- Objective and standardized student performance evaluation.

2.1.1 Use and Characteristics Simulation-Based Training Equipment

Skills training is described by Miller, (1990) as the practice for acquiring, enhancing, and retention of medical knowledge (knowing), competence (knowing how), performance (showing how), and actions (reacting)(Issenberg et al., 2005). Products to facilitate medical skills training are becoming increasingly available. Yet, due to the novel (and evolving) domain, we lack comprehensive understanding of the relations between required product characteristics and the different training -objectives and -conditions (Alinier, 2007). Thus, products used across healthcare education varies by functional traits, fidelity, and consequently, their ability to facilitate training of clinical skills. Furthermore, competency levels of trainees influence training outcome and requirements for the products utilized. For example will skilled practitioners using lower fidelity simulators acquire fewer new skills than novice trainees under the same conditions (Aggarwal et al., 2010). Lower fidelity could however, provide sufficient training conditions for the enhancement and/or retention of skills that have been obtained (Issenberg et al., 2005). Conversely, the benefits of using higher fidelity simulators are enhanced by trainees having the required knowledge and skillset to perform and utilize these in a scenario (showing how). This involves acquiring and elevating higher modalities of skills such as decision making, communication, and procedural strategies (reacting) (McGaghie et al., 2006).

A critical notion is however, that closeness to a real patient (fidelity), is an insufficient typology for describing simulation features and adjacent equipment functionality (Issenberg et al., 2005; McGaghie et al., 2006; Meller, 1997). As will be discussed in this thesis, the design and evaluation of novel training equipment will need to address fidelity multidimensionally, i.e., along functional, engineering, and physiological attributes. Furthermore, future SBMTE are required to attend emerging needs in healthcare training, as well as leverage the technological advancements shifting the healthcare paradigms and expectations for learning experiences at large.

2.1.2 Challenges and Opportunities for Future Medical Training Equipment

A core objective for developing new simulators is to emulate human physiology sufficiently to enable transfer of learning to clinical settings. How to achieve this is not evident, and lack of available products and resources often leads to standardized patients and human markers being used in current training. However, in many cases the training requires tactile abnormalities not found in healthy actors, or the procedure is too dangerous and painful to perform on real people. Therefore, there is a strong need for new training equipment that could effectively bridge the gap from training to clinical practice. Using and advancing technology to address *what* and *how* procedures could be simulated, is a core source for future medical training product opportunities.

The most important feature of simulation-based medical education is training feedback (Issenberg et al., 2005). Feedback can broadly be divided as either summative, for retrospective performance scoring, or formative, performance improvement during the learning activity (McGaghie et al., 2006). For this thesis, the formative feedback is of greatest interest, as it could be addressed by product functionality such as sensors and smart systems providing corrective suggestions to the trainees. Products providing sensor-based objective feedback also spawn opportunities for self-directed or peer-directed training, thus alleviating time and resource allocation for running simulation-based scenarios (Aggarwal et al., 2010; Issenberg et al., 2005). Moreover, feedback could be tailored to the performance level and learning objectives set for individuals. Needs for adaptive interplay between product functionality (difficulty) and trainee skill level (capability) should be considered across all the various passive (design), active (functions), and interactive (control) aspects of future simulators (Meller, 1997). This emphasizes the need for thoughtful integration of sensors, and opportunities for future simulator systems to become dynamic, adaptable, and autonomous.

Realism of simulators concerns multiple modalities of interactions in medical training. Lack of realism could suggest dehumanizing effects of increased use of medical simulation in healthcare education. Thus, concerning implications on interpersonal and communicative skills of future practitioners (Issenberg et al., 1999). This suggests a need for humanlike (and advanced) attributes in physical simulators, e.g., nonverbal communication traits, diversity, dynamic and patient specific characteristics. Such attributes could increase users' immersion, i.e., the experienced sensation of realistic patient interactions, and moreover, enhance transfer of learning to clinical settings. These considerations (and product features) however, spans beyond merely realism and simulator validity (Stunt et al., 2014). More importantly they considers scenario realism and ecological training validity of the interactions therein, which needs to be explored and ascertained (Kushniruk et al., 2013). These elements will be discussed further in chapter 4 of this thesis.

2.1.3 Development of Simulators

The indicated complexity and vast span of potential needs and opportunities within the product context suggests explorative and research driven activities to develop new technology, product concepts, and insights for medical training applications. To understand the needs and medical conditions in which new products could have an impact, the following sections will describe the development context, tools, approaches, and conditions where these insights could (and will) be gained.

2.2 Development Context and The Fuzzy Front End

New product development (NPD) includes creating products, services, and solutions, by exploiting hitherto unexplored opportunities and user-needs (Otto & Wood, 2001). This requires developing technology, conceptual ideas, and the underlying knowledge for future products (P. A. Koen et al., 2002; Ulrich & Eppinger, 2011). However, the dynamic relations between new problems and numerous potential solutions cause uncertainty. Hence, we often base decisions on a limited, informal, and ambiguous foundation, where various tools and methods are used to learn across different disciplines and knowledge domains (Clarkson & Eckert, 2010; O. Eris & Leifer, 2004; Meinel et al., 2011). While this uncertainty can impede decisions before specifications and requirements are fully understood, the ambiguous conditions also allow for more radical ideas to emerge (Brown & Wyatt, 2010; Kelley, 2001; Leifer & Steinert, 2011). Hence, early development efforts are associated with high risks and significant opportunities.

These preconditions describe the fuzzy front end (FFE) of NPD, and is considered the phase from when we first identify an opportunity until a concept is deemed ready for further development (Kim & Wilemon, 2002; P. A. Koen et al., 2002; Verworn et al., 2008). In the FFE it is critical for organizations, teams, and developers to front-load development efforts, make critical decisions, and learn about unknown future aspects for the eventual launch of new products (M. B. Jensen et al., 2017; Thomke & Fujimoto, 2000). Moreover, the FFE conditions are volatile, uncertain, complex, and ambiguous, and there are apparent risks of making premature or ill-informed decisions (Thomke & Reinertsen, 1998). These can cause costly rework or products failing for not meeting targeted users' requirements or needs (Elverum & Welo, 2015).

Consequently, Schrage, (1993) proposes utilizing prototype-driven specifications as opposed to specification-driven prototyping. This means, prototypes being used to elicit product specifications from empirical insights obtained through development activities (M. B. Jensen et al., 2017; Kriesi et al., 2016; Sutcliffe & Sawyer, 2013). Hence, prototypes are used in various roles and manifestations to inform decisions, communicate ideas, and facilitate learning under the ambiguous and rapidly changing conditions in which development takes place (Camburn et al., 2013; P. A. Koen et al., 2002; Lauff et al., 2018; Menold et al., 2017).

2.3 Prototypes and Prototyping

Prototypes are created and utilized across various disciplines and stages of development, and is considered one of the most critical aspects of NPD (Wall et al., 1992). While the importance of prototypes and prototyping is acknowledged, it is simultaneously one of the least formally explored areas of design (Camburn et al., 2013). *"There is still a lack of knowledge about the fundamental nature of prototypes due to their complex and dynamic nature"* (Lim et al., 2008). The utilization of prototypes and prototyping within professional settings, moreover, suffers few examples in current literature (Elverum & Welo, 2015; J. F. Erichsen et al., 2019). Furthermore, how prototypes are used to gain user feedback and facilitate communication between design teams and stakeholders, needs examples and formal exploration (Menold et al., 2017).

2.3.1 Prototype and Prototyping Definitions

The use of prototypes across different settings, disciplines and stages of development has resulted in divergent definitions of prototypes and their purposes (L. S. Jensen et al., 2016; Lim et al., 2008; Petrakis et al., 2019). While this indicate the vast applicability of prototypes across design practice, they are often considered approximations to evaluate the feasibility or viability of a design (Ulrich & Eppinger, 2011; Wall et al., 1992). However, the generative and creative roles of prototypes in conceptual design phases, have been highlighted in recent research (Camburn et al., 2013; Lauff et al., 2018; Menold et al., 2017). This supports that prototypes can facilitate both reinforcing existing knowledge or generating new knowledge (Lande & Leifer, 2009; Lauff et al., 2018). Furthermore, prototyping is involving and intended for different audiences (Bryan-Kinns & Hamilton, 2002). E.g., gaining feedback from external stakeholders or users (Buchenau & Suri, 2000; Houde & Hill, 1997; Menold et al., 2017), or facilitating communication by acting as boundary objects (Bogers & Horst, 2014; Carlile, 2002; Johnson et al., 2017; Rhinow et al., 2012). Prototypes can benefit internal learning activities by stimulating ideas, mitigate design fixation, and increase designer's self-efficacy (Dow et al., 2011; Lee & Ostwald, 2022; Linsey et al., 2010).

By the roles acting as internal learning tools, external communicational artefacts, or gauging features of eventual products, this thesis considers a broad definition that prototypes are tangible outcome of prototyping activities to answer questions (Schrage, 1993). Thus, the characteristics and attributes of prototypes are inherently dependent on the question that needs answering, which evolve throughout development projects. This suggests that any object, both physical and digital, and generative and analytical, could serve as a prototype if it informs aspects of the faced design question to the developers (Buchenau & Suri, 2000; Houde & Hill, 1997; Lim et al., 2008). Subsequently, prototyping is considered both cognitive and physical activities for learning, where designing, building, and testing can generate, transfer, and capture knowledge that gets embodied by the tangible output (i.e., prototypes) (J. F. Erichsen et al., 2019).

2.3.2 Prototype Resolution, Abstraction, and Fidelity

To leverage prototypes purposefully for answering open design questions, we need terminology describing their characteristics and intentions. This thesis describe prototypes by manifesting a design via abstraction, resolution, and fidelity considerations (Bryan-Kinns & Hamilton, 2002; Edelman & Currano, 2011). Furthermore, what prototypes represent and thus their use in design activities are considered by what they filter, and more specifically by their role, look and feel and implementation aspects (Houde & Hill, 1997; Lim et al., 2008). These prototype attributes, dimensions, and their relations are indicated in Figure 2.1. Resolution is the amount of detail a represented by a prototype. Abstraction considers the closeness to what is known, or completeness of a prototype. High fidelity is regarded as the closeness to an eventual design and is therefore affected by both the resolution and abstraction levels. The filtering dimensions relates to the intention of the prototype, and what aspect of the eventual product that is being prototyped (Edelman & Currano, 2011; Houde & Hill, 1997; Lim et al., 2008).

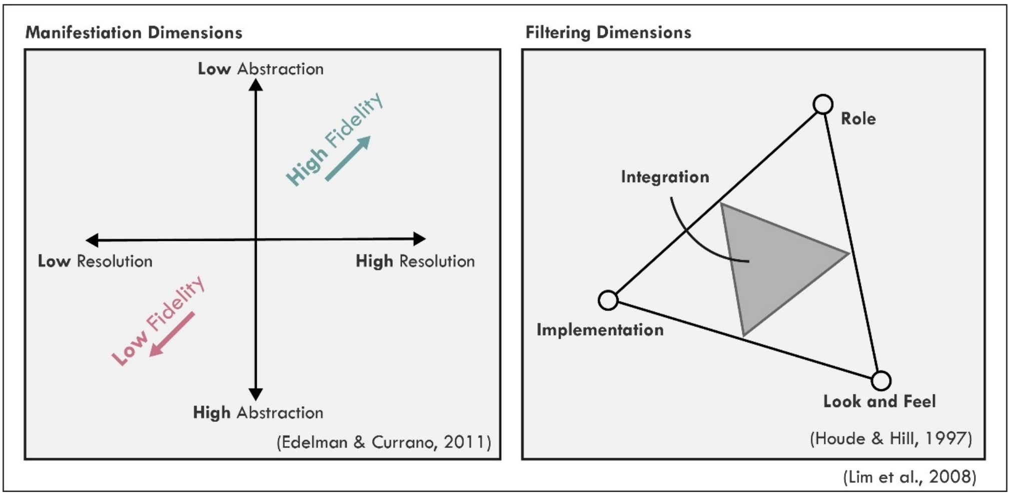


Figure 2.1 Prototype resolution, abstraction, and fidelity. Adapted and combined models from (Edelman & Currano, 2011; Houde & Hill, 1997; Lim et al., 2008).

2.3.3 How do we Decide the Required Prototype?

While prototypes and prototyping show vast applicability in conceptual design stages (Krishnakumar et al., 2022), deciding what a prototype should entail and how it could contribute to new insights being gained, can be challenging (Petrakis et al., 2021). To support deciding and planning prototypes several frameworks and suggestions exists. For example, Menold et al., (2017) suggests prototypes to be framed, built and tested throughout different stages of design, as a framework for assuring strategic prototyping is performed. Moreover as prototypes evolve throughout projects, visualizing this process could inform strategic planning as suggested by Hansen & Özkil, (2020). However, as will be described in the next section (and onwards in this thesis), causality and dynamic conditions makes prototyping especially challenging to plan for. This concerns the ambiguity of FFE, and how insights gained from prototypes, could reveal new obstacles and opportunities to address and pursue in projects.

2.4 Development Approach and TrollLABS mindset

An issue with formal and linear development models, that have raised the quest for more dynamic and agile process models, are the misconception of immutable truths and that pre-defining requirements and specifications ultimately mitigate risk. However, this might not be the case in the FFE, where intertwined problems, complexity, and causal ambiguity make knowledge contextual and dynamic as problems evolve and product requirements change (Clarkson & Eckert, 2010; Leifer & Steinert, 2011; Takeuchi & Nonaka, 1986).

2.4.1 TrollLABS Thinking and Mindset

As conditions change when new insights are obtained from prototyping activities, subsequently project plans ought to change or become obsolete. As we in the FFE experience vast uncertainty before insights are gained, it is therefore challenging to plan, rationalize, and generate informed design ideas. This describes the narrative for the mindset and development approach (wayfaring) advocated at TrollLABS. This starts with a bias towards action, and often performing rapid and uninformed development actions to gain preliminary insights. Opposed to thinking before doing, this means creating prototypes to stimulate creative thinking, and deploying a reflection in action mindset to learn (Schön, 2017). This behaviour, generating insights to progress development, is moreover supported by doing divergent and convergent design iterations (Ö. Eris, 2003). This means generating multiple alternatives to learn about the conditions, before deploying analytical and convergent logic to decide on the next steps.

2.4.2 Unknowns and Serendipity

As we obtain new insights from designing, building, and testing ideas, we uncover unknowns. Unknowns are obstacles and subsequent opportunities that can impede or pivot further development. By prototyping, we can elicit and make these tangible, and we distinguish between unknown knowns, known unknowns and the unknown unknowns (M. B. Jensen et al., 2017; Kriesi et al., 2016; Sutcliffe & Sawyer, 2013). The unknown knowns are knowledge we possess but were unaware of their relation to the current design challenge. Known unknowns on the other hand, are knowledge we don't possess, but are aware that we need to obtain it. Lastly unknown unknowns concern problems we are unaware of and lack the required knowledge to solve. Subsequently, unknowns are in projects obstacles and opportunities that needs elicitation before they can be addressed. As we are unaware of their existence before they are revealed, serendipity is both required and requested. In terms of prototyping activities this means, maximizing the chances of finding the critical unknowns for the given design challenge, and provide the flexibility and required toolbox to pursue and address them.

2.4.3 Wayfaring

As described, strict and formal development processes have limited affordance in the FFE, and thus this thesis relies on the use of wayfaring as development management approach. Providing flexibility and broad coverage of problem and solution space (Leimeister et al., 2021), wayfaring is suggested to accommodate complex and ambiguous problems in the FFE. Wayfaring, originating from the Hunter-Gatherer Model (Steinert & Leifer, 2012), allows project execution to follow nonlinear pathways towards eventual product concepts (Gerstenberg et al., 2015). This is facilitated by using sequential and iterative prototyping sprints, called probes. Probing enables stepwise progress of development, revealing unknown unknowns, and navigation using critical functionality towards an informed (potentially radical) idea (Kriesi et al., 2016). Each probe enlightens the design space by performing divergent (360 scan), and convergent (deciding direction) prototyping activities. Using wayfaring, rapid development and reflection are accommodated by utilizing adapted scrum principles (Atzberger & Paetzold, 2019; Schwaber & Sutherland, 2011). Timeboxing design activities (sprints) and regular interactions to share and reflect are purposeful. Furthermore, wayfaring is an opportunistic development approach, thus adapting different tools and methods to accommodate unique, complex, and ambiguous problems. For example, multidisciplinary integration, agile, and human centred design.

2.4.4 Problem Focus, Human Centred, and Multidisciplinary

Wayfaring provide flexibility for purposeful and rapid development activities to be performed in projects. Hence, the approach is driven and governed by the individual problems that emerge and needs solving (Leifer & Steinert, 2011). This often requires development teams to deploy a variety of tools and methods, ranging across different disciplines and knowledge domains (Gerstenberg et al., 2015). This multidisciplinary integration is critical to perform early and throughout projects. Problem focus and addressing user-needs is furthermore, proposed by design thinking that underlines that success of eventual products require wide and willing adoption by users and stakeholders (O. Eris & Leifer, 2004). Consequently, as NPD considers unique and evolving problems, wayfaring (as design thinking), suggests different thinking spaces, inspiration (problem), ideation (solutions), conceptualization (implementation) that the project moves through (Brown & Wyatt, 2010). Yet, these spaces are not addressed sequentially (Meinel et al., 2011). Hence, wayfaring is often chaotic when reviewed in retrospect, and fractal connections between different modalities of knowledge can occur as new prototypes are created. However, wayfaring being driven by prototypes, following the evolution of prototypes and their embodied knowledge, different intents and directional shifts are often easily visualized.

2.5 Organizational Knowledge and Triple Loop Learning

Learning, communicating, and informing decisions by prototyping concerns knowledge being generated, shared, or affirmed in NPD projects (Camburn et al., 2013; J. A. B. Erichsen et al., 2016; O. Eris & Leifer, 2004; M. B. Jensen et al., 2017; Lauff et al., 2018). Different prototype affordances (fidelity levels) are also shown to contribute to different knowledge dimensions (Real et al., 2022). This knowledge is furthermore, simultaneously generated at different organizational levels, both on a structure and formal process level, and informal designer level where the prototyping is being performed (Bogers & Horst, 2014). This suggests that prototyping can generate different kinds of knowledge, and that it is generated both within and across ontological planes of the organization. While not rooted in knowledge management, this thesis will utilize the following concepts and terms from organizational learning to describe empirical observations and insights from prototyping (Nonaka, 1994).

2.5.1 Tacit and Explicit Knowledge

Nonaka, (1994), describes types of knowledge as either tacit, being skills, experiences, and know-how of individuals. Or explicit, being articulated and unambiguous, and thus more easily conveyed and stored and can exist on both individual and organizational levels. Prototypes exerts both explicit (i.e., articulated and formal) and tacit (i.e., unarticulated and informal) attributes, and knowledge generated from prototypes contributes to both project insights, process insights, and organizational learning (J. A. B. Erichsen et al., 2016).

While knowledge is generated by individuals, the SECI-model, seen in Figure 2.2 proposed by Nonaka, (1994), describes how tacit and explicit knowledge, could be generated, shared, and amplified. This furthermore describes the different conversions of knowledge types, through socialization (tacit-tacit), externalization (tacit-explicit), combination (explicit-explicit, not considered by this thesis), and internalization (explicit-tacit). Schrage, (1993) states that for organizations to create better products, they must first learn to create better prototypes. It is therefore of interest to explore how prototypes facilitate knowledge being generated, shared, and re-used, both within and across development cultures, to enhance their prototyping and development capabilities (J. A. B. Erichsen et al., 2016; Leifer & Steinert, 2011; Nonaka, 1994).

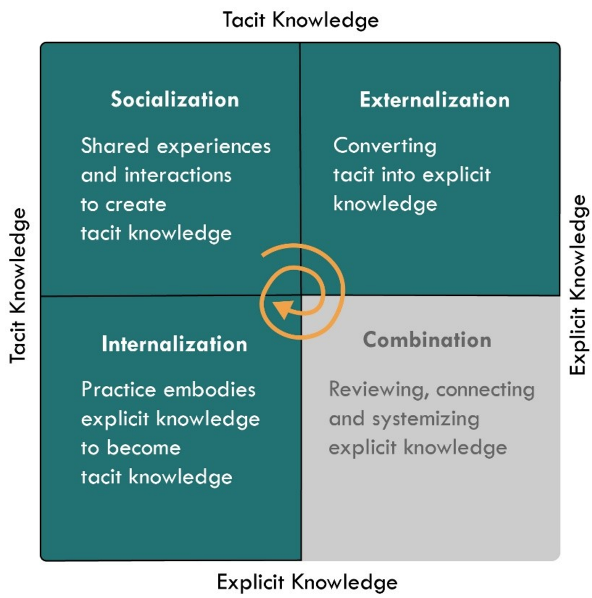


Figure 2.2 The SECI-model, adapted from Nonaka, (1994).

2.5.2 Triple Loop Learning

Prototyping broadly covers all operational levels (ontological planes)(Nonaka, 1994; Wall et al., 1992), and simultaneously both tacit and explicit knowledge is generated (Bogers & Horst, 2014). For looking at the holistic benefits, implications, and opportunities for prototyping in industrial contexts, this thesis will by structure and logic follow the concept of triple loop learning (Argyris & Schön, 1997; Tosey et al., 2012), as described by (O. Eris & Leifer, 2004; Leifer & Steinert, 2011). This can be seen in Figure 2.3, which visualize the operational, tactical, and strategic learning loops, and the relevant actors, activities, and their relations within these loops.

Loop 1 is the operational level; this is considering practitioners in teams applying prototyping in projects. This will be presented in chapter 4 by the aim of answering *what* we do in prototype-driven development projects. **Loop 2** takes place in the informal space of product development cultures, and concerns process improvements and interpretation. This will be presented in chapter 5 as the operationalization of prototype-driven development in a wayfaring framework. It attempts to inform *how we think* and how wayfaring could support NPD management in the FFE. **Loop 3** consists of the capture, transfer, and re-use of knowledge in development cultures. It stretches beyond the product development culture into the formal organizational structures. In chapter 6 a review and discussion of tacit and explicit knowledge generation, transfer and capture will be provided. Furthermore, this section drives a discussion on strategic use of wayfaring and how it has been attempted implemented.

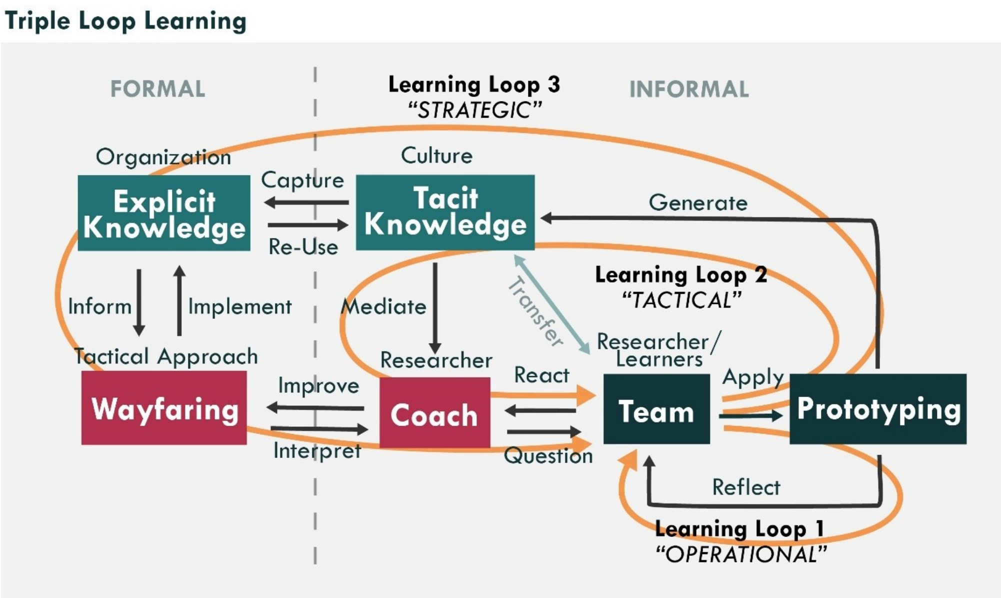


Figure 2.3 Triple loop learning. Loop 1 considers operational and design activities. Loop 2 considers tactical and informal process knowledge. Loop 3 considers organizational knowledge and strategic discussions. Adapted from (O. Eris & Leifer, 2004) and (Leifer & Steinert, 2011).

3

Cases:

Projects and Activities

This chapter describes the project and activity cases carried out during this PhD project. The cases are focused on two separate yet interconnected themes. Cases 1-5 and 10 are development projects where the author has either had a direct role as the primary developer or contributed to and coached the projects. The projects have aimed at improving medical training, by simulating various physiological aspects of a human patient. Cases 6-9 concern strategic and managerial activities to support prototyping, both within the context of TrollLABS, Laerdal Medical, and joint efforts between the two. Figure 3.1 highlights the anatomical areas relevant to the cases.

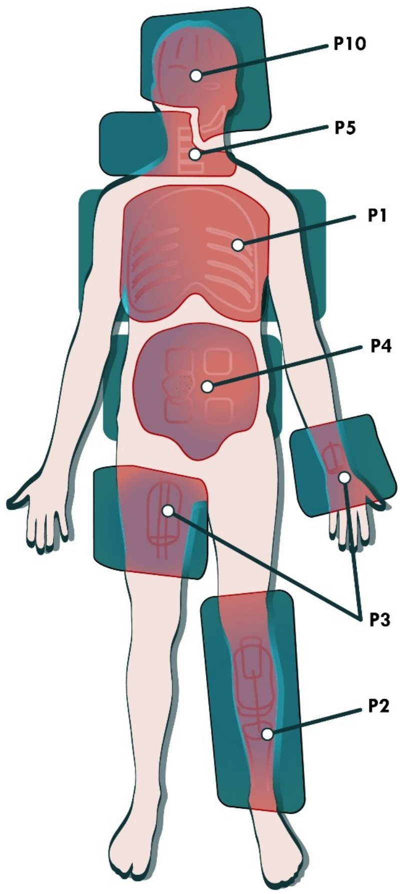
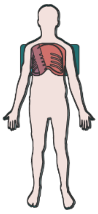


Figure 3.1 Overview of the case projects and their relation to areas of human physiology.

P1 Chest for CPR Training



Cardiopulmonary resuscitation (CPR) training can use mannequins that enable the practice of chest compressions and artificial ventilation. This case concerns a master's project, carried out by the author in 2018. The project was initiated to explore the needs and opportunities for improving CPR mannequins and, thus, the training of healthcare personnel. The project suggested improving the realism of current CPR mannequins by addressing characteristics not sufficiently replicating the experience of compressing a human chest. Over 80 prototypes were built and tested by internal experiments and external user tests. Progressive stiffness, viscoelastic recoil, and considerable variance between individuals was critical functionality the project attempted to incorporate. The project found that CPR also compromises the mechanical integrity of the chest and that reduced spring response and tactile feedback of ribs fracturing could affect CPR performance. As a result, a compliant polymer spring that emulates the shape and mechanical deformation (dynamic) of the human chest and a mechanism simulating the tactile experience of rib braking got developed. This case is further explained in C1 Auflem et al., (2019).

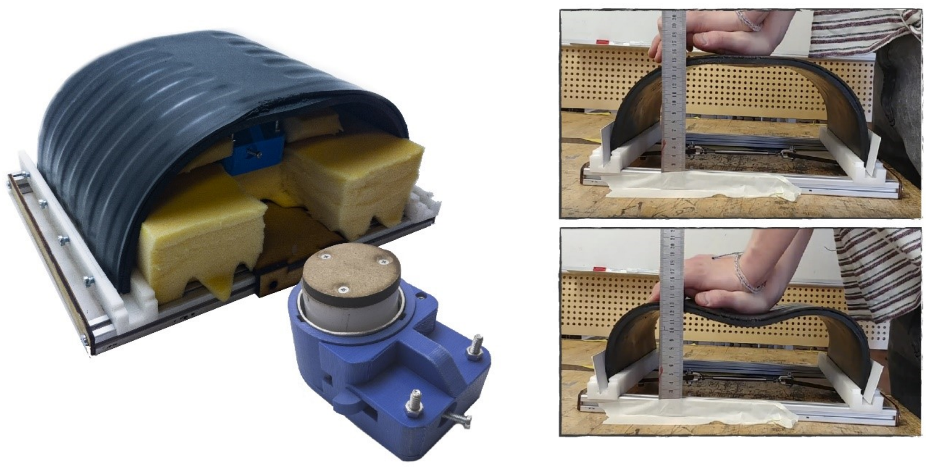
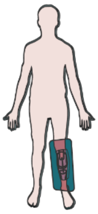


Figure 3.2 P1 Prototype of a chest with rib fracturing haptics device for CPR training. On the right demonstration and testing of deforming the chest structure by compressions is shown.

P2 Break-a-Leg



Open leg-fracture re-alignment (reduction) and stabilization are challenging to perform and require hands-on experience to master. This case describes a course project from TMM4245 Fuzzy Front End in 2018, initiated to explore the need for equipment to train fracture reduction procedures. The author assumed a coaching role in this project and had weekly scrum meetings as well as informal discussions with the team when appropriate and needed. The team used physical prototypes to enable users to describe how to perform the correct procedure. Furthermore, this revealed the need for realistic tactility and functionality to perform the simulated procedure adequately. The team developed a functional prototype, as seen in Figure 3.3, with adjustable and modular traits allowing the mechanical behaviour of the simulator to be configured according to different training scenarios. The simulator was tested and highly regarded as different fractures results in different tactile experiences. Thus, varying in difficulty and subsequently by required skills to perform the procedure sufficiently. This project is also utilized as a case in C1 (Auflem et al., 2019).

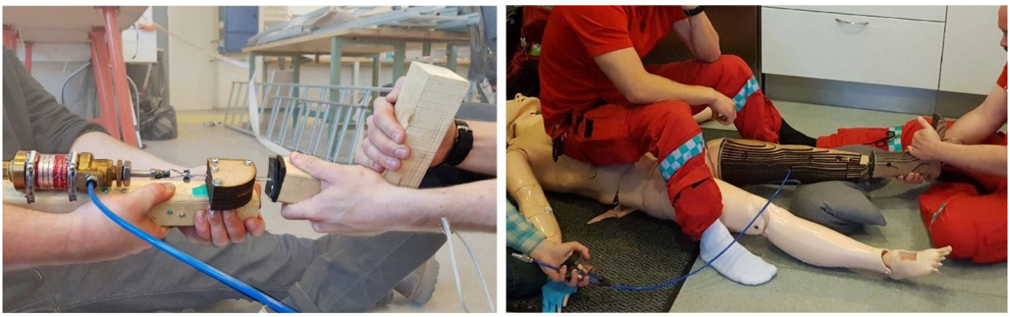
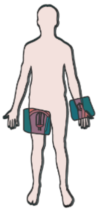


Figure 3.3 P2 Prototypes of a fractured leg task trainer. On the left, functional prototype using pneumatic cylinder and simple topology. On the right, a higher fidelity prototype (using the pneumatic system), integrated with an existing simulator for paramedics to test.

P3 Vascular Phantom(s)

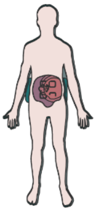


This case describes two projects concerning simulating artefacts of human vascular physiology. The first project aimed to create a vascular access task trainer for a Resuscitative Endovascular Balloon Occlusion of the Aorta (REBOA) procedure in 2018. The author had a contributing role in the explorative phase of this project. The second project, from 2022, entailed creating an arterial wrist phantom with sensing capabilities to investigate mechanical diameter waveforms caused by a simulated pulse. In this project the author contributed to both the explorative prototyping and conceptualizing of the technology.

The first project required prototyping towards the tactility of canulating a human groin. Furthermore, any solution was required to accommodate modern ultrasound imaging equipment, allowing a needle to be guided and placed correctly. This resulted in a viable prototype enabling training of users performing the procedure and a novel technology for ultrasound phantoms C10 (Steffensen, Kriesi, et al., 2022). The second project developed a novel concept of weaving sensors embedded in flexible elastomers to measure the diametric change of simulated arteries. This technology was utilized in a wrist phantom prototype to enable research and calibration of wearable sensor applications by sensing mechanical motion from pulse waveforms C8 (Steffensen, Auflem, et al., 2022).

Figure 3.4 P3 Vascular phantom prototypes. From left an ultrasound image of the simulated artery and vein, medical professional using handheld ultrasound probe guiding a needle insertion, vascular wrist phantom, and the braided sensor.

P4 Palpation Trainer(s)



Palpation is the task of performing diagnostics routines using hands and fingers to feel (palpate) for areas of tenderness, pain, or tactile abnormalities on the patient's body. In current medical training and curricula, palpation is a skill that requires repeated practice and dexterous hands-on experience. A challenge with palpation is the highly subjective and experience-based skill that needs to be practised, often relying on actual patients. In many situations, this can be dangerous and uncomfortable for the patient. This case entails two master's projects, coached by the author, spanning two semesters: the development of a palpation task trainer in 2020, and the development of a novel concept and technology for haptic palpation simulators in 2021.

In the first project, different prototypes got developed to calibrate a task trainer's required fidelity and functionality. This task trainer was tested to ensure an ecologically valid and usable solution described in C6 (Ege et al., 2021). In the second project, a novel technology using principles from soft robotics got developed. The concept was used to investigate the fundamentals of palpation interactions in healthcare and training through a semi-structured experiment to assess the tactile perception of novice users. Based on ferromagnetic granules and granular jamming, the technology resulted in a haptic interface that could render palpable objects of different shapes and hardness presented in C7 (Rørvik et al., 2021).

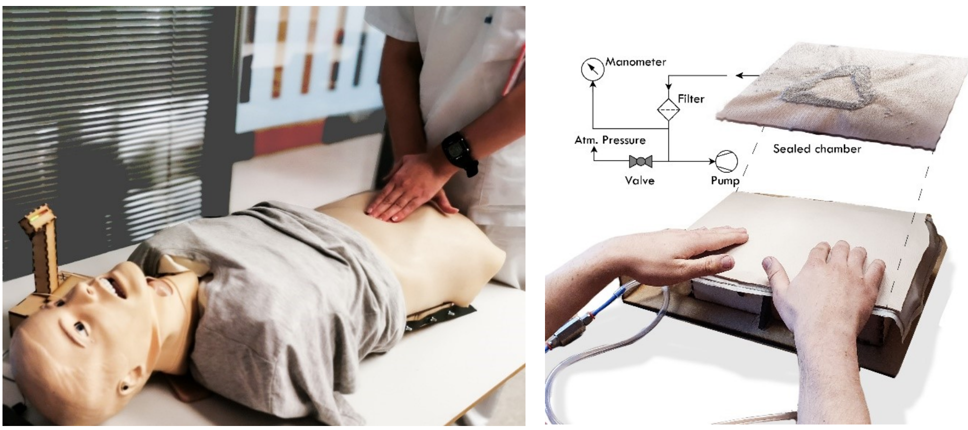
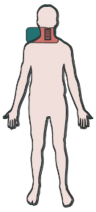


Figure 3.5 P4 Prototypes showing concepts for palpation training. On the left a novel palpation-trainer integrated in a mannequin torso. On the right the ferrogranular palpation interface.

P5 A Pain in the Neck



This case considers a master's project, carried out in 2022, to explore new neck concepts for patient simulators. The project was coached and supported by the author. Through various consultations with a sample of relevant stakeholders, it got pointed out that a realistic range of motion and (lack of) objective feedback is a current shortcoming, leading to standardized patients being the preferred alternative in many training scenarios. Furthermore, the importance of training for suspected neck injuries and pre-hospital trauma cases was emphasized.

A new neck concept with variable compliance, enhanced mobility and sensors measuring movement was developed and proposed. A prototype was piloted in an experiment with different users in a simulated scenario. The concept and results from the pilot experiment suggest the concept to be viable in simulation and experiments to research best practice procedures regarding spinal immobilization of suspected neck injuries. This experiment showcased how different techniques of patient transfer and immobilization yield different qualities in mitigating head movement. The new neck was suggested as a standardized tool that could facilitate further research to improve medical training and inform the best medical procedures practice regarding suspected neck injuries, presented in C11 (Henriksen et al., 2023).

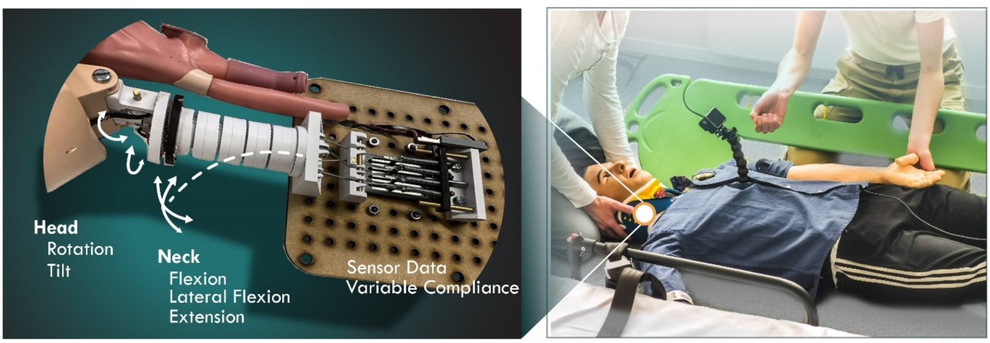


Figure 3.6 P5 Prototype of the neck concept. The testing setup with the concept integrated in a modified simulator shown on the right.

P6 TrollLABS Medical

In November 2018, a new lab environment was opened at the faculty of medicine and health sciences at NTNU. This lab is a co-location of a student-driven innovation organization called DRIV and medical project initiatives from TrollLABS. The physical space, as seen in Figure 3.7, is located near students, clinicians, and researchers on the hospital grounds. The author, and colleagues from TrollLABS, were involved in the creation of the space and especially in building up the workshop area and equipping it with tools and machinery. Of interest to this thesis is creating a multipurpose arena and how it contributes to interactions and collaboration with medical students and professionals as critical users in development projects. This space has thus been important in running experiments, simulating scenarios, and interacting with stakeholders apart from being a workshop for building prototypes.

Figure 3.7 P6 TrollLABS Medical facility. Top image showing before and after the lab being set up. Below, a separated workshop area equipped with workbench, laser cutter, 3D printer, and various prototyping tools and equipment.

P7 Joint Workshops and Sharing Sessions

An essential aspect of the collaborative efforts between Laerdal Medical and TrollLABS has been the regular and mutual exchange of development experience, mainly facilitated by two forms of physical gatherings. Firstly, workshop activities with Laerdal have been organized to bring engineers, designers, and developers from the organization to a new physical location and partake in development activities (considering problems outside of the regular scope of the organization). The second mode of physical interaction has been to bring and present projects from TrollLABS to Laerdal Medical. These interactions are structured as informal show-and-tell sessions, where encountered problems, new technology, and user interactions are showcased by presenting prototypes. The general idea has been to share and discuss tools and methods for prototyping in early project stages and to allow experimentation and utilization of new tools and methods in a new context and environment.

Figure 3.8 P7 project presentation and prototype showcase at the Laerdal office.

P8 Intern Technology Projects

Intern projects, where 5-6 students have worked on a dedicated task, have been carried out each year, during summer break, throughout this PhD. Typical for these projects is that they are provided with a dedicated space, given an open problem formulation, prototype-driven focus, and utilized simplified scrum principles. The development process has been guided by the project's progress, meaning that requirements and objectives are dynamic in the sense that they are uncovered and formalized throughout the project. Hence the outcome at the end of the projects varies. Results have ranged from proposed conceptual models handed over after testing with users, to a portfolio of technologies relevant for the organization to investigate further. Since these are unique projects with different tasks and participants, the case will consider the structure, development approach, findings, and organizational benefits of running "skunk-works" development initiatives in-house. The author has coached these projects with technology and managerial leads from Laerdal Medical. This has involved having daily scrum meetings, weekly presentations, and alignment and integration of projects within the Laerdal organization.

Table 3.1 Technology projects overview.

Project Year	Challenge	Findings and Results
2018	Energy Harvest	Concept for electromechanical generative braking of chest compression recoil.
2019	Technologies for active eyes	Range of prototypes showcasing various technologies to enhance realism in simulators.
2021	Skin texture and colour	Concepts for altering texture and colour in flexible materials, for simulated symptoms.
2022	Effort of breathing	Novel actuation concept to simulate breathing distress in a paediatric simulator.

P9 Product Development Sprints

The development sprints were (and still are) a concrete initiative to enhance the development activities Laerdal Medical performs in the fuzzy front end of new projects. A two-week development sprint was carried out in one of the development teams in the organization in January 2020. As two product opportunities had been identified, the team got split into two groups consisting of three developers each. The sprint aimed to formulate focused design questions, generate fast prototypes to gain relevant insights, and to test and interact with key users. In short, the development sprint followed simplified scrum principles starting each day with a stand-up meeting presenting current design questions and plans for the coming day. The author had a facilitating role, and together with a team manager and an additional facilitator meetings and project execution was managed. The sprint executed two phases from the Laerdal development model by working iteratively and prototyping under time constraints. The case and findings from the project sprint are further explained in the internal paper shared within the Laerdal organization i1 in the appendices.

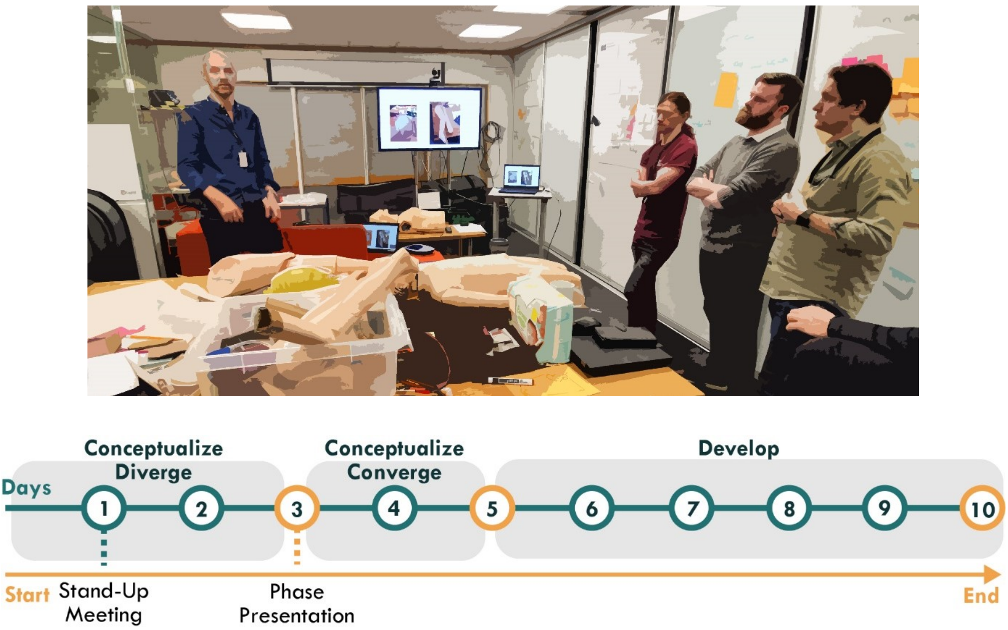
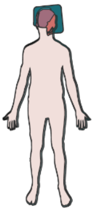


Figure 3.9 P9 Stand-up meeting and timeline of two-week development sprint.

P10 Expressive Simulator Face



This project set out to explore needs and concepts to enable nonverbal communication capabilities in patient simulators. Hence, understanding what aspects of nonverbal and social cues could facilitate meaningful interactions with simulated patients in healthcare training contexts. While the project's scope started broad, it quickly focused on investigating the face of simulators, currently static and non-responsive. Healthcare professionals supported this by stating that in many cases, they could recognize a patient not being well by entering the room, looking at the face, and assessing behaviour and expressive responses.

While robots with facial expression capabilities are not new, mechanisms, technology, and concepts considered in this context would need to accommodate the use cases of patient simulators. Especially considering both clinical functionalities bound to simulated physiological artefacts and the need for tactile interactions and manual interventions. The author has been the main contributor to this project.

Figure 3.10 P10 detailed view of the simulator face. Demo setup connected to a Laerdal mannequin. Image adapted and used with permission (Tolo, 2022).

The prototype-driven project resulted in several conceptual prototypes, one of which is seen in Figure 3.10. The proposed concept maintained the existing jaw, eye, and airway functionality and allowed the skin to be interchangeable, thus altering the appearance and representing different patients. By the functionality required to realize this concept, several novel findings and technologies were generated, e.g., a new composite material for emulated skin, novel actuation principle(s) Figure 3.11, and AI-assisted control and design evaluation as presented in C3 (Auflem et al., 2022) Figure 3.12. Furthermore, a spin-off project, also seen in this figure, investigated using facial responses to calibrate and provide feedback on cognitive assessment of simulated patients.

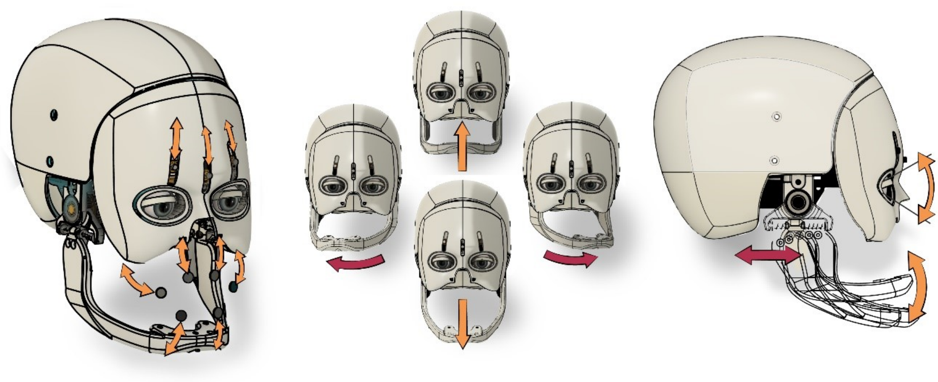


Figure 3.11 The mechanical design of the simulator face concept. Modes of actuation (orange) and passive compliance (red) are indicated with arrows.

Findings and insights from this project have been shared within Laerdal as —demos, a video and interview in 2022, and a presentation held at the Laerdal Medical Data Summit in Copenhagen in October 2022. Furthermore, this has resulted in several potential continuations and integration opportunities for this project, that could be applied in different areas of the Laerdal product portfolio.

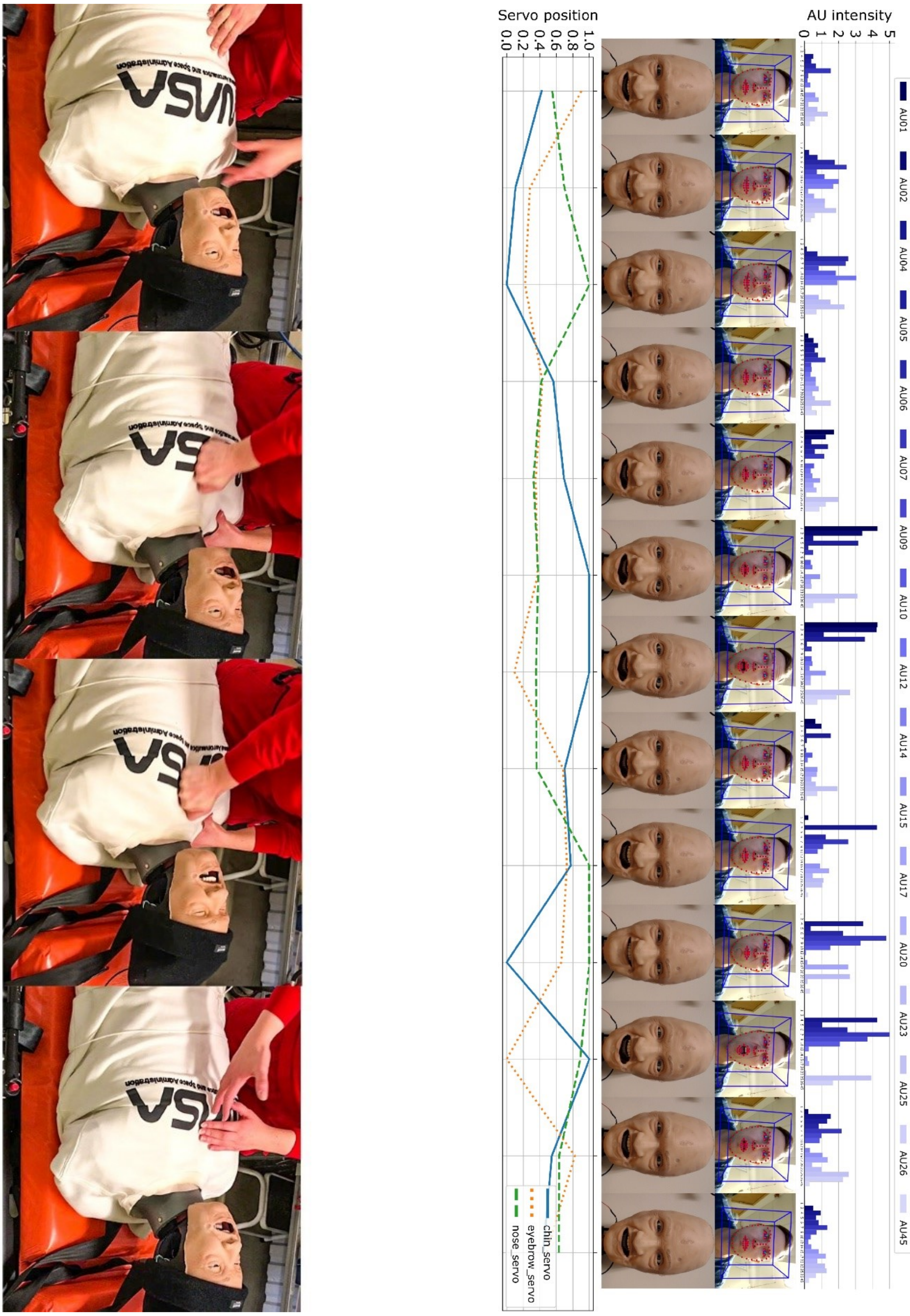
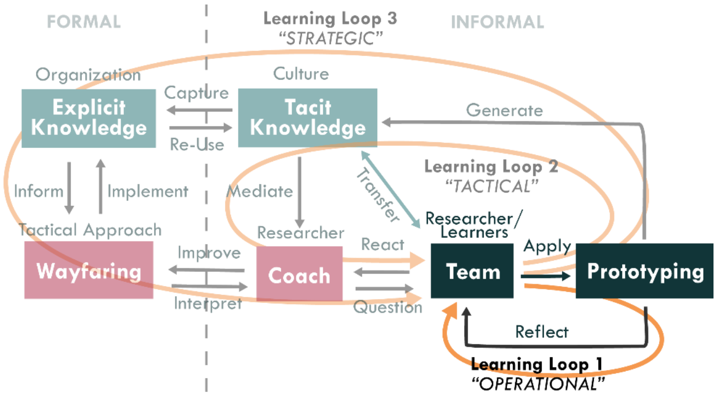


Figure 3.12 Experiments using the expressive robot face in P10

Right showcasing real-time re-enactment using FACS coding and automatic landmark detection and tracking (Auflem et al., 2022). Left shows cognitive assessment and pain stimuli using face expressions.



4

What we Do

in Prototype-Driven Development of Medical Training Equipment

This chapter accounts for the first learning loop of this thesis and will, from a development practitioner's perspective, provide insights and considerations for prototype-driven development in the fuzzy front end. Specifically, how designing, building, and testing prototypes in different contexts yield different insights to progress and inform further development of medical training equipment. Prototyping has, in the project cases, been a powerful asset in the development to obtain fast, deep, and tangible insights. Hence, the timeframe is immediate, considering development actions, reactions, and reflections.

The chapter aims to provide generalizable development approaches identified in the project workflows and describe the prototype and prototyping characteristics accompanying the different approaches and modes of thinking. More specifically, it will describe prototyping to explore problems, explore solutions, develop concepts, and pilot conceptual prototypes in experiments and user tests. Findings and project output across the various case projects will be presented through the lens of potential for improving medical simulation and training. Moreover, examples from the case projects will be used to provide insights into what prototype-driven development of medical training equipment entails in the earliest stages of development.

4.1 Thinking Spaces in the Fuzzy Front End

The Laerdal Development model divides development into phases to ensure that activities align with user needs and that these are made to deepen our understanding. The initial phases of this model, *explore* and *conceptualize*, is particularly important as it covers the fuzzy front end (FFE) of development, from identifying opportunities to creating a product concept. Rather than being rigid gates, these phases should act as development spaces, ensuring a flexible process that allows for quick iterations, as encouraged by design thinking (Brown & Wyatt, 2010; Leifer & Steinert, 2011; Meinel et al., 2011). Thus, the sections of this chapter will outline identifying opportunities, aligning with users through inspiration, generating and evaluating ideas through ideation, and establishing and iterating system prototypes through conceptualization. However, these development spaces and methods are not necessarily linear nor sequential, and prototyping will be highlighted as the main driver of the development. Finally, experiments and user tests will be shown to help generate insights for informing decisions for further development.

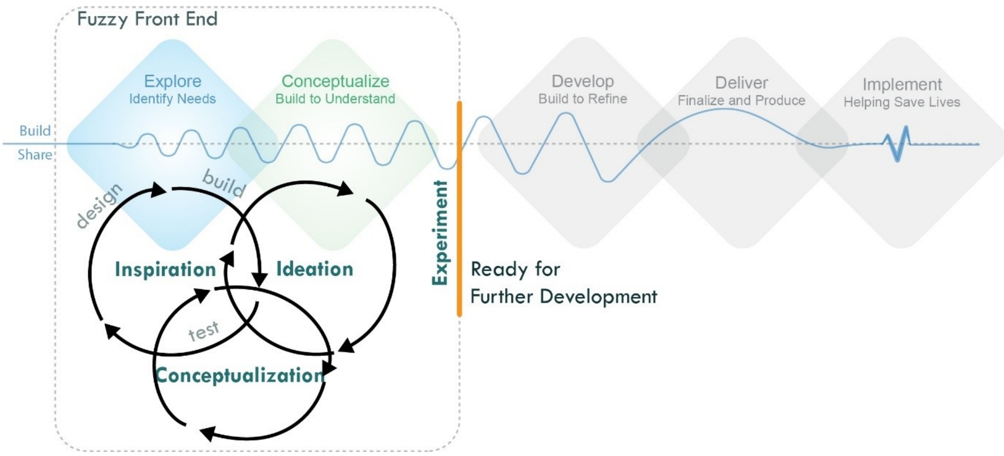


Figure 4.1 Prototype-driven development in the fuzzy front end contextualized in the Laerdal Development Model.

4.2 Identifying Opportunities

For new medical training equipment to help improve trainees' performance, skillsets, and knowledge, insights on various domains are required. Thus, to gain insights rapidly, design practitioners deploy different tools and actions in the fuzzy front end to learn. These entail actions to develop a new product opportunity and create concepts addressing the various needs for eventual products. Prototyping is at the core of these activities, and need-finding and prototype interaction contribute insights for generating concepts through prototype-driven development and assessing conceptual prototypes.

In the project cases, opportunities have been identified as either *role*, *look and feel*, or *implementation* issues or needs. The categories represent the spatial model for aspects of interactive product prototypes proposed by Houde & Hill, (1997). Opportunities identified as lack of available or adequate equipment are considered by role, and more specifically, opportunities for products to enable or enhance the training of medical tasks (considered by P2 and P4). The look and feel relate to a product's visual and tangible attributes. Subsequently, opportunities concern clinical realism and products sufficiently approximating human physiology, often derived from a gap between training and clinical experiences (considered by P5 and P1). Lastly, the implementation relates to technical and engineering opportunities for new products. These opportunities emerge from technology and methods that enable novel concepts to be created or that can enhance existing products (considered by P3 and P10).

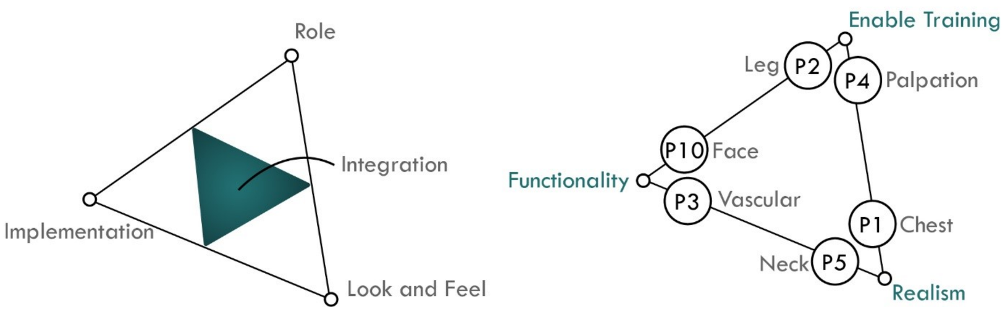


Figure 4.2 Triangle reframed for the development of medical training equipment. Model from the aspects of what prototypes prototype recreated from Houde & Hill, (1997) on the left, and contextualized to medical training equipment and opportunities elicited in the case projects on the right.

4.3 Exploring Opportunities and Problems – Inspiration

Inspiration is the prototyping activities deployed to understand and rethink the problem(s) and spark ideas for new solution and products. In the FFE, product opportunities are not always explicit and often require elicitation and exploration for conceptual ideas to emerge (P. Koen et al., 2001). This involves understanding the various use cases, users, critical functionality, and medical scenarios. Early user identification, and engagement, to obtain insights on the needs and implications for new products to succeed is therefore paramount. The goal of the initial problem exploration is, thus, to translate abstract training needs, tacit knowledge, and functional traits into design problems (and objectives).

4.3.1 Users, Stakeholders, and Extreme Users in Medical Sector

Understanding user needs is critical for solutions to be adopted and to cause positive behavioural changes (Leifer & Steinert, 2011). This is emphasized as human-centred design and is a critical element advocated in design thinking (Meinel et al., 2011). Notably, in our context, the users may not be the ones we first envision, or the curricula and guidelines differ from the initial understanding of the medical context. Therefore, exploring and deciding for whom and what the training should entail is required. Hence, designers ought to understand the ecosystem of clinical practice and the training and educational paradigms the development efforts could target.

There are many user segments of medical training equipment ranging from novice students to well-versed practitioners. For example, students in healthcare education attending low-dose high-frequency skill training, to experienced healthcare personnel attending routine practice, certification, or team simulation events. The latter can take the role of extreme users and are a crucial asset for designers to provide context and tacit knowledge regarding real patient encounters and medical procedures. Experienced healthcare practitioners (EHP) are, thus, essential in the earliest stages of development because of their clinical experience and knowledge about current training and educational paradigms.

In the project cases, various professions and occupations from the healthcare workforce have acted as extreme users. E.g., projects P1 and P3 interacted with anaesthesiologists and physicians with experience from air medical services, thus, benefitting having prior encounters with trauma scenes and patients in conditions relevant to the projects. User segments that have contributed to the various projects are paramedics (P2, P5), nurses (P1, P4), general practitioners (P4), rescue services (P1, P2, P3), anaesthesiologists (P1, P10), and neurologist (P10). Contributors to the various cases have been chosen based on availability, stage of development, medical context, simulation context, and relevant clinical experience.

Insight on Users: Finding and understanding the needs for new solutions can benefit extreme users —in context of this thesis being experienced healthcare practitioners (EHP). These can provide context and relate problems (and opportunities) to clinical practice, training and educational paradigms, and personal experience from real patient encounters. This insight was gained from interacting with EHP in the project cases.

4.3.2 Prototyping to Communicate and Gain User Insights

To explore opportunities to improve, or enable, training in medical procedures, several aspects of the interaction between the trainee and simulated patient (or equipment) need to be understood:

- What medical procedures are the EHP performing, and what do they entail? These could be cognitive, psychomotor, or intrapersonal aspects.
- What are the critical functions, as in simulated features and physiological attributes, required to train in the medical scenario?
- What are (if existent) the training routine's learning objectives and success criteria proposed by current curricula?

Insights are needed to explore how these questions translate between the medical and engineering domains. Hence, they are not easily answered and require efforts to make them tangible and understandable through meaningful interactions. Prototyping has in the projects facilitated relevant insights by extreme users (EHP) interacting with physical artefacts. These artefacts have enabled communication by acting as boundary objects and tapping into EHP's experience and tacit knowledge (Buchenau & Suri, 2000; Rhinow et al., 2012). Users, especially EHP, are scarce resources, and it is essential to maximize the learning output from each interaction. Prototypes considered for these use cases are rapid and crude models, where the focus is on gaining fast feedback and learning rather than assessing the attributes of the artefact itself. Hence, these prototypes address specific assumptions about the problem without paying significant attention to what an eventual solution would entail.

Exemplified by project P1, the tactility of chest compressions was explored, utilizing two prototypes of a simplified chest cross-section. EHP compared real CPR experiences with the characteristics of the prototypes, which enabled tacit knowledge from medical experiences to translate into engineering terminology for the designers. Prototypes established a common ground and enabled a common vocabulary by describing and decoding using analogies and prototypes, thus bridging the disciplinary gap C1 (Auflem et al., 2019). For demonstrations and showcasing clinical equipment, P2 deployed a simple contraption (using wood beams and springs) strapped to a team member. This enabled paramedics to perform and describe the complete procedure of repositioning an open leg fracture C1 (Auflem et al., 2019). In P4, a container with water-filled bladders was covered by a rubber sheet, which enabled a general practitioner and medical students to perform palpation exercises and routines C4 (Ege et al., 2020). In P3, a simple silicone block encasing two tubes of water enabled the use of both hand-held ultrasound imaging equipment and large gauge needles to showcase the procedure of ultrasound-guided vascular access. In all these cases, simple artefacts facilitated descriptive demonstrations and brought special attention to critical aspects of the experience that would need to be included or captured realistically by eventual products.

4.3.3 Learning by Doing

Ultimately, the problem exploration should elicit the experience of medical interactions that are critical to replicate in training. This could entail tactile interactions, physiological behaviours, and patient characteristics. Replicating these experiences is essential to ensure that simulated training does not negatively affect learning outcomes and transfer to real scenarios. To understand the experience of clinical interactions, designers can thus leverage experience prototyping (Buchenau & Suri, 2000). This was done in the projects by acting out different interventions and training routines (internally) by using readily available artefacts and activities. As it is the experience of interactions that is being prototyped, the objects merely acted as enablers. E.g., performing actual needle insertions for the feel of vascular access (P3), using piglet cadavers for chest compressions (P1) and fracture realignment (P2), and ultrasound testing of both human and inanimate objects to compare image quality (P3). Experience prototyping, activities and interactions as seen in Figure 4.3, provides insights and tacit knowledge to be utilized in further development.

Figure 4.3 Prototyping interactions and prototyping experiences in the case projects. P4 showing palpation, P2 fracture realignment (users ↑, pigs' leg ↓), P3 ultrasound demonstration (↑) and needle insertion (↓), P1 chest compressions performed on a piglet cadaver.

4.3.4 Design Problems

Early problem exploration generates insights on training needs, medical context, and tactile interactions. These insights could, and arguably will, spawn inspiration for future products, novel equipment, or opportunities to improve current training experiences. All of these are design problems —concretized development opportunities that could solve or address current training needs. However, to reveal the relevance, importance, or applicability of design problems, they need to be made tangible for deciding whether to pursue or discard them. Eliciting these design problems are thus crucial and is summarized by the following insight gained from the inspiration activities carried out in the project cases:

Insight on Inspiration: Meaningful interactions with users can be facilitated by prototyping. Prototypes can enable communication by acting as boundary objects, being tangible artefacts to describe and decode tacit knowledge. Prototypes can make complex experiences tangible for designers by enabling users to act out routines and demonstrating procedures, or by designers learning by doing. Design problems can be established and informed by prototyping.

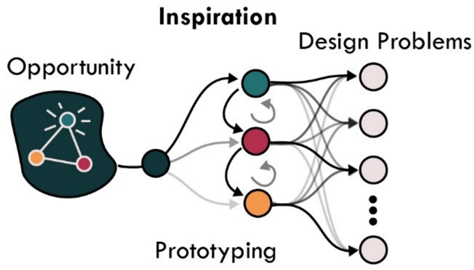


Figure 4.4 Inspiration.

4.4 Exploring Solutions – Ideation

Ideation is prototyping activities deployed to generate, evaluate, and iterate ideas. Exploring solutions that could address design problems (needs and opportunities) requires ideas to be generated and evaluated. Hence, solution exploration should reveal what design problems are worth tackling, and how we can sufficiently address them. These insights could be gained from ideas made tangible by prototypes, artefacts used to reveal the potential, limitations, and application of ideas. As both problems and contextual conditions are unique and evolve throughout projects, it is evident that we require different prototyping efforts to obtain insights from various sources concerning different product aspects.

4.4.1 Diverging, Converging, and Iterating

Faced with (multiple) design problems, each containing subsets of potential solutions, speed and agility are essential. Early solution exploration should therefore leverage generative prototyping and divergent ideation. Hence, we should gain insights by exploring the different design problems and generating multiple alternative solutions in parallel. Prototypes should, therefore, only be as comprehensive as necessary to reveal or produce the required insights. In most cases, this is sufficient knowledge to decide whether a (preliminary) proposed solution is purposeful to explore further or discard —often regarded as failing fast. However, as we explore problems and solutions on a limited, preliminary, and ambiguous foundation, this could also lead to great potentials being elicited. These findings are often serendipitous and could both influence our understanding of the problems, ideas, and vision for future products.

Iterating entails pursuing revealed insights by creating subsequent and more informed prototypes. Enhancing and investigating aspects of the design that shows potential to progress in development often requires more comprehensive prototypes for these insights to be gained. As consecutive prototypes generate new insights, the number of unsolved design problems and alternative solutions decreases. Hence, only the most promising idea(s) will progress, and the least successful ideas, according to their potential, prototype performance, and alignment with revised problem understanding, are discarded. Iterations hereby assure convergence based on the increased understanding of the problem(s) and solution(s).

4.4.2 Deciding Resolution and Abstraction across Disciplines

Creating purposeful prototypes is critical when aiming to diverge and converge on ideas rapidly. However, what makes a prototype purposeful is not evident as different prototypes (and their attributes) yield different affordances, typically differentiated by their closeness to an eventual design (i.e., fidelity). While fidelity considers perceived, intended, and functional refinement, it results from the amount of detail achieved by a prototype (i.e., resolution) and how comprehensively it represents the concept idea (i.e., abstraction). Hence, purposeful prototypes are created by managing resolution and abstraction parameters following the intended learning and intended use case of the prototype C2 (Auflem et al., 2020).

Resolution and abstraction describe a prototype's characteristics across multiple disciplines, which means that when prototyping, we decide on these parameters to generate a purposeful outcome. However, iterations do not necessarily suggest higher resolution and lower abstraction simultaneously across all disciplines. They are altered interchangeably, dependent on what the prototype is prototyping, as suggested by Houde & Hill, (1997). I.e., while improving on individual prototype characteristics, for example, higher hardware resolution, rapid iterations could be assured by deciding on low resolution and high abstraction across disciplines not critical for the prototype. To the left in Figure 4.5, are prototype examples from P10 mapped according to their resolution and abstraction concerning hardware design.

Furthermore, on the right, multiple disciplines involved in a prototype test has been mapped by the same dimensions.

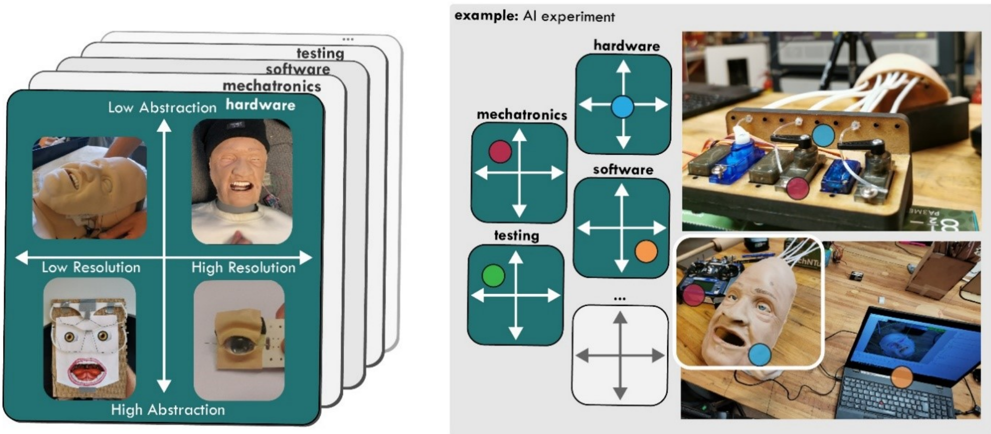


Figure 4.5 Resolution and abstraction across knowledge domains. Exemplified by experimenting with AI from P10.

4.4.3 Prototype to Learn

To leverage prototypes in learning and obtain project insights, we investigate individual aspects of an eventual product through focused experiments. E.g., learning how a product could enable training of medical tasks by models to facilitate relevant and clinical interactions. Contributions C1, C4, C9, and C11, showcase how low abstraction and low resolution prototypes have yielded insights for novel training opportunities (Auflem et al., 2019; Ege et al., 2020, 2022; Henriksen et al., 2023). Furthermore, how to achieve a realistic appearance or tactility, exploring material concepts using high resolution and high abstraction could be purposeful. E.g., the development of novel material concepts for emulating skin (P10), ultrasound phantom (P2) C10 (Steffensen, Kriesi, et al., 2022), and a novel material technology for haptics described in C7 (Rørvik et al., 2021). Similarly, learning how technical functionality (e.g., sensors or actuators) high resolution within a mechatronics domain might be sufficient to gain relevant insights as exemplified by Steffensen, Auflem, et al., (2022) . In C2 (Auflem et al., 2020), we describe how the selection of methods and mediums for prototyping causes unavoidable trade-offs. This suggests that intended learning by prototypes should decide the need for resolution and abstraction across the prototyping disciplines.

Ideation and prototyping insights support establishing an informed product vision, meaning an explicit understanding of what the product will become and what it needs to accomplish. What it needs to accomplish are the attributes required for achieving sufficient usability, realism, and functionality. These aspects are the critical functions a prototype should aim to achieve and, thus, become core objectives to pursue and realize in a concept. Hence they govern and guide further development (Auflem et al., 2019; Kriesi et al., 2016). The following insight on creating prototypes to learn and ideation has been obtained from the case projects:

Insight on Ideation: Aspects of eventual designs and products can be explored by diverging, converging, and iterating solutions with prototypes. Varying resolution, and abstraction across disciplines, sufficient prototypes can provide insights for establishing critical functions, progress development, and inform a future product vision.

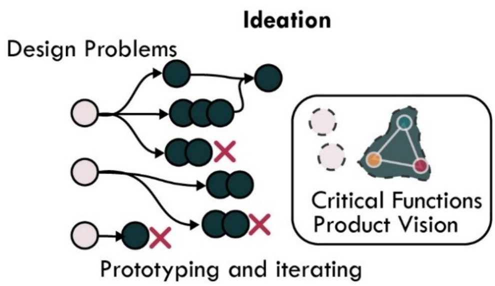


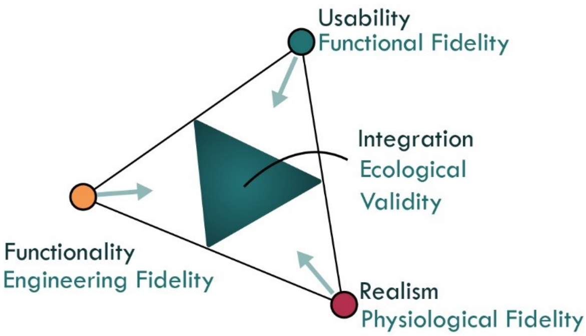
Figure 4.6 Ideation.

4.5 Prototype Systems and Fidelity – Conceptualization

Conceptualization is prototyping activities to combine and evolve concepts and decide sufficient fidelity given the use-case and stage of development. Prototyping enhances our understanding of what a new product should entail by means of —usability in training scenarios, the realism it should encompass, and how it should realize and integrate functionality. However, having addressed individual design problems (that are deemed purposeful to pursue), conceptual prototypes, i.e., systems, are required to provide insights into the actual product vision(s). Conceptualization includes integrating and utilizing the solutions and insights we have obtained to create a prototype that could be tested. As conceptual prototypes are approximations of the eventual product, we need to consider (and decide) how closely it needs to approximate —human physiology, a production-ready model, and a finished product experience. Deciding on the required fidelity along these dimensions is used to progress development towards informed and ecologically valid product proposals.

4.5.1 Ecological Validity

Ecological validity considers the setting, task, users, and scenario representativeness to clinical events (Kushniruk et al., 2013). Subsequently, ecological validity considers a prototype's ability to enable training of relevant medical tasks by presenting functionality and attributes necessary for users to realistically perform them, C5 and C6 (Ege et al., 2020, 2021). This means a prototype collectively representing sufficient fidelity along the three dimensions, given the training context, use-case, and intended users. These dimensions consider the aspects an eventual product encompasses, representing functional, physiological, and engineering fidelity.



***Ecological validity** is a prototype's ability to **enable training** of relevant medical tasks, by having the necessary **functionality** for it to be **realistically** performed.*

Figure 4.7 Ecological validity and dimensional fidelity.

4.5.2 Functional Fidelity

Functional fidelity is the prototype's usability and whether it facilitates required training interactions (Buchenau & Suri, 2000). By functional fidelity, we explore if included functions sufficiently enable interactions by the intended users and accommodate clinical instruments and equipment. These include physical interactions, i.e., palpation, interventions, or practice of psychomotor skills—and cognitive tasks, i.e., assessment, diagnostics, and decisions. Furthermore, how the prototype provides training feedback or performance metrics to enable skills acquisition, improvement, and retention. Enhancing and iterating this dimension ensures that the concept is usable and includes the relevant features to facilitate training.

4.5.3 Physiological Fidelity

Physiological fidelity considers the clinical, visual, and tactile realism of prototypes and how closely they approximate physiological traits. This furthermore addresses the look and feel of prototypes, meaning sufficient inclusion of anatomical landmarks, textures, tactility, or physiological responses. The realism of prototypes is essential for immersion and ensuring sufficiently realistic training conditions. These effects could contribute to, or impede, the transfer of learning to clinical scenarios (Maran & Glavin, 2003).

4.5.4 Engineering Fidelity

Engineering fidelity is how close to a finished product a prototype is considered and perceived (Houde & Hill, 1997; Maran & Glavin, 2003). This relates to the included functionality and decisions made to realize the prototype. E.g., integrating sensors to obtain objective measurements on prototype performance or behaviour. Furthermore, it considers robustness, reliability, and repeatability. The engineering fidelity is important for evaluating prototypes and whether flaws are caused by prototype trade-offs or concept idea limitations. For users, engineering fidelity concerns the build-quality and refinement of a prototype and its resemblance to a finished product. This is a governing factor as the engineering fidelity of an artefact can influence users' subjective evaluation. For example, a crude prototype will be more easily discussed on its potential, while a refined prototype will be more easily critiqued on its flaws (Edelman et al., 2009; Houde & Hill, 1997).

4.5.5 Iterating Towards Sufficient Fidelity

The goal of creating conceptual prototypes is to inform decisions for what a product should do and how it should do it. This is considered underlying knowledge for further development to eventually launch the product. Hence, the aim is to determine the needed refinement of concepts given the intended use-case. As different use-cases yield different fidelity requirements, iterations of the concept should therefore be made to achieve sufficient attributes for experiments and user testing to ascertain the concept.

As examplified in Figure 4.8, projects P2, P4, and P5 used different approaches and sequences of addressing required fidelity. In P2, an initial prototype investigated the use case (1), which elicited tactile realism needed to perform fracture reduction, requiring iterating to find a solution providing realistic tactile sensation (2). Prototypes to implement and sufficiently integrate this solution were developed (3), leveraging insights for functions elicited from the initial prototype. In P4, tactility for palpation training was explored (1). Sensor-based feedback was revealed as critical functionality that could enhance training (2). Prototypes and insights were used to create a proof-of-concept (3), to explore use cases and training contexts. In P5, a concept idea for new functionality in simulator necks was developed (1). Iterations to achieve the required realistic feel and movement when interacted with resulted in (2). Creating a prototype that could be used to evaluate different use cases was performed (3), using both insights and preceding iterations. These examples show that functional, physiological, and engineering fidelity can, and will, be addressed at different stages and sequences based on the elicited critical functions and needs. The following insight summarize conceptualization using prototoypes:

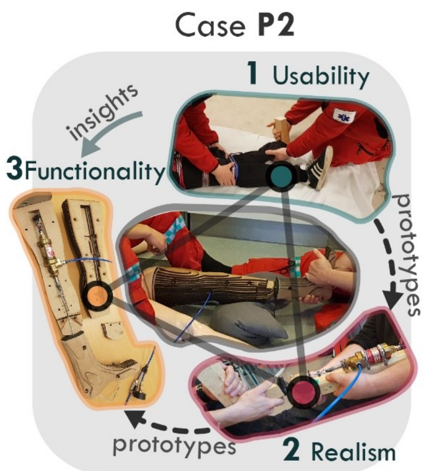
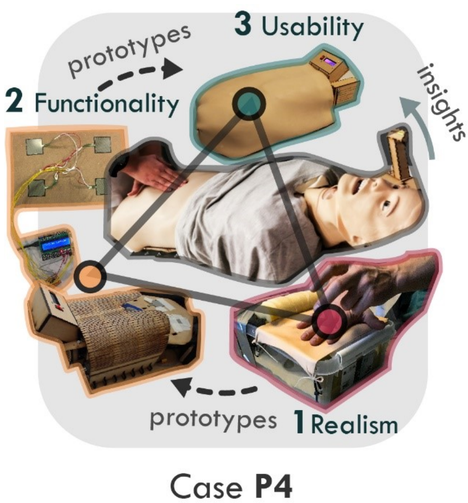
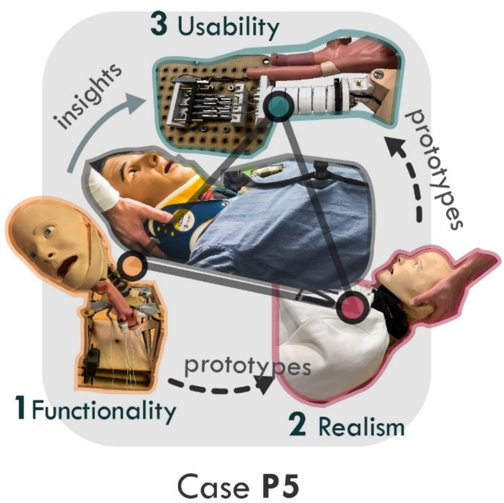


Figure 4.8 Prototypes used to aim the required fidelity in functional, physiological, and engineering dimension. Showcased by projects P2, P4, and P5.





Insight on Conceptualization: Insights on the applicability, viability, and usability of product ideas, requires concept prototypes to explore sufficient and required fidelity along functional, engineering, and physiological dimension. This progress development by creating and iterating system prototypes that can be tested and evaluated for ecological validity and inform further development.

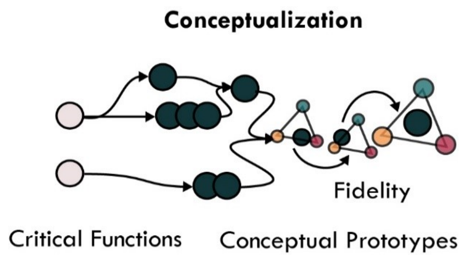


Figure 4.9 Conceptualization.

4.6 *Experiments and User Testing*

Tests and experiments are affirmative activities performed to understand, describe, verify or validate prototypes (Batliner et al., 2022). These activities vary throughout the development and are determined by the prototypes and intended learning outcomes of the tests. When the prototypes remain simple and numerous, tests remain simple, e.g., internal activities designers perform based on the current understanding of the problem. However, as knowledge about the design problem is obtained, fidelity and functional attributes of prototypes are subsequently enhanced. As these prototypes and their intended interactions become increasingly complex, increasingly comprehensive test procedures are required C4 and C5 (Ege et al., 2020; Lilleløkken et al., 2020).

Prototype testing can alleviate designer bias by obtaining external subjective evaluations by user testing, objective measurements, or combinations between the two. Furthermore, experiments and user tests are also considered prototyping activities. Hence abstraction and resolution should be decided from the intended learning outcomes and allocated timeframe. Hence, testing should be performed early and iteratively to learn about future implications and iterate prototypes and tests in parallel (Thomke & Bell, 2001). Test procedures, user involvement, environment, and protocol, should be altered and shaped to reflect the current state of the prototypes (Vestad & Steinert, 2019).

4.6.1 Prototyping Experiments in the Lab

In a controlled lab environment, we can rapidly alter testing conditions and prototypes with control over external factors and stimuli. Prototype environments and tasks can filter distinct aspects of the intended use-cases and be tailored to specific functional attributes of the prototypes we are testing. This enables multiple tests to be easily performed repeatably and reliably. Lab experiments have especially contributed to the understanding and describing of prototype functionality concerning sensor applications (P1, P3, P5, P10). An example is showcased in Figure 4.10, by the evolution of prototype, test, and sources of data in project P3.

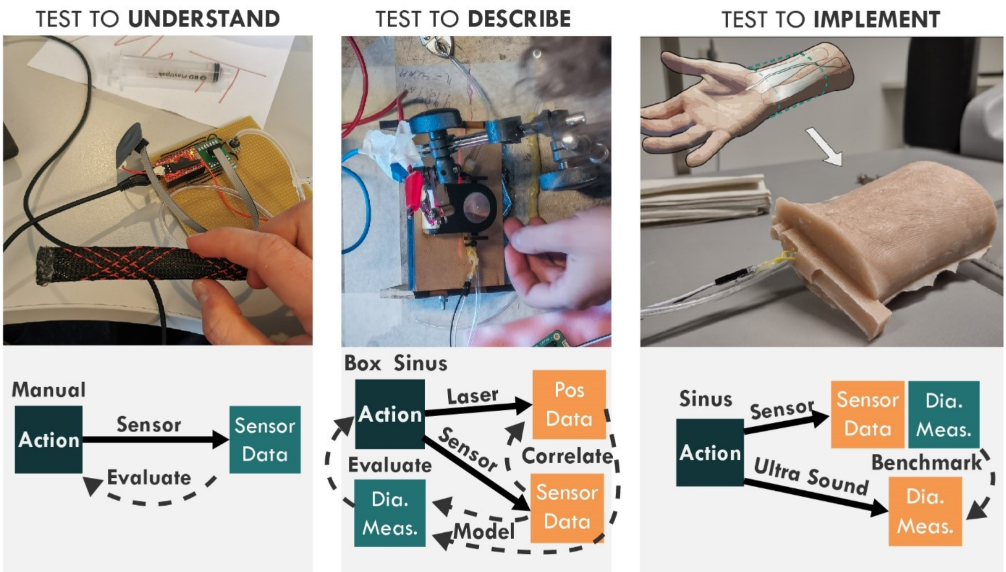


Figure 4.10 Iterative testing of vascular induction sensor P3. Increasing prototypes engineering fidelity, and test comprehensiveness accordingly.

Test to understand are the simplest experiments, performed to learn and suggest further development opportunities. Describing prototype functionality, requires more comprehensive tests, where additional objective data (sources) could be required for results to inform further development or be communicated to external stakeholders. Implementation requires benchmarking where we can correlate our results to an objective and grounded variable. Hence, having multiple variables that can be correlated is often purposeful.

Experiments have furthermore revealed novel applications of tools and methods for prototype testing. Exemplified are image classification and machine vision applications used to measure and compare complex user prompts (drawings) C7 (Rørvik et al., 2021). Object tracking leveraged to correlate movement with sensor readings C8 (Steffensen, Auflem, et al., 2022). Machine learning applications enabled human-like analysis and predictions used to approximate human behaviour C3 (Auflem et al., 2022). In experimental prototype tests, the rapid and nuanced analysis that is enabled by these tools can inform decision making and thus, improve design outcome P10 and C3 (Auflem et al., 2022).

4.6.2 Prototyping User Tests

User testing is to pilot conceptual prototypes within the intended environment, under similar conditions, and with the targeted users to assess and inform aspects of an eventual product. In the project cases a flexible and multifaceted framework for designing and performing structured experimental tests have been leveraged C5 (Lilleløkken et al., 2020). This helps deciding tasks to address the functions we are testing, who needs to contribute to the tests, and how the tests should be performed. Hence, it requires designing an environment (location, tools, and users) to tests within, and scenario tasks and objectives to be performed. Figure 4.11, shows the environment in TrollLABS Medical (P6), and the tasks performed in the scenario (P5) described in C11 (Henriksen et al., 2023).

Figure 4.11 User testing of neck immobilization and patient transfer in P5.

4.6.3 Performing User Tests

An example of a user test protocol is shown in Figure 4.12. Execution of user tests has, in the projects, involved creating a set of actions the users must perform. These are tasks and objectives that can be assessed and measured objectively, e.g., a simulated clinical task with correct answers as measurable output. Baseline and free interaction are actions that allow for flexibility in the execution of the tests and a source of subjective and descriptive data. By a mixed method approach, the obtained results can be used as both assessment parameters, and descriptive feedback. The results can, moreover, be structured and weighed to answer the aspects of usability, realism and functions included in the prototype. Analysing results can ascertain, describe, or indicate areas of improvement using ecological validity as the collective goal.

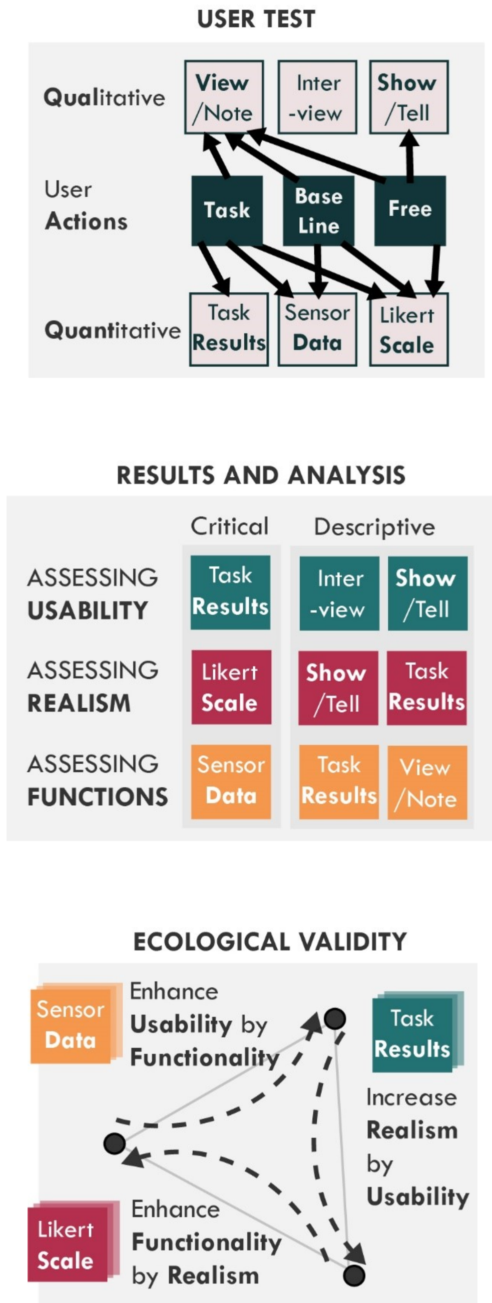


Figure 4.12 User testing approach, modes of result, and ecological validity considerations.

Usability by **functional fidelity** is assessed by whether users can successfully perform the task given the included functionality of the prototype. Hence, the prototype should provide objective and quantitative results, which could require sensors to be implemented, or obtaining additional data from sources e.g., surveys, drawings, or signalling cues given by the participants. Subjective evaluation can assess how functionality is applicable in a training scenario. This includes comparing it with clinical experiences, and if the required functionality is present to perform routines suggested by curricula.

Furthermore, if sufficient feedback is provided to understand and successfully perform tasks.

Realism by **physiological fidelity** is assessed by subjective evaluation and by users comparing functional traits and attributes with clinical experience and anatomical knowledge. This evaluation of the prototype's realism can benefit from using tools such as a Likert scale and questionnaires to structure and analyse results. However, tactile perception is both subjective and hard to benchmark, so open dialogue and unstructured interviews during the interaction are purposeful to reveal areas for further development.

Engineering fidelity considers the decided **functionality** and how it has been realized, this dimension is valuable to generate insights for further development. Hence, including sensors and utilizing sensor data can be valuable. Objective and descriptive data from sensors are valuable to correlate results with the other fidelity dimensions. This could be a source for further improvements by leveraging prototyping with data through explorative data analysis.

Ecological validity considers if sufficient functionality is in place for users to perform the required tasks. Furthermore, if the usability allows for realistic interactions, and that these interactions sufficiently replicate clinical experiences. Lastly, if the functionality is sufficiently realistic for users to perform the tasks and is deemed a sufficient training solution. How user testing and testing protocols are adapted to different product and testing conditions are further elaborated on in contributions C4, C6, C7, C9, and C11.

Insight on Testing: Structured testing of prototypes entails both lab experiments or user tests. Lab experiments can test specific aspect and functions of prototypes to understand, describe, or implement. User testing addresses collective functions of a prototype and requires a multifaceted test approach to evaluate ecological validity and areas of improvement concerning functional, physiological, and engineering fidelity.

4.7 Prototype-Driven Development

Prototyping activities have, in the previous sections, been shown to make problems tangible through various modes of interaction, facilitate ideation, and create and test conceptual prototypes. However, in projects, these activities are seldom organized and consecutive but rapid and entwined. In practice, this means prototypes are designed, built, and tested in rapid cycles where each cycle generates new ideas, needs, and constraints for subsequent prototypes. This emphasizes that prototypes are the main drivers of project progress and for obtaining the required insights to do so; hence the working mode is prototype-driven.

4.7.1 Exemplified by Project Case P10

The development of a new face for patient simulators in P10, as seen in Figure 4.13, will be used as an example to describe prototype-driven development project workflow. The figure shows prototypes built throughout the project and how characteristics and aims are developed over time. This can be noticed by resolution and abstraction shifting between prototypes looking like faces to inanimate objects testing abstracted design problems. Indicated along the project timeline are findings and milestones. Findings are serendipitous or novel discoveries made from prototyping. Milestones are conceptual prototypes for affirming or shifting development direction from testing the applicability of decided solutions.

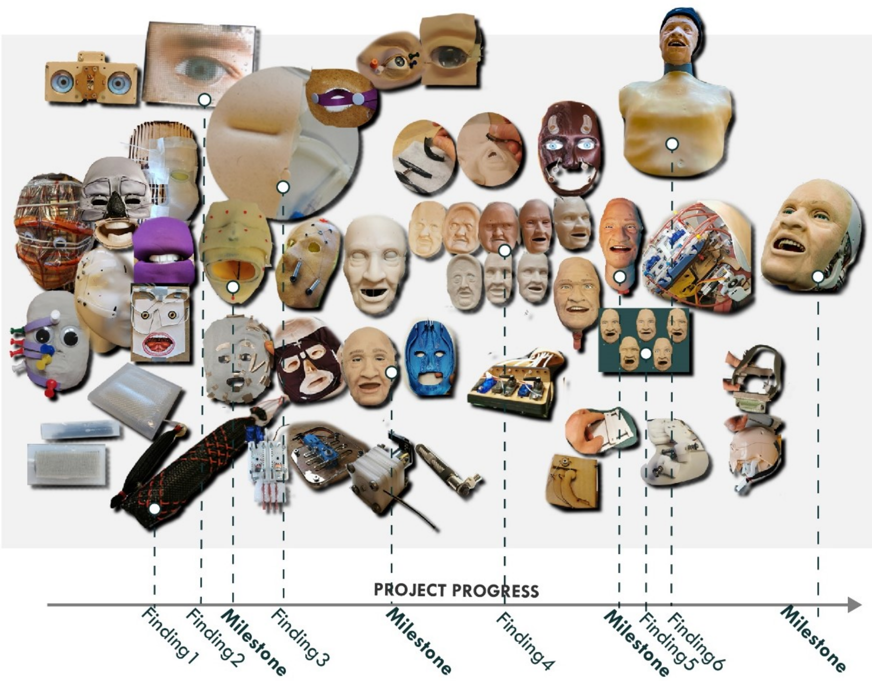


Figure 4.13 Prototypes developed throughout the P10. The project has investigating nonverbal communication and facial expressive simulator functionality. Findings and milestones are indicated along the project timeline.

A retrospective development mapping through prototypes, problems, iterations, and testing activities are seen in Figure 4.14. Although interconnected, exploration, conceptualization, and experimentation phases are indicated. The cadence of diverging and converging is illustrated in the lower band graph. Observations from this mapping are:

- While appearing throughout the project, more findings occur before the solution space is confined (1st. divergent-convergent cycle). Later findings, however, contribute more to the final concept (2nd and 3rd. cycle)
- Exploring problems, solutions, and experiments all endure divergent and convergent working modes.
- Functionality expands as conceptual prototypes are re-iterated (2nd. divergent cycle). Suggesting more opportunities and ideas emerge as new insights are obtained.

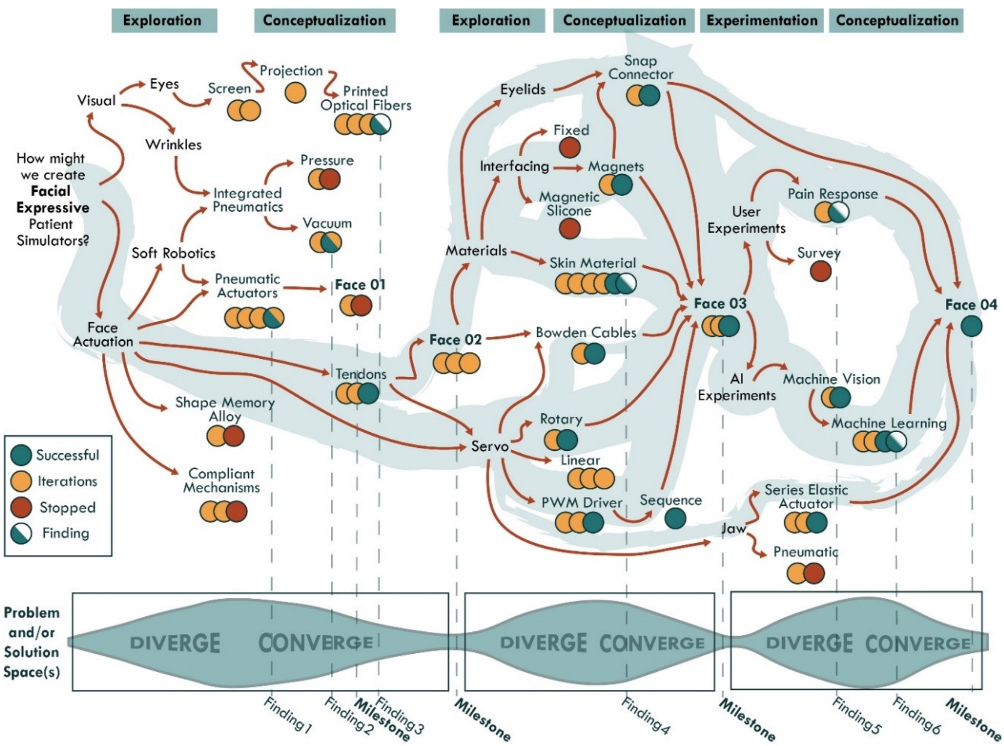


Figure 4.14 Prototype-driven development and relations between prototypes in P10.

These observations are made from a singular project and not correlated with comparable project examples. Nevertheless, the project provides an overview of how development, prototypes, and phases evolve, i.e., by nonlinear, divergent, and convergent cycles. The following insight summarize this as:

Insight on Prototype-Driven Development: Prototypes progress development by stimulating new ideas and eliciting findings to explore. We utilize different tools, and prototyping approaches sporadically and opportunistically in prototype-driven development. The purposes of prototypes are furthermore not singular, and hence the process can seem chaotic, pacing through sequences of divergent and convergent efforts.

4.8 Project Outcomes

This section structures findings across the development projects and relates them to different product aspects and potentials for improving medical training, as visualized in Figure 4.15. How new concepts **enable training** means identified needs in training and education and how projects have attempted to fill a gap in currently available equipment. Concepts that can enable and provide equipment for simulated training and hands-on practice in current education have been emphasized in P5 (neck immobilization), P4 (palpation exercises) and P2 (fraction reduction). Furthermore, accommodating the different challenges found in clinical settings requires altering training difficulty. Thus, introducing variance and uncertainty in scenarios, and individual differences to be captured by **variable functions**. This means that training equipment can be tailored to case complexity, scenario diversity, and user competency as exemplified by P1 (different physiological patient characteristics), P2 (variable muscle tension), and P5 (different injuries, scenarios, and timeframes).

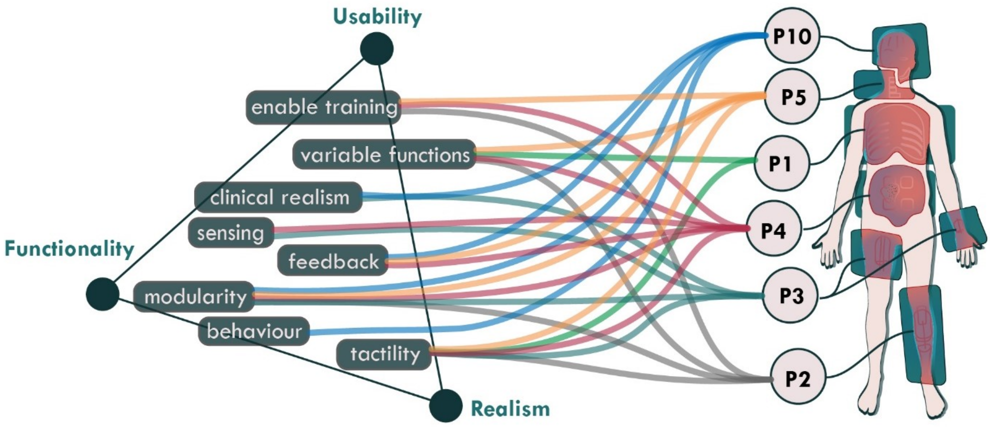


Figure 4.15 Synthesized findings for developing simulation-based medical training equipment.

Modularity involves the opportunities for concepts to be altered and adapted to scenarios and patient characteristics. This was realized in P2 by having interchangeable modules that could emulate different fracture characteristics. P10 focuses on having a swappable appearance, thus allowing for greater diversity in simulated patients. With a modular design, concepts can be developed faster to contribute value and functionality to existing mannequin architectures. Especially in the development context, this is purposeful for integration in existing simulators to achieve higher fidelity and enable experiments and user tests to be carried out quickly.

Tactility, tactile realism, and haptic functionality are essential to ensure familiar hands-on experiences and transfer of learning to clinical scenarios. Hence, it has been a core focus to approximate the tactility of human physiology for different simulated interventions. This has led to the development of concepts and technologies such as skin-like composite for P10, Ultrasound phantom material P2, and ferrogranular jamming for haptic simulation P4. **Clinical realism** entails how realistic a concept allows for clinical interventions to be performed. In P3 this is stressed by the importance of enabling the use of real clinical equipment, such as vascular access by needles and ultrasound imaging. The face project P10 was focused on maintaining existing clinical features such as the jaw, airway accessibility and mobility while adding features to enable communication and **behavioural** features. In P10, communication and simulation of cognitive states are both important and, in different scenarios, critical to ensure realistic and immersive simulation experiences.

Feedback is crucial in training and to facilitate and enhance learning. Human-like feedback by facial responses were requested in both P2 and P4 and developed in P10. This was furthermore tested in a cognitive response experiment using sternum rub to assert patient awareness. This suggests that future medical simulators need clinical and social cues in combinations to ensure immersion and explore broader use-cases to alleviate unfamiliarity effects and enhance training output.

Sensing has been critical in enabling feedback and spanning beyond triggering responses. For example, how different formative (learning and correction) and summative (scoring and evaluation) feedback metrics could be obtained and used was explored in P4 and P5. This provides an outlook on how more self-directed training solutions, increased realism, and enhanced functionality could allow seamless and uninterrupted training. This, however, requires increased autonomy and assuring feedback is empirically grounded and evidence based. Furthermore, this will require an even deeper understanding of the routines and interventions that must be practised, how they are performed in clinical practice, and how they can be recreated in a simulated setting.

4.9 Recap, Reflection, and Transition

This chapter has provided insights and approaches to prototype concepts for simulation-based medical training. Firstly, through methods to explore problems and opportunities by meaningful interactions and learning by doing approaches. Secondly prototyping to learn by creating purposeful prototypes and deciding required abstraction and resolution. Conceptual prototyping has been described to ensure required usability, functionality, and realism. Experiments and user testing, approaches have been presented to ascertain these aspects. This can be visualized as Figure 4.16, suggesting relations between methods, process, and learning in prototype-driven projects.

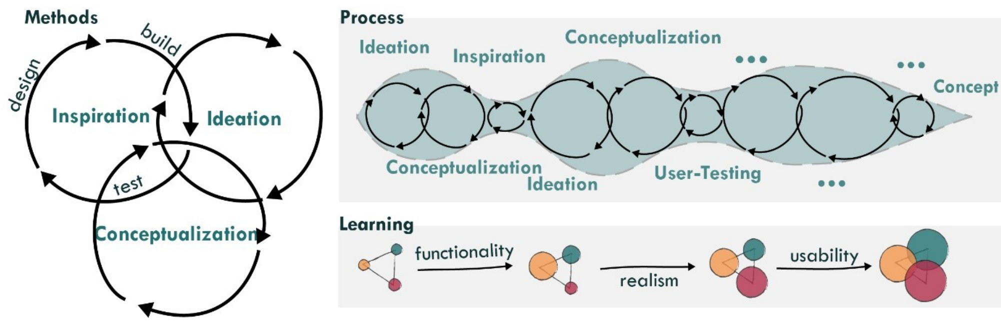
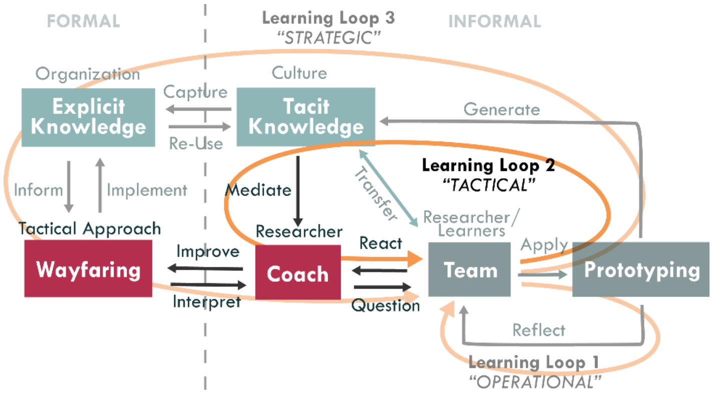


Figure 4.16 Visualized methods, process, and learning in prototype-driven projects.

However, while having presented methods and approaches sequentially, prototype-driven development (as exemplified by P10) is complex, opportunistic, and dependent on the problems faced and insights gained. Hence, by the uniqueness of projects, supporting prototype-driven development from a managerial perspective is not evident. The next chapter concerns coaching and managing pre-requirement projects, and how to use wayfaring to support prototyping in the FFE.



5

How we Think

to Facilitate New Product Development, Prototype-Driven Projects, and Wayfaring

This chapter accounts for the second learning loop by assuming a coach or development management perspective. It describes how NPD projects could be initiated and facilitated — given the conditions of FFE, development using prototype-driven approaches, and product context being simulation-based medical training equipment. It involves asking questions to facilitate development teams using framing, filtering, and manifesting of design ideas to produce purposeful output. Project progress, autonomy, and sharing is suggested facilitated by time-boxed development activities. Furthermore, considerations for assembling development teams and their roles are presented. These modes of working and thinking are operationalized in a proposed wayfaring framework.

5.1 Managing Development and Prototype-Driven Requirements

From a managerial perspective, development efforts should be made to mitigate the risk of concepts not meeting the target users' needs or not being feasible to deliver. This means development activities sufficiently enlighten critical aspects of product realization (usability, realism, and functionality), and that these are collectively sufficed before further decisions are made. Frontloading these activities limits the risk of encountering obstacles later in the development, which would require both costly and time-consuming rework. (Thomke & Fujimoto, 2000; Thomke & Reinertsen, 1998). Moreover, effective, and efficient development should create new knowledge and generate relevant and actionable insights. This requires sufficient room for teams to do creative exploration while at the same time managing project scope and development increments (Elverum & Welo, 2015).

Prototype-driven requirements described in C1 (Auflem et al., 2019), suggest not constraining the solution space, but rather elicit product requirements based on insights obtained from prototyping activities (Schrage, 1993; Sutcliffe & Sawyer, 2013). This allows for being flexible in response to encountering unforeseen obstacles or novel opportunities in development (M. B. Jensen et al., 2017). Furthermore, critical product functionality and subsequent requirements will follow and reflect the shifting trajectory of development as new insights are obtained (Kriesi et al., 2016).

A visualization of ill-informed requirements, unbalanced project scope, and insufficient exploration is shown in figure 5.1. This suggests the importance of ensuring open problem and solution spaces, by instigating projects with broad and open-ended task formulations as described in the project cases. Concerning project execution, the following sections will describe how asking questions could influence project scope, direction, and prototyping activities. Furthermore, managing time by development sprints, will be described for operationalizing prototype-driven development in a wayfaring framework.

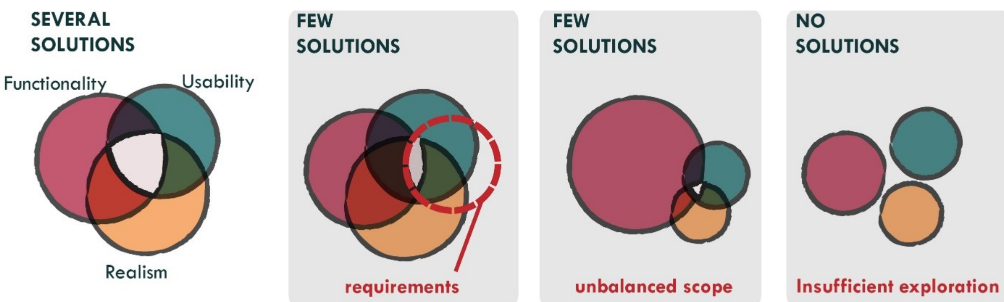


Figure 5.1 Visualizing project exploration, scope, and requirements.

5.2 Asking Questions

In NPD, we aim to generate new knowledge by utilizing what we know (skills, tools, and methods) to elicit and inform what we do not know (problems, opportunities, tools, requirements, and needs). Asking questions is thus an integral part of product development as it is in problem-solving (Schrage, 1993). The questions concern uninformed aspects of the product opportunity, encountered problems, or eventual design aspects designers try to address. By this notion, questions govern the sequential steps towards a successful concept, and manifest the current state of the project and the contextual knowledge the designers possess.

From the cases and synthesized insights, an argument can be made that we wish to elicit critical questions and their relation to the task as early as possible. The nature of these questions could be *unknown knowns*: questions we have solutions to or ways of solving, but we need to know their relevance to the challenge. *Known unknowns*: questions we still need to learn how to address but are aware of their relevance to the task. And *unknown unknowns*: that are questions we are not yet aware of and do not know how to solve (M. B. Jensen et al., 2017; Sutcliffe & Sawyer, 2013). The *unknown unknowns* might be the ones most critical to elicit, as these can impede decisions, and significantly pivot the development direction once encountered.

The questions we ask address the uncertainty we face at a given time and stage of development. The characteristics of these questions also reflect and influence the aim and scope of the development efforts we perform. With little knowledge about the problem and solution relations, "how might we" questions (HMW) allow for rethinking problems and eliciting new opportunities. Generative design questions (GDQ), like "how many ways can we" questions, ensures that multiple alternatives are considered and enables divergence by their open-ended characteristics. Convergence and inductive reasoning, on the other hand, are facilitated by affirmative and deep-reasoning questions (DRQ), like "can we use …". DRQ are required to make decisions based on an increased understanding of the problem and solution relations (Ö. Eris, 2003).

Teams should focus on generating and eliciting questions that govern progression, and make sure to pose questions that (individually and/or collectively) target critical aspects of a future product. This involves questions that informs what sufficient usability, realism, and functionality entails by the given problem as described in chapter 4. Moreover, asking questions are important as they act as a communicative two-way street. From a designer's perspective, communicating questions can concretize problems and suggest approach for addressing them (reflective or affirmative). Similarly, from a management perspective, posing questions can stimulate certain activities (divergent or convergent) and suggest direction and scope of these activities in projects.

5.3 Answering Questions by Prototyping

Questions are in development projects addressed by prototyping (Schrage, 1993; Wall et al., 1992). Contradictory to being merely the creation of prototypes, prototyping is considered a learning activity. Prototyping is both cognitive and physical, where iterative design, build, and test cycles are used to bring answers to, or generate new, questions C1 (Auflem et al., 2019). *Designing* is the activity of generating ideas and concepts. *Building* is realizing ideas and deciding upon medium and methods to formalize them C2 (Auflem et al., 2020). *Testing* are activities performed to evaluate prototypes and provide answers to, or generate new, questions to progress the development C6 (Ege et al., 2021). As projects progress, it is evident that questions will be generated, answered, or become irrelevant, as new insights are obtained. Therefore, prototyping is iterative, and each iteration inform sequential steps by generating new knowledge.

Prototype-driven development, as described in chapter 4, is dynamic and often chaotic. Fractal connections occur between prototypes, learning outcomes, user-insights, and other sources of learning. These connections can be leveraged by reflective thinking, learning from prototypes, failures, and allowing serendipitous findings to occur. Therefore, it is challenging to plan prototyping, but the thinking process that goes into the activity and decisions and considerations made by developers can be supported. Hereby, three categories addressing *what is being prototyped*, *why a prototype is needed,* and *how the prototype should be made,* shown in figure 5.2, will be described. Firstly, *what* means filtering aspects of the design idea that the prototyping will address and deciding on dimensional fidelity. Secondly, *why* means the strategic rationale for prototyping by framing the target audience and prototyping intent. And thirdly, *how* is manifesting design ideas by prototypes, deciding on artefact abstraction and resolution across disciplinary domains. These categories can be used to formulate questions, that ensure purposeful, and strategic, prototyping is performed in projects.

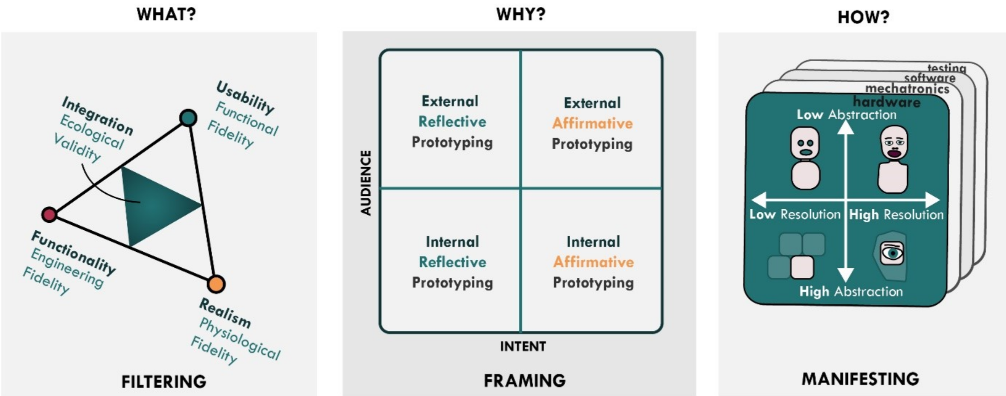


Figure 5.2 Prototyping cheat sheet. Filtering, framing, and manifesting considerations for formulating questions that stimulate, inform, and assure purposeful prototyping. Filtering adapted from Lim et al., (2008) and Houde & Hill, (1997). Framing adapted from J.A.B Erichsen et al., (2016).

5.3.1 Filtering —What?

Described by Lim et al., (2008), filtering is used to abstract aspects of design ideas to be made tangible by prototyping. The dimensions to filter can be described as Houde & Hill, (1997) suggests, by the role, look and feel characteristics and implementation aspects of eventual products. In this thesis these aspects concern the usability, sense of realism, and the functionality of prototypes respectively. Closeness to a final design along functional, physiological, and engineering fidelity, determine the attributes for prototypes based on the underlying knowledge about a problem and the need for filtering design ideas. Filtering enables thoughtful consideration of what aspects of the product idea we are targeting, and equally important which aspects we deliberately leave out. Questions should be asked to determine *what* the prototype is prototyping and how it could generate insights. Filtering is hereby purposeful for converting product ideas into prototype ideas.

5.3.2 Framing —Why?

Framing means doing decisions based on why prototypes are needed, by deciding the role prototypes will play and who we are targeting with them (Edelman & Currano, 2011; Lauff et al., 2018; Menold et al., 2017). Inspired by the SECI-model by Nonaka, (1994), J. A. B. Erichsen et al., (2016) describes prototyping activities by intent and audience in a two-by-two matrix, as seen in Figure 5.2. The intent relates to the need for prototyping, and how prototyping can facilitate learning. This is distinguished between prototyping for learning through reflective activities, and prototyping activities for affirming knowledge. The audience is divided by internal and external, depending on whether users or external stakeholder involvement is needed or required. Framing is hereby essential for planning the prototyping activity, meaning how we are addressing the filtered aspect of product ideas. Intent furthermore suggests if the outcome should be enhanced knowledge to generate new questions, or if it is made to affirm ideas and doing decisions by answering questions. This explanation can help pose and decode better questions to stimulate certain activities and aims for project teams.

5.3.3 Manifesting —How?

Manifesting is deciding how a prototype can represent an idea by resolution and abstraction considerations across disciplinary domains. It involves selecting appropriate methods and mediums for prototyping, albeit opportunities and methods for creating prototypes are abundant. Questions should thus be made to ensure purposeful prototypes are created given the context and conditions for why the prototype is needed. Hence, the intended use and context should govern the selection of prototyping methods and approaches, —ranging from creating visual representations using sketching or computer aided design (CAD) to fully functional models using 3D printers or machined parts. However, skills, resources, timeframe, and tools accessible to the designers govern the applicability of the various approaches and methods. These introduce trade-offs that developers must consider, as described in C2 (Auflem et al., 2020). Although these are derived from prototyping methods with non-rigid materials, they are relevant to a broader scope. Hence, the following parameters for prototyping methods should be considered when addressing questions by prototyping:

- Design freedom, meaning the design constraints a method causes.
- Resolution, meaning amount of detail that is obtainable using a method.

- Abstraction, meaning completeness of the representation of an eventual design.
- Modifications, meaning how easy it is to alter parts or designs after it is created.
- Speed, the timeframe to generate prototypes using a method.

5.4 Revisiting Wayfaring

Managing prototype-driven development projects can be challenging because of the fractal and evolving relations between various sources and forms of knowledge. Given the high uncertainty and lack of design constraints in the fuzzy front end, process models are required to accommodate novelty, iterative adaptability, and complexity (Wynn & Clarkson, 2018). Wayfaring is suggested to accommodate this ambiguity, ensuring agility by divergent and convergent development modes (Leifer & Steinert, 2011). It is described to support "finding the really big idea" by exploring various problems and solutions continuously, iteratively and in parallel through prototyping (Gerstenberg et al., 2015; Steinert & Leifer, 2012). This section intents to show how prototype-driven development can be operationalized in a wayfaring framework using filtering, framing, and manifesting of design ideas. This is done by deploying probes, iterative development sprints to obtain new knowledge in projects. The aim of this section is to facilitate question-focus and prototype-driven development to become manageable and leveraged in NPD. Thus, this section will revisit and adapt the mechanisms, and mental models of wayfaring to support the development of new products in the context of medical training equipment.

5.4.1 Using Wayfaring in Projects

The analogy of early-stage development being like wayfaring, as opposed to travelling, suggest a mental space that needs to be explored and traversed. Travelling is transportation through space from location A to location B, by the best possible route, concerning time, resources, and effort. Wayfaring on the other hand, is exploring the unknown, thus finding one's way through adapting the route to the obstacles and opportunities that arise throughout the journey. By this analogy, traversing development space will, and should, elicit hitherto hidden obstacles, and make unknown opportunities tangible and thus actionable.

Wayfaring is in projects carried out by probing the development space as seen in Figure 5.3. The development space is a mental model of all potential problems and solutions and represents a mapping of the project pathways, through the questions that are being elicited and answered. Probes are development activities performed as sprints by the development team to obtain insights. These are carried out by doing design activities and reflect upon the results to spawn new questions (progress), answer questions (affirming solutions), or identify dead ends and alternatives (redundancy). In addition, insights, and enhanced capabilities (tacit knowledge) can affect later development efforts or enable new product ideas to emerge through serendipity. As insights are obtained, the conceptual ideas or product visions are altered, thus changing the trajectory and aim of the project (Gerstenberg et al., 2015; Kriesi et al., 2016).

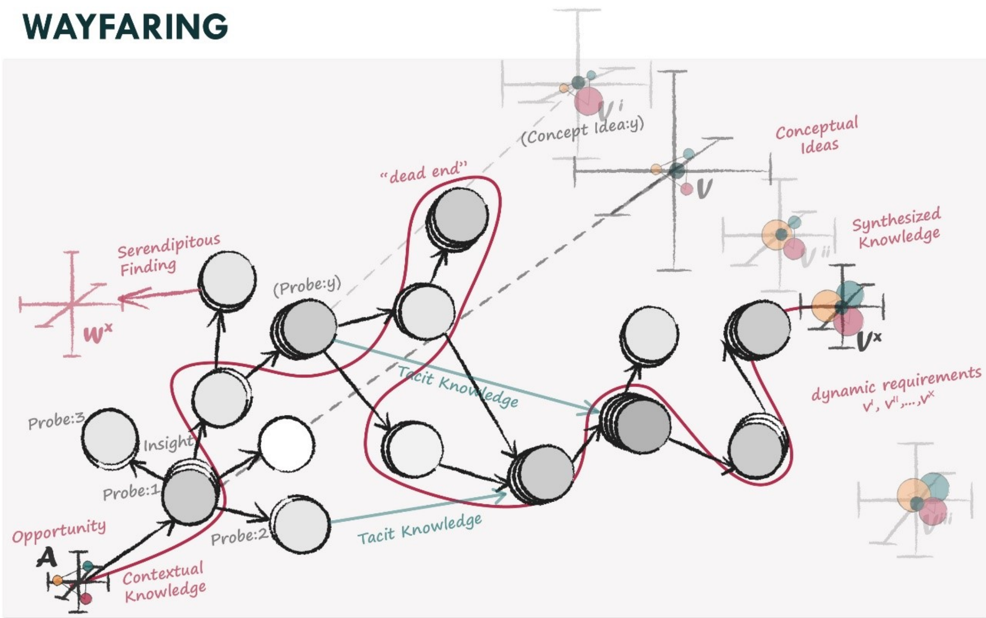


Figure 5.3 Wayfaring to manage projects by probing the development space. Adapted from the "hunter-gatherer-model" and wayfaring (Gerstenberg et al., 2015; Steinert & Leifer, 2012).

5.4.2 Probing by Prototyping Activities

A probe, as seen in Figure 5.4, is a sprint performed to address a question in development. This distinguishes a probe from a prototype, by that multiple prototyping iterations could be performed within each probe. Consequently, each probe symbolizes a development activity, where the learning output, by questions or answers, are used as inputs for proceeding probes.

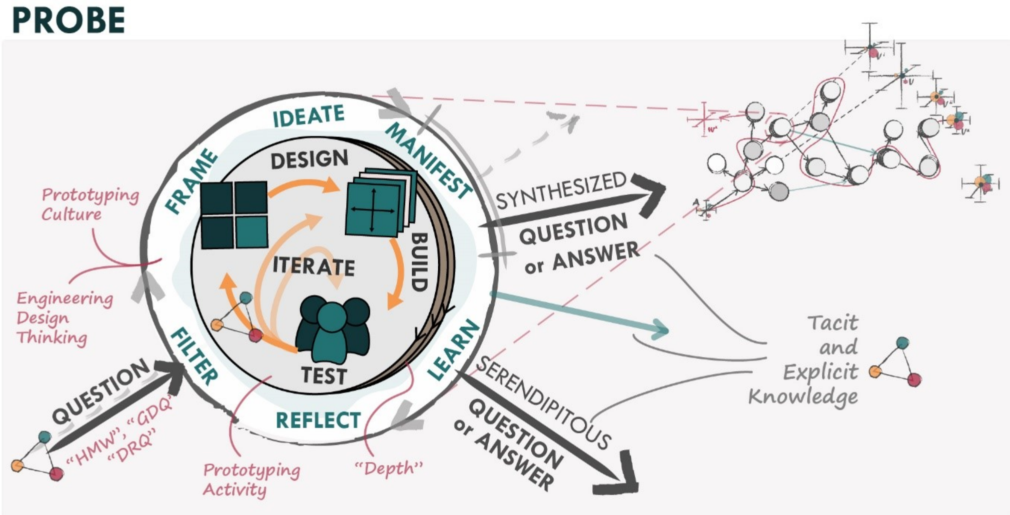


Figure 5.4 Probing by addressing questions in development.

The core of each probe is prototyping activities, which are iterative design build and test cycles performed by the development team. The depth of each probe describes the extent and effort gone into the prototyping, i.e., number of iterations, amount of fidelity, disciplines included, or comprehensiveness of tests. Depth illustrates the length of a sprint, highlighting that different questions require different prototyping approaches and efforts to be answered (Schrage, 1993).

Surrounding the core, is the engineering design thinking mindsets and prototyping culture in which the development takes place. This abstract layer illustrates the thinking, strategic decisions, and logical reasoning that supports the activity. Filtering, framing, and manifesting considerations are made in the outer layer. This layer is more complex by involving other stakeholders within the same prototyping culture (organization, lab, discipline), and could provide knowledge and capabilities not embedded in core team itself. Furthermore, the layer of prototyping culture concerns how we discuss, share, and learn from design activities. and what is being shared outside of the core can differ across project teams, organizations, and stages of development.

The output of each probe in projects is twofold. Firstly, it is the tangible and actionable insights, here described as questions and answers. These triggers subsequent probes and deciding where the project is heading. Secondly, it is the tacit and explicit knowledge generated. This knowledge shapes a concept idea, informs dynamic requirements, and mitigates uncertainty by informing the future product's usability, realism, and functionality.

While being based on the probing and wayfaring ideas proposed by (Gerstenberg et al., 2015; Steinert & Leifer, 2012), wayfaring is in this thesis adapted to be applicable in NPD management and project coaching. It highlights the interplay between other actors, e.g., management, in the prototyping culture and development team. This emphasize the notion that knowledge can be utilized without being generated by the core activity itself through enhanced inspiration, skill-sharing, and serendipity (Leifer & Steinert, 2011; Vestad et al., 2019). Subsequently, knowledge is generated without contributing directly to problem in question (tacit), and how this is absorbed and utilized by the prototyping culture through organizational learning could strengthen its development capabilities (O. Eris & Leifer, 2004; Nonaka, 1994; Schrage, 1993).

5.5 Time and Sprints in Wayfaring

Compared to linear product development process models, often found in later stages of development, wayfaring in the FFE introduce significant complexity making it challenging and often counterintuitive to manage. As opposed to optimizing a process for repeating a task, NPD aims to do something novel once (O'Donovan et al., 2005). Consequently, each project is unique, requiring dynamic and flexible processes. Hence, it is evident that managing time, resources, and the outcome is challenging. The outcome is complex and dynamic in how different opportunities require different efforts, and the opportunities might change characteristics as new insights are obtained. Likewise, the required resource allocation is dynamic and depends on the tools, technology, and skills required to generate appropriate prototypes to facilitate learning. Time is essentially the single parameter we can sufficiently control and manage in the FFE of projects. This can be achieved by timeboxing development activities by probes and enable agile development through wayfaring.

The projects considered in this thesis have been performed using timeboxing and iterative development using adapted principles from scrum and agile development (Atzberger & Paetzold, 2019; Schwaber & Sutherland, 2011; Thomke & Reinertsen, 1998). These principles involve sprinting on a day-to-day and week-to-week basis. For the intern technology projects (P8), short stand-up meetings between the core team and coaches have been performed daily. In longer timeframes such as projects P1-P5 spanning over one or two semesters or P10 exceeding two years, less frequent meetings and weekly sprints have been sufficient. Probes are not necessarily carried out over a single sprint interval, but they are timeboxed activities that are being reviewed and shared during the scrum meetings. From a management perspective, each project and stage of development might require different follow-up strategies. After completing a sprint, it is purposeful to arrange sharing sessions where reflection and integration within the development culture is achieved. This can align plans and findings with organizational efforts and share knowledge broadly within the prototyping culture. This is done by communicating current questions, and by agreeing on formulated questions team should aim to pursue next.

5.6 Development Team and Roles in Wayfaring

5.6.1 Team

The core development team should be kept small and agile. This means, to have few enough people for decisions to be made without support, and thus ensuring autonomy (Nonaka, 1994). In the project cases, individuals (P1, P5 and P10), pairs (P4), three person (P2, P3) and small groups (P8, P9) has successfully utilized wayfaring. However, even individual, or small team projects have benefitted multidisciplinary skillset and prototyping culture contributions by being integrated and sharing during scrum meetings. The benefit of having people with a diverse skillset facilitate a variety of tools and methods to be deployed when exploring. However, the team needs to be open to learn, and acquire new skills, to remain flexible. This could mean learning new tools, understanding medical procedures and how to perform them, or exploring product aspects outside of the core competency.

These characteristics are referred to as T-shaped people (Meinel et al., 2011), having an in-depth core competence that is relevant to the task, describing the vertical dimension. The horizontal dimension is the design thinking behaviour. It is people's ability to interact, communicate, collaborate, bridge disciplinary and knowledge boundaries by deploying a diverse set of tools. Arguably this horizontal dimension is also facilitated and enhanced by prototype-driven activities and the behaviours it requires and induces.

5.6.2 Coaches and Management

A coach (in many of the cases represented by the author), play a role in the projects as facilitator and asking questions to stimulate activity and strategic thinking and interactions to occur. Furthermore, the coaching role is important to broker knowledge within and between projects (O. Eris & Leifer, 2004). This means connecting existing problems with new solutions, and new problems with existing solutions. This however requires a coach to have prior and diverse experiences and being able to successfully convey experiences as actionable ideas for the team to inspect and explore.

Managers are in the context of the team considered decision makers and governing resource allocation and managing timeframes and progress. The role of the manager in wayfaring, is to ensure organizational acceptance and playroom, based on the directional and progressive output of prototyping activities. They will need to ensure teams performing autonomously, doing decisions, and playing out creative pathways, before being required to formally report and/or converge. Reporting and managing decisions, or corrections, are thereby made during the formal scrum meetings and should not influence day-to-day prototyping activities more than necessary.

5.6.3 Users

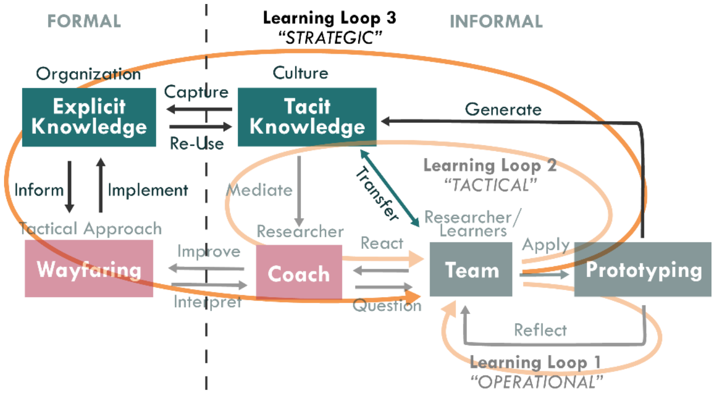
Users and stakeholders are not considered parts of the core development team but are as previously discussed crucial assets for the teams to develop new ideas. Having integration and involvement in development from stakeholders (and in the project cases EHP) can be crucial for execution of tests, meaningful interactions, and ultimately acceptance and adoption within the context it was intended. Therefore, gaining and maintaining close contact with users, stakeholders and experts is crucial, and could be achieved by a formal integration and locality within hospital environment such as the project P6 TrollLABS Medical describes. Users are however a scarce resource, where time and meaningful interactions are of essence. As previously described, this means focus and aim should be on gaining as much insight from each interaction as possible.

5.7 Recap and Transition

This chapter has provided considerations for initiating NPD projects and addressed the contradiction of development needing creative freedom to explore and managerial need of specifications and project control. Broad project scopes and unconfined problem formulations are suggested for facilitating prototype-driven specifications and dynamic requirements. Furthermore, questions are highlighted and described as valuable tools to communicate project objectives, and moreover, influence development activities. Filtering, framing, and manifesting are suggested as dimensions to aid better questions to be formulated, and thus assuring purposeful and focused prototypes being built.

These thinking and management perspectives are utilized in a revised and proposed wayfaring framework. Wayfaring facilitates that projects progress, pivot, and adapt as new insights are obtained. This framework utilizes probing to time-box development sprints and highlights the interplay between development culture and development team within each probe. Finally, considerations for assembling teams, by having t-shaped individuals, the coaching role, and management suggestions are made without elaborating on how these should be integrated in formal process models. Hence this chapter provides a bottom-up approach for facilitating prototype-driven development to occur in the FFE, rather than project management and process control.

Prototype-driven development facilitated by wayfaring, will in the following chapter be presented and discussed within the context of strategic and organizational learning. Hence, it concerns the underlying mechanisms of how knowledge generated, shared, and captured contributes to organizational knowledge. Moreover, it will provide strategic suggestions for how to leverage wayfaring, and how it has been implemented in the development practice at Laerdal Medical.



6

Why we Prototype

Prototyping to Learn and Wayfaring in Product Development Organizations

As a recap, generating, communicating, and reassuring knowledge by prototyping have been pronounced topics throughout this research project. This chapter accounts for the third learning loop and will, from a strategic perspective, present mechanisms of organizational learning facilitated by prototyping in the product and development context described. How these mechanisms have acted within and between the development cultures, proposes generated and shared tacit knowledge as an essential output of prototype-driven development projects. Furthermore, the chapter provides insights into the strategic utilization of wayfaring and how it has been implemented in Laerdal Medical. Finally, this chapter discusses the limitations and areas for further research and development considering the product, development and methodological contexts described.

6.1 Prototyping to Generate, Share and Capture Knowledge

This section views prototyping to generate, share, and capture knowledge by relating obtained insights from chapter 4, to the dynamic model of knowledge creation as proposed by Nonaka, (1994). Thus, considering the internalization, socialization, and externalization of knowledge through prototyping. As development in the FFE inherently requires decisions to be made on a limited, informal, and ambiguous foundation, tacit knowledge is important to understand and leverage. This section arguments that prototyping to learn expands knowledge domains of creators, suggesting prototyping to facilitate tacit knowledge being shared, captured, and re-used in, and across, development cultures (O. Eris & Leifer, 2004; Leifer & Steinert, 2011).

6.1.1 Generating new Knowledge by Prototyping

In product development, we intend to expand our tacit and explicit knowledge by obtaining new insights, thus extending into the unknown domain, as seen in Figure 6.1. In the cases this has included exploring opportunities, problems, solutions, or affirming solutions and the problems they address. To gain insights into these unknowns we can leverage different prototyping strategies and prototype output as described in chapter 4. Prototypes are based on explicit knowledge, i.e., problem prompts, contextual knowledge, disciplinary rules, and engineering principles. Moreover, in creating prototypes, we experiment, reflect, and utilize tacit knowledge, i.e., skills, experiences, problem understanding, and inspiration (Schön, 2017). Prototyping as a learning activity, employs and enriches tacit knowledge (internalization) embedded in the artefacts by *how* and *why* they are created (Nonaka, 1994).

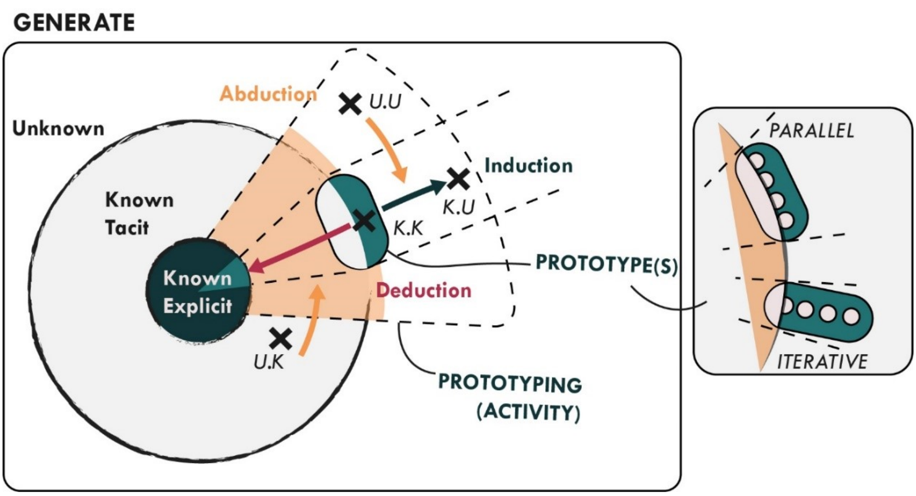


Figure 6.1 Knowledge generation through prototyping to facilitate reasoning.

While prototypes are tools to affirm or discard formalized ideas (by making them tangible), prototypes and how we use them, in prototyping, can spawn diverse insights and motivate theories. Theories are formed from assumptions and reasoning, thus progressing beyond what the prototype manifests and concerns gaining vision and insights on what a product might become. For example, evaluating prototype performance according to our assumptions through deductive reasoning. In projects this includes testing prototypes against design questions, which can inform decisions and establish known knowns (K.K). However, explicit insights depend on the features of prototypes (what we can measure and evaluate), thus suggesting only a small unknown domain could be targeted through deductive reasoning alone.

From the projects, more insights have been found to originate from prototyping activity, i.e., the interactions, internal testing, experimenting, to stimulate inductive and abductive reasoning (Gerstenberg et al., 2015; Leifer & Steinert, 2011). For example, can known unknowns (K.U) which are problems, opportunities, or solutions, become apparent by inductive reasoning. These unknowns are accessible because they appear within our understanding of the problem and can thus be addressed by enhancing prototype characteristics through iterations or more focused prototypes. This uncovering of unknowns can also reveal unknown knowns, meaning problems we can solve but were unaware of their relevance to the task in question.

Unknown unknowns (U.U), are obstacles or opportunities that need to be encountered before we can address them, and thus become known unknowns as soon as they are revealed (M. B. Jensen et al., 2017; Sutcliffe & Sawyer, 2013). U.U can be elicited by abductive reasoning or serendipity, thus bringing these into the viewport of our current development efforts. Abduction is described as lateral extension of reasoning by the use of metaphors (Bateson, 1979), which is purposeful when we have no adequate representation of design idea or lack a product vision. Hence, prototyping can facilitate abduction by contributing to creating something completely new, by sparking inspiration or novel concept ideas from prototypes. The directional reasoning through induction and deduction is used to revise or develop these concept ideas further (Nonaka, 1994).

As exemplified by observations from the various projects, different strategies and prototype outputs yield different outcomes regarding gained insights. These are differentiated by prototyping iteratively to pursue known unknowns by gaining a deeper understanding of problems located further away from our known domains. This suggests more refined and focused prototypes being deployed. Conversely, by utilizing parallel efforts and less refined prototypes, broader exploration and, thus, utilization and generation of more diverse tacit knowledge could be achieved. Broad and divergent exploration is purposeful to increase chances for serendipitous findings. As shown in project P10, prominently producing more novel findings before the solution space is confined and development efforts focus on a defined product vision. These findings support that internalization can be facilitated by early and exploratory prototyping to generate tacit knowledge and encourage serendipitous findings.

6.1.2 Transferring Knowledge by Prototyping

While knowledge is generated by individuals, tacit knowledge can transfer between individuals facilitated by prototyping as seen in Figure 6.2. This knowledge can entail both contextual and domain knowledge or experiences (external) —or tools, methods, and techniques deployed to build prototypes (internal). Strategically, sharing and accumulating such tacit knowledge can enrich the toolbox of available methods and know-how for generating prototypes. Thus, it can be leveraged in new contexts, use cases, or design problems if shared and obtained across the development culture.

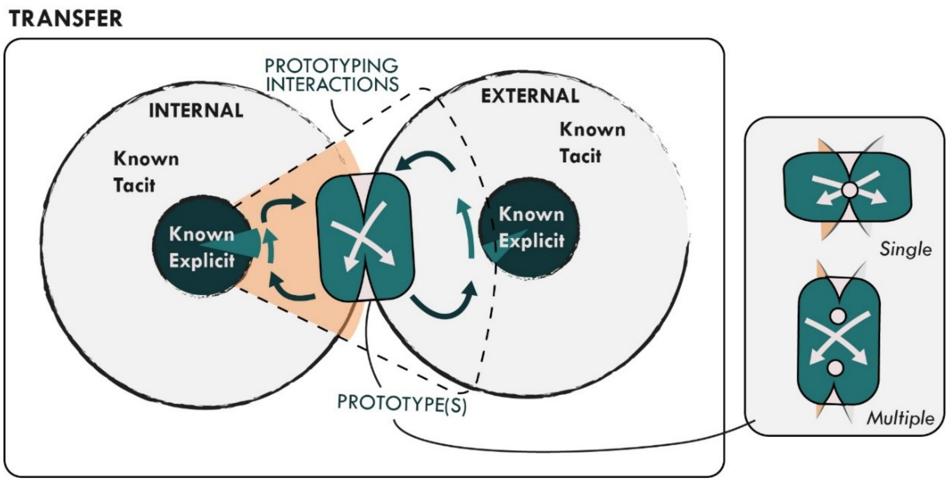


Figure 6.2 Knowledge transfer through prototype interactions.

The case projects highlight users interacting with prototypes and transferring informal domain and procedural knowledge of medical procedures and interventions. This can be complex, experience-based knowledge that is abstracted and translated into engineering terms via prototypes. Such transfer of tacit knowledge benefits different strategies and, thus, prototype characteristics. For example, will lower fidelity and multiple prototype representations enable a broader interface to bridge disciplinary gaps (evaluating the idea). In comparison, higher fidelity and focused prototypes can gain deeper insights (evaluating the prototype) but are less likely to access as diverse tacit insights. However, tacit knowledge might be required to be made tangible, which could be facilitated by gaining first-hand experiences (learning by doing). This involves contextualizing knowledge within a new domain and make it actionable, which has been essential when exploring new opportunities, needs, and problems. These findings support that reflective prototyping activities enables tacit knowledge to be shared, received, and understood, thus facilitating internalization and socialization (Nonaka, 1994).

Similarly, skills or design inspiration can be transferred through interactions between individuals discussing, showing, and interacting with prototypes in development cultures. This facilitates tacit knowledge being both transmitted and received, which can enhance both alignment within project teams, but furthermore foster adoption of tools, methods, approaches, and techniques outside the project group. This can be seen as inspiration spreading within (or between) organizations, but just as important, skill sharing and learning. Generative prototyping can contribute this learning by increasing the amount of knowledge generated within prototyping cultures. E.g., development sprints can be made solely to learn a new machine or to explore a novel technology and its potential applicability for future prototypes. Sharing could, thus, enhance prototyping capabilities of organizations (Schrage, 1993), and underline transparent development activities, sharing not only finalized products but learnings and prototypes from ongoing projects C9 (Ege et al., 2022). These findings argue that prototype-driven cultures can gain benefits from the collective output of prototypes and knowledge by it being shared, received, enhanced, and re-used.

6.1.3 Capturing Knowledge by Prototyping

Prototypes are tools to obtain insights and facilitate learning where attributes and generated knowledge have both tacit and explicit characteristics. For example, materials, size, visual traits, and functions are explicit and unambiguous attributes. However, filtering, framing, and manifesting, are tacit considerations made by developers explaining —what the prototype represents, what it intends to achieve, and how it performs accordingly. Unaware of these considerations, the prototype only conveys its explicit attributes, and tacit features remain ambiguous and need interpretation and imagination. Thus, to inform decisions or convey insights and knowledge beyond interactions within groups, cultures, or organizational boundaries, tacit knowledge needs to be captured.

Capturing tacit knowledge, i.e., experiences, findings, and insights, involves making them explicit by generating information and descriptive metadata through various lenses, as seen in Figure 6.3. This intends to capture enough information by sufficient detail for the tacit components to become self-explanatory and possible to re-use or adapt to new conditions or contexts. Why tacit knowledge needs to be captured however, depends on the projects, to whom it is addressed, and what the knowledge entails.

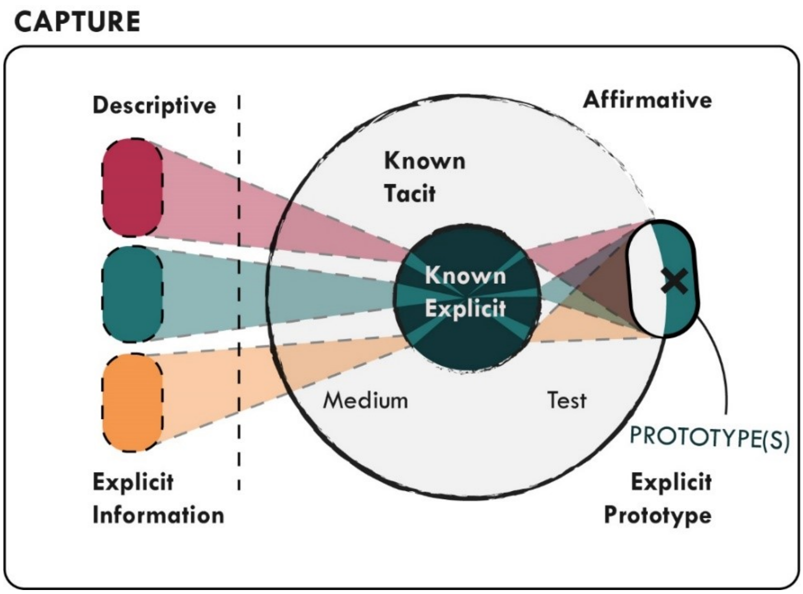


Figure 6.3 Knowledge capture through testing.

In the academic and research-focused projects performed at TrollLABS, the need for capturing insights obtained through prototyping and outcome is based on sharing results with an external audience and research community. This requires efforts to capture tacit knowledge from interactions, usability, realism, and functional abilities as exemplified by user-testing and lab experiments. This involves obtaining results from multiple sources, to gain a broad image of what the prototype represents, intends, and achieves. This is particularly important in, e.g., open-source hardware, where a greater community is also acting as an organization to adopt the project findings and the tacit learnings spanning beyond the technical blueprints of the design C9 (Ege et al., 2022).

Prototype tests, are moreover, important for the evaluation and ascertaining of functions to provide a broad foundation for decisions to be made (affirmative). This has included, e.g., evaluating sensor applications by objective and quantifiable metrics to describe or benchmark developed technology. Moreover, performing user tests for objective results to support evaluation and descriptions of informal and complex prototype interactions. This highlights how multifaceted test approaches could inform strategic decisions or further development by providing nuanced results from different viewpoints and sources. Testing considerations are critical for generating knowledge assets that supports and describe prototypes to be stored, shared, and re-used in organizations.

6.1.4 Knowledge as a Development Output

By looking at the interactions between project cases presented in this thesis, prototyping generate knowledge that exceeds what is presented in final concept proposal. This is new knowledge that is shared, received, and adopted into the development culture, i.e., skills, methods, and contextual knowledge that spread and accumulates. This supports development cultures in creating better prototypes, suggesting organizations become equipped to create better products. Examples of concrete knowledge transitions between projects in and between the Laerdal and TrollLABS development cultures can be seen in Figure 6.4. This provide concrete examples of mechanisms of learning in prototype-driven cultures (Schrage, 1993).

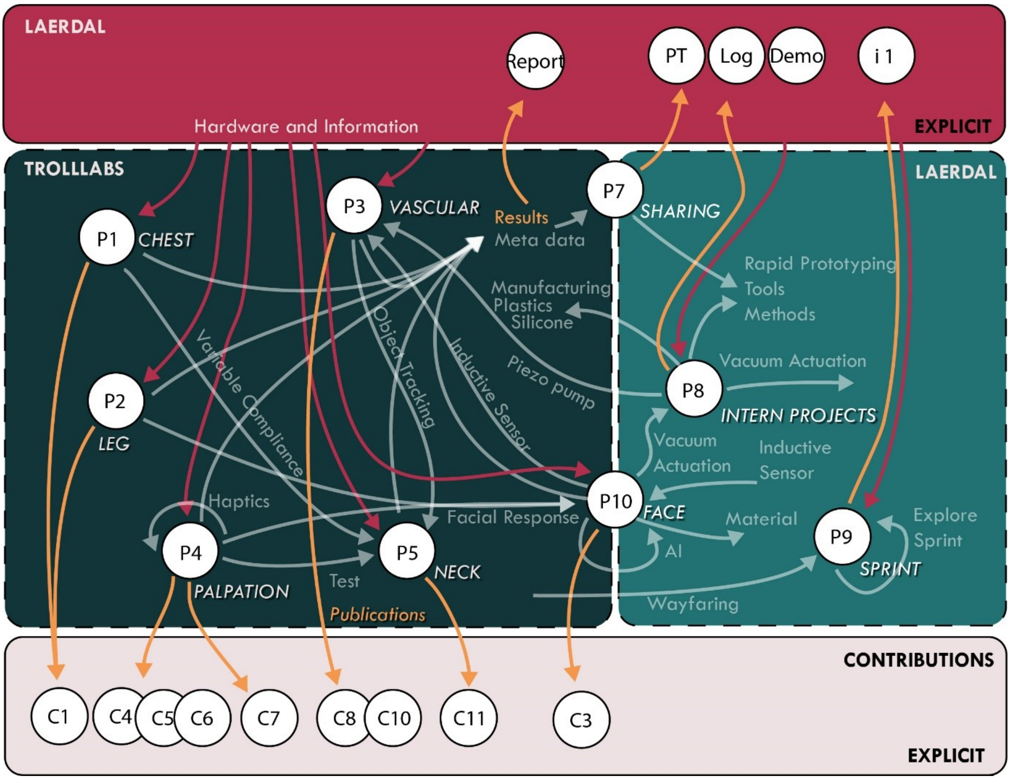


Figure 6.4 Knowledge generated, shared, and captured throughout this research project.

These findings, and the mechanisms of knowledge generation within (and between) development cultures by prototyping, the following considerations are presented:

- Tacit knowledge is applied in product development as skills, tools, and methods for prototyping, and is generated from designing, building, and testing prototypes.
- This tacit knowledge can be shared through prototyping interactions between individuals and can strengthen internal prototyping culture, or help interpreting tacit knowledge of external stakeholders.
- Ensuring transparency and sharing culture by integration events, sharing prototyping experiences and interactions between people and prototypes, can stimulate tacit knowledge generation, transfer, and re-use.
- Making tacit knowledge explicit, can be done through testing prototypes and should include creating descriptive representations and explicit metadata information through various lenses (multifaceted frameworks).

6.2 Using Wayfaring in Development Projects

As described in chapter 5, wayfaring is not based on strict formalized processes, but rather tools and guidelines suitable ambiguous problems encountered in the FFE. This section discusses two strategic advantages of using wayfaring, before highlighting how the approach has been tested and adopted by Laerdal Medical through the case projects. Firstly, using wayfaring in the projects has helped to obtain insights by efficient and effective prototyping. Secondly, wayfaring has facilitated eliciting serendipitous findings and unknown unknowns, but has moreover provided flexibility for these findings to be pursued and made tangible.

6.2.1 Efficient and Effective Prototyping by Wayfaring

As previously emphasized, time and timing through iterative development cycles is crucial in NPD in the FFE. This also includes project execution and process design, where time is associated with the risk of not meeting deadlines, not successfully solving the right problems, and consequently running up project costs. Subsequently, value can be considered the knowledge generated (in projects) and accumulated (in the organization), thus mitigating uncertainty and inherent risks. Hence, wayfaring should be utilized to ensure effective and efficient prototyping activities, to mitigate risk and maximize value.

Efficient prototyping is regarded as maximizing the value generated over the project period. Wayfaring contributes to doing rapid iterations and short sprints (probes) to generate multiple insights by failing early rather than later. As illustrated by Figure 6.5, this means manifesting prototypes purposefully by using low fidelity levels, through higher abstraction and lower resolution representations (collectively or interchangeably). These considerations ensure rapid sprints and multiple iterations, which furthermore enables different aspects of the eventual product to be filtered. Consequently, wayfaring ensures rapid and broad coverage of the development space, as described in chapter 5, and mitigates uncertainty by that numerous insights are gained early in the project workflows.

Effective prototyping on the other hand, relates to the quality of the insights we obtain from performing appropriate and diverse development activities. This means prototyping to reveal and inform the different aspects of an eventual product idea, through filtering different use-cases, physiological realism, and functional attributes. Prototypes should be manifested to accommodate the appropriate audience and intent, for tests or meaningful interactions to provide deeper insights. Opposingly, if prototypes are only created (and iterated) within a narrow filtering aspect, numerous yet insignificant insights might be gained. Therefore, wayfaring should ensure that broader coverage and deeper insights are obtained for informed designs to be established.

Collectively, these strategic considerations for wayfaring in projects are used for more quickly and with less risk, develop informed concepts that can be affirmed. Testing and ascertaining concepts and solutions, however, requires higher levels of fidelity concerning both prototypes and their testing conditions. Thus, the time spent on affirmative activities is often longer and associated with greater risks. However, the accumulated insights and preceding prototypes from wayfaring, provide redundancy. Meaning multiple alternative ideas, that can be brought forward if an idea is tested and deemed insufficient or a dead end.

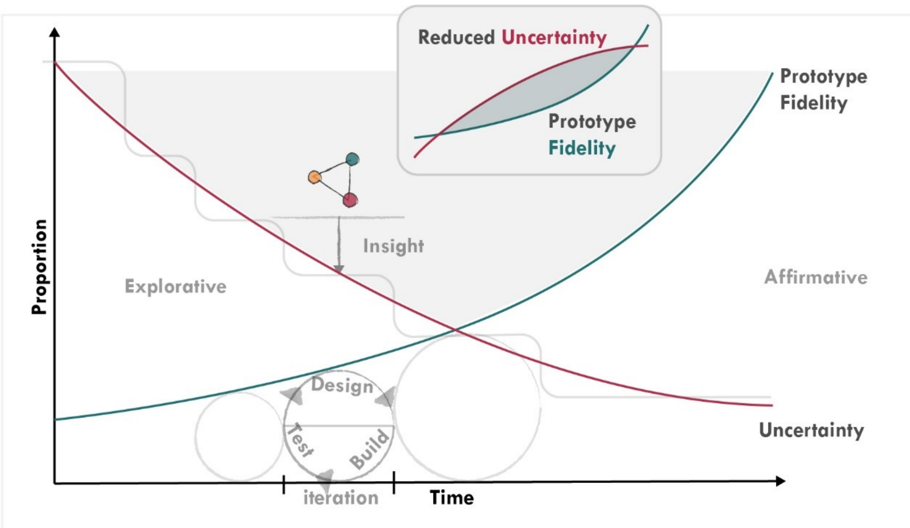


Figure 6.5 Mitigating uncertainty by obtaining insights by prototype-driven development.

6.2.2 Flexibility, Serendipity, and Unknown Unknowns

Deploying wayfaring in the FFE, we are faced with conditions that are ambiguous and complex. This makes is hard to plan for where the project will end up, and consequently which findings that will remain relevant as the project progress. This is especially the case with serendipitous findings and unknown unknowns, being problems, ideas or solutions that could have great potential, but they appear outside of our initial project scope and direction. Wayfaring could provide greater chances of eliciting these findings, by ensuring explorative and divergent efforts are made when uncertainty is high. Furthermore, the cost of investigating serendipitous findings remain low, as rapid probes can make them tangible and reveal their relation to the projects as either obstacles or opportunities.

Having this flexibility in early development efforts should be ensured on a strategic level, because findings outside of the initial problem prompt, can provide novel and significant opportunities. In the project cases serendipitous findings have pivoted the project plan by novel solutions being identified (P4 ferrogranular principle). Or they have enabled spin-off projects such as the inductance sensor identified in P10, developed and conceptualized in P3, and presented in C9 (Steffensen, Auflem, et al., 2022). Through the case projects and specifically exemplified by P10, the occurrence of unknown unknowns and serendipitous findings, are more frequent in the earliest stages where prototypes remain crude and before the development space gets confined. However, as design sprints can (and should) be kept short and agile, also later findings can be pursued and evaluated. This suggests that wayfaring and strategic prototype-driven development can reveal potentials or obstacles through serendipity. Furthermore, these can greatly benefit projects or the organization at large, through knowledge for novel technology, capabilities, or product opportunities, being generated, explored, and shared.

6.2.3 Discussion on Implementing Wayfaring

To encourage prototype-driven development and adoption of wayfaring within Laerdal, three activities described by the cases P7, P8 and P9 has been carried out throughout this research project. This was despite prototypes already having a critical role in current development practice in the organization and were used to inform decisions (gate reviews) and convey information throughout projects (alignment and user-tests). However, questions on how to improve the pre-requirements initiatives and foster broader adoption and utilization of the development tools and approaches presented in this thesis has been stressed. Moreover, how we could implement these modes of working in project and process workflows has been (and still is), a matter discussed and wished for by the development organization.

Firstly, the workshops and sharing sessions (P7) carried out throughout the length of this research project, have exemplified and shared prototyping tools, methods, and mindset between TrollLABS and Laerdal Medical. These interactions have furthermore fostered a wider adoption (beyond engineers and developers) of the user-centred, question focus, and prototype-driven approaches. E.g., a workshop with the materials department revealed that wayfaring and prototype-driven approaches could be beneficial to use in the pursuit and testing of novel materials for future products.

Secondly, the intern technology projects (P8) have provided an arena for trying out ideas for running and managing projects, and furthermore integrate these within the organization. The projects have shown how inexperienced interns dwelling in uncharted development territories, could quickly (8 weeks) obtain significant insight and broad problem understanding. The projects have been showcased by show and tell presentations after completion. But more importantly been required to share and interact with representatives from different branches of the development organization on a weekly basis. The agile and rapid development modes have generated novel solutions and revealed novel problem spaces for the organization. This has in several of the projects both challenged, and informed, the future product visions (and opportunities) for the product portfolio of Laerdal Medical.

Lastly, the product development sprints (P9) was initiated, run, and reviewed to share methods and findings broadly across the organization, as seen in the internal paper i1. While this only concerns a single 2-week project sprint, it showcased potentials and benefits of adopting question-focus and prototype-driven development facilitated by wayfaring. Consequently, the way of working, allocating a limited period in the front-end of new projects for “concept discovery” has been both adopted and utilized in other projects (lastly fall 2022). This has not only shown the applicability of the methodology, but regarding the outcomes of these initiatives has revealed that it fits the purpose also beyond the education and research setting at TrollLABS.

6.3 Discussing Limitations and Further Research Opportunities

6.3.1 Medical Training Implications, Validity, and Grounding

While considering multiple fields, disciplines, and their interfaces, this thesis is grounded in engineering design. Consequently, the research is limited by a one-sided approach where the implications for medical training equipment are based on empirical insights and knowledge obtained through the cases alone. Subsequently, the developed solutions have yet to be tested in clinical experiments where transfer learning effects and clinical performance are gauged. This suggests the research could have benefitted a more diverse grounding by including researchers from other fields, such as healthcare education. This could have enlightened transfer of learning from using the equipment or clinical skill improvement, thus strengthening the argument and validity of the various concepts and results. However, the results and equipment presented could be valuable starting points for areas worth researching further. Moreover, characteristics and functionality provided by the equipment by sensors, adaptability, realism, and ecological validity considerations suggest these as platforms for further research to be conducted. Exemplified by pilot experiments in the contributions, opportunities for further development and hypotheses could facilitate further research in medical training and medical education domains.

6.3.2 Research Methods and Results

The gathered and synthesized insights on prototype-driven development and wayfaring, are not only obtained from either academic or professional practice, but they are viewed in unison. Hence, providing examples on how research projects could generate knowledge to be integrated in professional settings. While not specifically an aim for this thesis, this accounts for a benefit and further opportunity for why industrial research projects in NPD are purposeful, and how findings from further research projects could help bridge the gap between engineering design research and engineering design practice.

Concerning the methods and results presented in this thesis it is evident that a designer and researcher bias is unavoidable given the participatory roles as coach and main contributor assumed by the author in all the cases. Moreover, a question on the causal effects of methods and results could be made, questioning whether different methods on both operational, tactical, and strategical levels could have yielded similar results. The results are moreover obtained through empirical observations, where most are gathered from chaotic projects utilizing wayfaring. This could question both the repeatability and objective truthfulness of the observations and synthesized findings. However, results and findings show coherence by the tools, methods, and approaches, despite the uniqueness across every case regarding problem, involved people, stakeholders, and solutions. This strengthens the argument for applicability of the tools and methodology presented, and moreover suggests it to be applicable in other domains. However, this is an area for further research both concerning implementation effects, implications, and applicability in other development contexts.

6.3.3 The Role of the Developer, Coach, and Researcher

Triple loop learning, in this thesis, involves different organizational levels and, thus, has required different viewpoints and perspectives. The author's roles have enabled these learning loops as developer, coach, and researcher. The developer role has involved deploying prototyping activities, reflecting on outcomes, and synthesizing insights. In a coaching role, contradictions or coherence in these insights has been identified by observing and supporting development projects, thus from a different perspective. Findings and approaches are proposed by operationalizing these insights within a wayfaring framework for managing projects.

These synthesized findings have laid the ground for discussing the strategic organizational benefits and the third learning loop. However, as strategic advantages span beyond project success or novel findings (addressed by the two first goals of this thesis), this loop and the learnings therein are proposed areas for further research. The author argues for a similar role to be applied within organizational structures (e.g., Laerdal Medical), thus enabling a bottom-up approach informing development processes. This could include prototype-driven working modes and processes informing management of FFE projects. Hence, deriving and informing development strategy based on prototyping practice and how this contributes to novel outcomes and successful projects.

6.3.4 Further Research and Development Opportunities for Laerdal Medical

Identified product and technology opportunities for future simulation-based medical training equipment are described by the findings in chapter 4. How these could be leveraged is however, governed by additional factors such as future strategic needs, cost of integration or launch, and business strategies. This section is therefore aimed at how suggested development activities could be leveraged and why methods and approaches presented in this thesis should be considered by Laerdal Medical.

Close access to users (EHP) has in the projects been emphasized, to obtain relevant and tangible insights. These insights are critical for developing the required functionality and could be gained by manifesting prototypes of adequate fidelity along the dimensions in question. The following opportunities and suggestions follow the findings of this thesis:

- **Early and tangible interactions**. Gaining access to and interact with users has shown important to understand the product needs, and spark inspiration for novel solutions.
- **Testing beyond validation and verification**. Entails testing multiple aspects of the eventual product and obtain relevant and tangible results that could be utilized to inform further development.
- **Access and space to facilitate interactions**. Access means contact with relevant healthcare personnel that, in person, can contribute to projects. Space means location and facilities where interactions can happen, for example TrollLABS medical (P6). This should be a flexible space that facilitates test setups, co-development, and rapid and unplanned interactions.

Technology and novel concepts are not only generated because of needs and suiting applications being found. Serendipitous findings have been identified to spread between individuals, projects, and in this research project, organizations. These are often a result of tacit knowledge and inspiration spreading as prototypes are tested and shared within the culture they are made. Hence, the following suggestions are made to facilitate these benefits and effects:

- **Prototype-driven and explorative**. Large output of prototypes and insights, thus generating and making tacit knowledge tangible for others. Facilitated by question-focus and wayfaring as described in chapter 5.
- **Sharing culture and transparency**. Regular sharing sessions where show is emphasized over tell. Using scrum principles where day-to-day and week-to-week sprints are used to communicate and share findings within and between teams and stakeholders.
- **Multi- and cross-disciplinary integration**. Teams that embark on early-phase exploration should consist of multiple disciplines. Different disciplines should be encouraged to contribute to early exploration activities where seen purposeful.
- **Utilize new technology in prototyping**. To make an impact in future products, new technology could be developed and explored (opportunistically) by prototyping. For example, as AI and machine vision applications was utilized to aid the development of a new face in P10 (Described by C3). This opened doors for similar approaches where the technology not only contributes to prototyping, but also to the actual functionality of the envisioned product.

This thesis has showcased a sample of development methods and approaches, yet this is merely a snapshot of possibilities. Moreover, as stressed, individual problems and projects entail different needs and could leverage different methods for and by prototyping. However, what is clear from this thesis, is how the mutual and widespread benefits of activities being performed in a prototyping culture enhance the overall throughput of design ideas and novel findings. By utilizing the methods showcased in this thesis, leveraging the fundamental principles for deciding the prototyping approach, and managing these through wayfaring, enhancement of said culture can (and probably will) happen. Hence, increased prototyping would be a considerable development opportunity in itself to pursue, —and enhancing our knowledge about what it entails and contributes would be valuable to research further.

7

Final Reflections

The three goals of this research project have been to develop concepts for new healthcare training applications, novel technology to facilitate the realization of these applications, and enhance methods and process knowledge for developing them. Subsequently, the project has generated research output within three categories; 10 project contributions (used as cases), 11 academic research contributions, and this thesis. The thesis synthesizes insights and findings to holistically describe *what*, *how* and *why* prototyping and prototype-driven development yield project insights and new knowledge. The author emphasizes these learnings through a triple loop learning approach where prototyping is considered —within design activities to facilitate insights, within projects to progress development, and within organizations to facilitate the generation, sharing, and capture of tacit knowledge. Thus, through the collective contribution, this research emphasizes the generative, communicative, and learning benefits of prototyping across different organizational levels. This is especially considered within the scope of this thesis, where the complexity of interactions, multiple fields, disciplines, and knowledge domains requires integration and alignment. This suggests excellent opportunities for utilizing, facilitating, and leveraging prototyping in similar endeavours, and moreover, how more research could benefit the product, development, and methodological context.

This thesis, the academic contributions, and case projects described, concerns an industrial research project between Laerdal Medical and TrollLABS NTNU. Such industrial PhD projects in NPD organizations and adjacent research fields enables establishing collaborative research outposts. This provides great opportunities as well as challenges. The initiation of similar projects requires mutual integration of the project in both organizations, i.e., university and company. Arguably the requirements for this integration to be successful is that both parties provide sufficient playroom and eases the collaboration between and within both development cultures. The author highlights collaborating across research projects and domains within TrollLABS as a success factor. This has enabled serendipitous discoveries and inspiration to emerge, not only from within projects, but across multiple projects and activities within the same confined development culture. Furthermore, the close localization and integration within the hospital campus, has ensured readily access to users and experienced practitioners required to inform the development projects. Equally important, the integration and playroom provided by Laerdal Medical, and opportunities enabled by the vast and highly skilled technology and development departments therein, has been crucial for this project. As transit between locations has inhibited continuous and rapid interactions, the transparent and tangible properties of prototype-driven projects have been important. These integrations have enhanced capacities and skills, of both the author and the project at large, thus accelerating the prototype-progress and conceptual knowledge required to facilitate the multitude of project cases and contributions described in this thesis.

Significant benefits and further research opportunities of industrial PhD projects are the duality of flexibility enabled in research, and the real-world applications and challenges provided by the company context. These benefits, and the previously described opportunities and limitations, could moreover be progressively taken advantage of by scaling up the research operation. Specifically, the author suggests a tri-core research project, using the filtering dimension of medical training equipment, described in chapter 4 and 5, as a starting point. This suggestion requires a larger team to account for areas of expertise limited or revealed purposeful for further research by this project and the contributions. These suggestions and what they could entail are visualized in Figure 7.1.

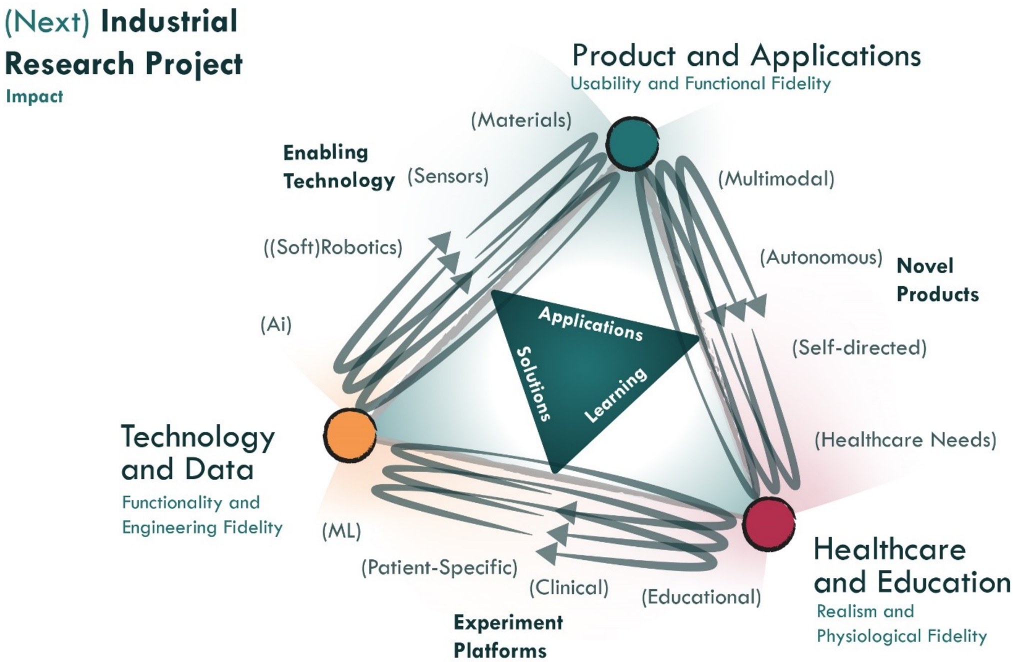


Figure 7.1 Recommendations and opportunities for next industrial research project.

- **Novel Products**: Needs, opportunities, and solutions to be explored and exploited by researching new products and application in the healthcare and education domain.
- **Experiment Platforms:** Enhance knowledge and impact by tests, experiments, and results from simulated scenarios using experimental product platforms.
- **Enabling Technology:** Develop new technology that can enhance simulation-based medical product functionality, usability, and realism.

The figure is underlining the interface between core research areas of this PhD to be opportunities for further research. For the eventual next industrial research project, wayfaring approaches can support the overarching research directions in conjunction with managing individual development projects therein. Prototype-driven methods can leverage opportunities and generate tangible insights to connect opportunities and research areas. Furthermore, to enhance our understanding and continuously re-evaluate and enhance what prototyping contributes, how we facilitate it, and the benefit for organizational deployment by triple loop learning is an advantageous research approach. Hence, similar approaches are encouraged to enhance the empirical and methodological foundation on which future innovations within healthcare training will be prototyped.

References

Aggarwal, R., Mytton, O. T., Derbrew, M., Hananel, D., Heydenburg, M., Issenberg, B., MacAulay, C., Mancini, M. E., Morimoto, T., Soper, N., Ziv, A., & Reznick, R. (2010). Training and simulation for patient safety. BMJ Quality & Safety, 19(Suppl 2), i34–i43. https://doi.org/10.1136/qshc.2009.038562

Alinier, G. (2007). A typology of educationally focused medical simulation tools. Medical Teacher, 29(8), e243–e250. https://doi.org/10.1080/01421590701551185

Aalto, P. (2022). TrollLABS space [3D scanning].

Argyris, Ch., & Schön, D. A. (1997). Organizational Learning: A Theory of Action Perspective. Reis, 77/78, 345–348. https://doi.org/10.2307/40183951

Atzberger, A., & Paetzold, K. (2019). Current Challenges of Agile Hardware Development: What are Still the Pain Points Nowadays? Proceedings of the Design Society: International Conference on Engineering Design, 1(1), 2209–2218. https://doi.org/10.1017/dsi.2019.227

Auflem, M., Boe, H. H., Erichsen, J. F., & Steinert, M. (2020). ON PROTOTYPING METHODS TO LEVERAGE NON-RIGID MATERIALS IN THE EARLY STAGES OF ENGINEERING DESIGN. Proceedings of the Design Society: DESIGN Conference, 1, 2445–2454. https://doi.org/10.1017/dsd.2020.120

Auflem, M., Erichsen, J. F., & Steinert, M. (2019). Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators. Procedia CIRP, 84, 572–578. https://doi.org/10.1016/j.procir.2019.04.202

Auflem, M., Kohtala, S., Jung, M., & Steinert, M. (2022). Facing the FACS—Using AI to Evaluate and Control Facial Action Units in Humanoid Robot Face Development. Frontiers in Robotics and AI, 9, 887645. https://doi.org/10.3389/frobt.2022.887645

Bateson, G. (1979). Mind and nature.

Batliner, M., Boës, S., Heck, J., & Meboldt, M. (2022). Linking Testing Activities with Success in Agile Development of Physical Products. Procedia CIRP, 109, 146–154. https://doi.org/10.1016/j.procir.2022.05.228

Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2020). Are Ideas Getting Harder to Find? American Economic Review, 110(4), 1104–1144. https://doi.org/10.1257/aer.20180338

Bogers, M., & Horst, W. (2014). Collaborative Prototyping: Cross-Fertilization of Knowledge in Prototype-Driven Problem Solving. Journal of Product Innovation Management, 31(4), 744–764. https://doi.org/10.1111/jpim.12121

Brown, T., & Wyatt, J. (2010). Design Thinking for Social Innovation. Development Outreach, 12(1), 29–43. https://doi.org/10.1596/1020-797X_12_1_29

Bryan-Kinns, N., & Hamilton, F. (2002). One for All and All for One?: Case Studies of Using Prototypes in Commercial Projects. Proceedings of the Second Nordic Conference on Human-Computer Interaction, 91–100. https://doi.org/10.1145/572020.572032

Buchenau, M., & Suri, J. F. (2000). Experience Prototyping. 424–433. https://doi.org/10.1145/347642.347802

Camburn, B. A., Dunlap, B. U., Kuhr, R., Viswanathan, V. K., Linsey, J. S., Jensen, D. D., Crawford, R. H., Otto, K., & Wood, K. L. (2013). Methods for Prototyping Strategies in Conceptual Phases of Design: Framework and Experimental Assessment. Volume 5: 25th International Conference on Design Theory and Methodology; ASME 2013 Power Transmission and Gearing Conference, V005T06A033. https://doi.org/10.1115/DETC2013-13072

Carlile, P. R. (2002). A Pragmatic View of Knowledge and Boundaries: Boundary Objects in New Product Development. Organization Science, 13(4), 442–455. https://doi.org/10.1287/orsc.13.4.442.2953

Clarkson, J., & Eckert, C. (2010). Design Process Improvement: A review of current practice. Springer Science & Business Media.

Dow, S. P., Glassco, A., Kass, J., Schwarz, M., Schwartz, D. L., & Klemmer, S. R. (2011). Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. ACM Transactions on Computer-Human Interaction (TOCHI), 17(4), 18:1-18:24. https://doi.org/10.1145/1879831.1879836

Edelman, J., & Currano, R. (2011). Re-representation: Affordances of Shared Models in Team-Based Design. In Design Thinking (pp. 61–79). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-13757-0_4

Edelman, J., Leifer, L., Banerjee, B., Sonalkar, N., Jung, M., & Lande, M. (2009). Hidden in Plain Sight: Affordances of Shared Models in Team Based Design. https://www.designsociety.org/publication/28602/Hidden+in+Plain+Sight%3A+Affordances+of+Shared+Models+in+Team+Based+Design

Ege, D. N., Auflem, M., Lilleløkken, O., & Steinert, M. (2021). User involvement in early-stage design of medical training devices – case of a palpation task trainer prototype. Design for Health, 5(2), 214–232. https://doi.org/10.1080/24735132.2021.1980270

Ege, D. N., Auflem, M., & Steinert, M. (2022). Lost in Transit: Implications and Insights for Making Medical Task Trainer Prototypes with an Open Source Hardware Paradigm. Proceedings of the Design Society, 2, 1253–1262. https://doi.org/10.1017/pds.2022.127

Ege, D. N., Lilleløkken, O., Auflem, M., & Steinert, M. (2020). Dealing with Ecological Validity and User Needs when Developing Simulation Based Training Equipment – Case Study of a Medical Palpation Task Trainer. Procedia CIRP, 91, 722–727. https://doi.org/10.1016/j.procir.2020.03.116

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. Academy of Management Review, 14(4), 532–550. https://doi.org/10.5465/amr.1989.4308385

Elverum, C. W., & Welo, T. (2015). On the use of directional and incremental prototyping in the development of high novelty products: Two case studies in the automotive industry. Journal of Engineering and Technology Management, 38(Supplement C), 71–88. https://doi.org/10.1016/j.jengtecman.2015.09.003

Erichsen, J. A. B., Pedersen, A. L., Steinert, M., & Welo, T. (2016). Using prototypes to leverage knowledge in product development: Examples from the automotive industry. 2016 Annual IEEE Systems Conference (SysCon), 1–6. https://doi.org/10.1109/SYSCON.2016.7490586

Erichsen, J. F., Wulvik, A., Steinert, M., & Welo, T. (2019). Efforts on Capturing Prototyping and Design Activity in Engineering Design Research. Procedia CIRP, 84, 566–571. https://doi.org/10.1016/j.procir.2019.04.303

Eris, Ö. (2003). Asking generative design questions: A fundamental cognitive mechanism in design thinking. DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design, Stockholm, 587-588 (exec. summ.), full paper no. DS31_1640FP.

Eris, O., & Leifer, L. (2004). Facilitating Product Development Knowledge Acquisition: Interaction between the Expert and the Team. International Journal of Engineering Education, 19, 142–152.

Gerstenberg, A., Sjöman, H., Reime, T., Abrahamsson, P., & Steinert, M. (2015). A Simultaneous, Multidisciplinary Development and Design Journey – Reflections on Prototyping. 409–416. https://doi.org/10.1007/978-3-319-24589-8_33

Glaser, B. G., & Strauss, A. L. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine.

Hansen, C. A., & Özkil, A. G. (2020). From Idea to Production: A Retrospective and Longitudinal Case Study of Prototypes and Prototyping Strategies. Journal of Mechanical Design, 142(3). https://doi.org/10.1115/1.4045385

Henriksen, E. M., Auflem, M., & Steinert, M. (2023). A pain in the neck: Prototyping and testing of a patient simulator neck for spinal immobilization training. Design for Health, 7(1), 42–63. https://doi.org/10.1080/24735132.2023.2184995

Houde, S., & Hill, C. (1997). What do Prototypes Prototype? In Handbook of Human-Computer Interaction (pp. 367–381). Elsevier. https://doi.org/10.1016/B978-044481862-1.50082-0

Issenberg, S. B., McGaghie, W. C., Hart, I. R., Mayer, J. W., Felner, J. M., Petrusa, E. R., Waugh, R. A., Brown, D. D., Safford, R. R., Gessner, I. H., Gordon, D. L., & Ewy, G. A. (1999). Simulation Technology for Health Care Professional Skills Training and Assessment. JAMA, 282(9), 861–866. https://doi.org/10.1001/jama.282.9.861

Issenberg, S. B., Mcgaghie, W. C., Petrusa, E. R., Gordon, D. L., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review. Medical Teacher, 27(1), 10–28. https://doi.org/10.1080/01421590500046924

Jensen, L. S., Özkil, A. G., & Mortensen, N. H. (2016). Prototypes in engineering design: Definitions and strategies. DS 84: Proceedings of the DESIGN 2016 14th International Design Conference. https://www.designsociety.org/publication/38892/PROTOTYPES+IN+ENGINEERING+DESIGN%3A+DEFINITIONS+AND+STRATEGIES

Jensen, M. B., Elverum, C. W., & Steinert, M. (2017). Eliciting unknown unknowns with prototypes: Introducing prototrials and prototrial-driven cultures. Design Studies, 49(Supplement C), 1–31. https://doi.org/10.1016/j.destud.2016.12.002

Johnson, M. P., Ballie, J., Thorup, T., & Brooks, E. (2017). Living on the Edge: Design artefacts as boundary objects. The Design Journal, 20(sup1), S219–S235. https://doi.org/10.1080/14606925.2017.1352771

Kelley, T. (2001). Prototyping is the shorthand of innovation. Design Management Journal (Former Series), 12(3), 35–42. https://doi.org/10.1111/j.1948-7169.2001.tb00551.x

Kim, J., & Wilemon, D. (2002). Focusing the fuzzy front–end in new product development. R&D Management, 32(4), 269–279. https://doi.org/10.1111/1467-9310.00259

Koen, P. A., Ajamian, G. M., Boyce, S., Clamen, A., Fisher, E., Fountoulakis, S., Johnson, A., Puri, P., & Seibert, R. (2002). Fuzzy front end: Effective methods, tools, and techniques. The PDMA Toolbook 1 for New Product Development.

Koen, P., Ajamian, G., Burkart, R., Clamen, A., Davidson, J., D'Amore, R., Elkins, C., Herald, K., Incorvia, M., Johnson, A., Karol, R., Seibert, R., Slavejkov, A., & Wagner, K. (2001). Providing Clarity and A Common Language to the "Fuzzy Front End". Research-Technology Management, 44(2), 46–55. https://doi.org/10.1080/08956308.2001.11671418

Kriesi, C., Blindheim, J., Bjelland, Ø., & Steinert, M. (2016). Creating Dynamic Requirements through Iteratively Prototyping Critical Functionalities. Procedia CIRP, 50(Supplement C), 790–795. https://doi.org/10.1016/j.procir.2016.04.122

Krishnakumar, S., Berdanier, C., Lauff, C., McComb, C., & Menold, J. (2022). The story novice designers tell: How rhetorical structures and prototyping shape communication with external audiences. Design Studies, 82, 101133. https://doi.org/10.1016/j.destud.2022.101133

Kushniruk, A., Nohr, C., Jensen, S., & Borycki, E. M. (2013). From Usability Testing to Clinical Simulations: Bringing Context into the Design and Evaluation of Usable and Safe Health Information Technologies: Contribution of the IMIA Human Factors Engineering for Healthcare Informatics Working Group. Yearbook of Medical Informatics, 22(01), 78–85. https://doi.org/10.1055/s-0038-1638836

Laerdal stock images. (2022). [Photograph]. www.laerdal.com

Lande, M., & Leifer, L. (2009). Prototyping to learn: Characterizing engineering student's prototyping activities and prototypes. DS 58-1: Proceedings of ICED 09, the 17th International Conference on Engineering Design, 1.

Lauff, C., Kotys-Schwartz, D., & E. Rentschler, M. (2018). What is a Prototype? What are the Roles of Prototypes in Companies? Journal of Mechanical Design, 140. https://doi.org/10.1115/1.4039340

Lee, J. H., & Ostwald, M. J. (2022). The relationship between divergent thinking and ideation in the conceptual design process. Design Studies, 79, 101089. https://doi.org/10.1016/j.destud.2022.101089

Leifer, L. J., & Steinert, M. (2011). Dancing with ambiguity: Causality behavior, design thinking, and triple-loop-learning. Information Knowledge Systems Management, 10(1–4), 151–173. https://doi.org/10.3233/IKS-2012-0191

Leimeister, J. M., Dickhaut, E., & Janson, A. (2021). Design Pattern as a Bridge Between Problem-Space and Solution-Space. In S. Aier, P. Rohner, & J. Schelp (Eds.), Engineering the Transformation of the Enterprise: A Design Science Research Perspective (pp. 137–150). Springer International Publishing. https://doi.org/10.1007/978-3-030-84655-8_9

Lilleløkken, O., Ege, D. N., Auflem, M., & Steinert, M. (2020). Proposing a novel approach for testing complex medical task trainer prototypes. DS 101: Proceedings of NordDesign 2020, Lyngby, Denmark, 12th - 14th August 2020, 1–12. https://doi.org/10.35199/NORDDESIGN2020.57

Lim, Y.-K., Stolterman, E., & Tenenberg, J. (2008). The Anatomy of Prototypes: Prototypes As Filters, Prototypes As Manifestations of Design Ideas. ACM Trans. Comput.-Hum. Interact., 15(2), 7:1-7:27. https://doi.org/10.1145/1375761.1375762

Linsey, J. S., Tseng, I., Fu, K., Cagan, J., Wood, K. L., & Schunn, C. (2010). A Study of Design Fixation, Its Mitigation and Perception in Engineering Design Faculty. Journal of Mechanical Design, 132(4). https://doi.org/10.1115/1.4001110

Maran, N. J., & Glavin, R. J. (2003). Low- to high-fidelity simulation – a continuum of medical education? Medical Education, 37, 22–28. https://doi.org/10.1046/j.1365-2923.37.s1.9.x

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2006). Effect of practice on standardised learning outcomes in simulation-based medical education. Medical Education, 40(8), 792–797. https://doi.org/10.1111/j.1365-2929.2006.02528.x

Medicine, I. of, & America, C. on Q. of H. C. in. (2000). To Err Is Human: Building a Safer Health System. National Academies Press.

Meinel, C., Leifer, L., & Plattner, H. (Eds.). (2011). Design Thinking. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-13757-0

Meller, G. (1997). A typology of simulators for medical education. Journal of Digital Imaging, 10(1), 194–196. https://doi.org/10.1007/BF03168699

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

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. Academic Medicine, 65(9), S63-7.

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. Organization Science, 5(1), 14–37.

O'Donovan, B., Eckert, C., Clarkson, J., & Browning, T. R. (2005). Design planning and modelling. Design Process Improvement, 60–87.

One Million Lives by Laerdal Medical: A Shared Goal for 2030. (2020, October 11). One Million Lives. https://one-million-lives.com/

Otto, K., & Wood, K. (2001). Product Design: Techniques in Reverse Engineering and New Product Development: Techniques in Reverse Engineering and New Product Development. https://research.aalto.fi/en/publications/product-design-techniques-in-reverse-engineering-and-new-product-

Petrakis, K., Hird, A., & Wodehouse, A. (2019). The Concept of Purposeful Prototyping: Towards a New Kind of Taxonomic Classification. Proceedings of the Design Society: International Conference on Engineering Design, 1(1), 1643–1652. https://doi.org/10.1017/dsi.2019.170

Petrakis, K., Wodehouse, A., & Hird, A. (2021). Prototyping-Driven Entrepreneurship: Towards a Prototyping Support Tool Based on Design Thinking Principles. The Design Journal, 24(5), 761–781. https://doi.org/10.1080/14606925.2021.1957531

Real, R., Snider, C., Goudswaard, M., & Hicks, B. (2022). Dimensions of Prototyping Knowledge: Characterising Prototype Evaluation Methods and Their Contributions to Design Knowledge. https://doi.org/10.13140/RG.2.2.22653.84961

Rhinow, H., Koeppen, E., & Meinel, C. (2012). Design Prototypes as Boundary Objects in Innovation Processes. DRS Biennial Conference Series. https://dl.designresearchsociety.org/drs-conference-papers/drs2012/researchpapers/116

Rørvik, S. B., Auflem, M., Dybvik, H., & Steinert, M. (2021). Perception by Palpation: Development and Testing of a Haptic Ferrogranular Jamming Surface. Frontiers in Robotics and AI, 8. https://www.frontiersin.org/articles/10.3389/frobt.2021.745234

Schön, D. A. (2017). The Reflective Practitioner: How Professionals Think in Action. Routledge. https://doi.org/10.4324/9781315237473

Schrage, M. (1993). The Culture(s) of prototyping. Design Management Journal (Former Series), 4(1), 55–65. https://doi.org/10.1111/j.1948-7169.1993.tb00128.x

Schwaber, K., & Sutherland, J. (2011). The scrum guide. Scrum Alliance, 21(19), 1.

Steffensen, T. L., Auflem, M., Vestad, H. N., & Steinert, M. (2022). Embedded Soft Inductive Sensors to Measure Arterial Expansion of Tubular Diameters in Vascular Phantoms. IEEE Sensors Journal, 22(7), 7240–7247. https://doi.org/10.1109/JSEN.2022.3155071

Steffensen, T. L., Kriesi, C., Steinert, M., Lafrenz, T., Brede, J. R., & Auflem, M. (2022). Ultrasound phantom (United States Patent No. US20220304922A1). https://patents.google.com/patent/US20220304922A1/en

Steinert, M., & Leifer, L. (2012). 'Finding One's Way': Re-Discovering a Hunter-Gatherer Model based on Wayfaring. International Journal of Engineering Education, 28, 251–252.

Stunt, J., Wulms, P., Kerkhoffs, G., Dankelman, J., van Dijk, C., & Tuijthof, G. (2014). How valid are commercially available medical simulators? Advances in Medical Education and Practice, 5, 385–395. https://doi.org/10.2147/AMEP.S63435

Sutcliffe, A., & Sawyer, P. (2013). Requirements elicitation: Towards the unknown unknowns. 2013 21st IEEE International Requirements Engineering Conference (RE), 92–104. https://doi.org/10.1109/RE.2013.6636709

Takeuchi, H., & Nonaka, I. (1986). The new new product development game. Harvard Business Review, 64(1), 137–146.

Thomke, S., & Bell, D. E. (2001). Sequential Testing in Product Development. Management Science, 47(2), 308–323. https://doi.org/10.1287/mnsc.47.2.308.9838

Thomke, S., & Fujimoto, T. (2000). The effect of "front-loading" problem-solving on product development performance. Journal of Product Innovation Management, 17(2), 128–142. https://doi.org/10.1016/S0737-6782(99)00031-4

Thomke, S., & Reinertsen, D. (1998). Agile Product Development: MANAGING DEVELOPMENT FLEXIBILITY IN UNCERTAIN ENVIRONMENTS. California Management Review, 41(1), 8–30. https://doi.org/10.2307/41165973

Tolo, J. S. (2022). Simulator face demo [Photograph].

Tosey, P., Visser, M., & Saunders, M. N. (2012). The origins and conceptualizations of 'triple-loop' learning: A critical review. Management Learning, 43(3), 291–307. https://doi.org/10.1177/1350507611426239

Ulrich, K. T., & Eppinger, S. D. (2011). Product Design and Development. McGraw-Hill Education. https://lib.hpu.edu.vn/handle/123456789/32894

Verworn, B., Herstatt, C., & Nagahira, A. (2008). The fuzzy front end of Japanese new product development projects: Impact on success and differences between incremental and radical projects. R&D Management, 38(1), 1–19. https://doi.org/10.1111/j.1467-9310.2007.00492.x

Vestad, H., Kriesi, C., Slåttsveen, K., & Steinert, M. (2019). Observations on the Effects of Skill Transfer through Experience Sharing and In-Person Communication. Proceedings of the Design Society: International Conference on Engineering Design, 1(1), 199–208. https://doi.org/10.1017/dsi.2019.23

Vestad, H., & Steinert, M. (2019). Creating your Own Tools: Prototyping Environments for Prototype Testing. Procedia CIRP, 84, 707–712. https://doi.org/10.1016/j.procir.2019.04.225

Wall, M. B., Ulrich, K. T., & Flowers, W. C. (1992). Evaluating prototyping technologies for product design. Research in Engineering Design, 3(3), 163–177. https://doi.org/10.1007/BF01580518

Wynn, D., & Clarkson, P. (2018). Process models in design and development. Research in Engineering Design, 29, 161–202. https://doi.org/10.1007/s00163-017-0262-7

Yin, R. K. (2012). Case study methods. In APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological (pp. 141–155). American Psychological Association. https://doi.org/10.1037/13620-009

C1

Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators

Procedia CIRP 84, 572-578: CIRP Design 2019 Conference

Marius Auflem, Jørgen Falck Erichsen, Martin Steinert (2019),

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 84 (2019) 572–578

www.elsevier.com/locate/procedia

29th CIRP Design 2019 (CIRP Design 2019)

Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators

Marius Auflem[a]*, Jorgen Falck Erichsen[a], Martin Steinert[a]

[a] NTNU Trondheim, Richard Birkelands veg 2B, Trondheim 7491, Norway

* Corresponding author. Tel.: +47 909 17 081, *E-mail address:* mariuauf@stud.ntnu.no

Abstract

This paper attempts to exemplify prototype-driven development in the early stages of product development, the stages before requirements and specifications are fixed. This pre-requirement phase provides opportunities and uncertainties for the design team to explore, and this paper shows how this could be (and has been) done through extensive use of explorative prototyping. Prototyping, in this context, is the activity building and experimenting with various concepts with the aim of producing tangible insights as fast as possible. In prototyping, prototypes are tangible artifacts built to answer specific questions, in order to explore and gain new insights as the project requirements emerge. The context for this article is product development of patient simulators used in medical training, referred to as 'Mannequins'. Mannequins are widely used in medical training to enable practice of treatment for conditions too rare or dangerous to perform on real patients. From this context, specific examples on prototype-driven development are shown through two case projects; Development of a chest for the training of cardiopulmonary resuscitation, and a fractured leg in order to train on realigning and stabilizing displaced fractures. These projects are user-centered design challenges within the medical education field. This paper shows how a prototype-driven development approach could be utilized on a project level and provides insights on prototyping to gain answers, learning and inform decisions. The paper argues that before requirements and specifications are fixed, a more exploratory and prototype-driven approach is needed, in order to provide more informed requirements and specifications. This way, prototypes are the drivers of the development and the iterations impact the direction of the ongoing development. Specific aspects of prototype-driven development such as user-interaction, prototype resolution, evaluation and testing are also discussed in this paper.

Peer-review under responsibility of the scientific committee of the CIRP Design Conference 2019.

Keywords: Prototyping; Prototype-Driven Development; User-Centered Design; Emerging Requirements

1. Introduction

When exploring new opportunities within a product domain, the ambiguity and lack of constraints can lead development teams into doing premature decisions in projects. This could result in costly rework and products failing due to not meeting the targeted users' requirements or needs [1]. In this pre-requirement phase of product development, the uncertainty and opportunities facing the design teams are important to explore in order to do informed decisions. Upcoming challenges and opportunities remain hidden unless elicited or made explicit in the ongoing development [2,3]. Hence, how to leverage unknown opportunities and accommodate future challenges is not evident—yet important—in product development [4].

By presenting two case projects we exemplify how prototyping have been utilized to explore and gain answers before requirements and specifications are made tangible or fixed. The cases are gathered from two early stage development projects focusing on development of medical training simulators further referred to as *mannequins*. In these projects, the design teams set out with no fixed or predetermined product requirements, and the goal was to investigate needs and corresponding opportunities for mannequins to improve or introduce new functionality for medical training and simulation.

2212-8271 © 2019 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the CIRP Design Conference 2019.
10.1016/j.procir.2019.04.202

1.1. Research Question

Prototypes serve various purposes in product development and the importance of prototypes is frequently highlighted in research [5,6]. Schrage [7] propose that in order to create better products, organizational cultures must learn to create better prototypes. Further, it is discussed how companies should derive their product requirements from prototypes as a contrast to requirement driven prototyping [7]. While these statements are based on interviews with industry actors, there is a call for empirical data to support the statements. This paper will contribute to how prototypes could be utilized to explore and establish product requirements on a project level. By presenting examples and findings on the use of prototypes from two case projects, we will answer the following research question; *How can prototypes be used to explore and establish informed requirements as opposed to using prototypes for meeting set requirements?*

1.2. Research scope

This research article provides examples and insights on the use of prototype-driven development on a project level. Hence how prototyping could be utilized for exploring and gaining answers in product development projects. Prototypes are an important aspect of research on early stage product development, product development methodology and managerial frameworks in this context [1,8]. This paper will, however, focus on how prototypes could be developed and utilized to provide designers and developers with examples for tackling the uncertainty of projects before product requirements and specifications have been made tangible.

1.3. Prototyping and Roles of Prototypes

The use of prototypes in different settings, disciplines, and stages of development has resulted in several frameworks for defining prototypes and their purposes [9]. While some see prototypes as product approximations or tools for testing and verifying early designs, the generative role of prototypes and prototyping activities is of interest when exploring potentials in the early phases of product development. From case studies, [10] have derived three roles of prototypes within companies, where they present how prototypes serve as tools for communicating, learning and for informing decision making.

As roles of prototypes and how prototypes are utilized in projects are described, prototyping is often explained as the creation and utilization of such artifacts [11]. The authors argue that the importance of prototyping ranges further than just the activity of creating prototypes. In this paper we define prototyping as the designing, building and testing of new concepts and ideas. Hence prototyping is considered a learning activity, cognitive and physical, and can enable new insights and generate knowledge when exploring a solution space [12]. The outcome of prototyping is therefore generated knowledge and prototypes, tangible artifacts embodying this either explicit or tacit knowledge [13].

1.4. Answering Design Questions

As prototyping is a tool for acquiring new insights, prototypes are built and tested to answer questions [5]. Hence, the prototyping medium is determined by the questions that need answering and both, physical, digital and analytical models can serve the purpose as prototypes [7]. The importance of prototypes is not how they are created or their closeness to a final product, but rather how they are utilized to gain answers to important open design questions [14].

In the context of this paper—i. e. products designed for interaction with users—many design questions require external feedback to be answered. An example is prototyping to answer how a product would serve a role in a user's life or how the interaction is perceived by the look and feel of an artifact [14]. Prototypes are a mode of communication and they enable interactions and design teams to explain concepts in a tangible matter and gain feedback [10]. As boundary objects, prototypes can be used to establish a common ground for this communication to happen by bridging both disciplinary and knowledge gaps.

1.5. Prototyping Strategies

In product development, the generative role of prototyping is effective when trying to come up with novel ideas and multiple alternatives for exploring a solution space. This concept generation is a divergent approach seeking out the potential solutions before converging down on one or multiple concepts to develop further. Eris [15] propose that divergent and convergent thinking could be achieved by subsequently asking generative design questions and deep reasoning questions in development projects. Generative design questions are open-ended, seeking to identify multiple possibilities not tied to the logical nature of the problem, while deep reasoning questions could measure the applicability of revealed alternatives and sort out unfeasible solutions or concepts [15].

In the early (i.e. pre-requirement) phase of product development, designers could benefit from using low-resolution prototypes to gain rapid answers and insights. We consider the resolution of prototypes as the level of detail. Note that this is often differentiated from fidelity, as the latter is considered the closeness to the eventual (final) design [14]. Utilizing low-resolution prototypes their rough construction and unfinished attributes allows playing with the ideas, possibilities, and potentials rather than verifying design [7]. Also, using a lower resolution makes it easier to get inspiration and change or generate concepts from the gained insights, all which could prohibit designers from prematurely fixating on design solutions [16].

When investigating the potentials of ideas and proposed concepts, a higher resolution might be necessary in order to gain unbiased or unclouded feedback, as many questions require external answering in the design process. Designers must be aware and reflective what prototypes they present, and to what audience, as prototype attributes and intent not necessarily is communicated by the artifact itself [14].

2. Case Projects

The development projects used as cases for this article were requested by a medical company and performed by two teams of graduate students. The first project is the development of a mannequin chest for training of cardiopulmonary resuscitation (CPR) and the second; the development of a leg for training of displaced bone fracture realignment. Mannequins are widely used in both skill training and education of health care providers. The aims of these projects were to create safe and repetitive training environments, that would appear realistic enough to enable users to transfer skill and knowledge into real-world medical scenarios.

2.1. Case 1: Resuscitation Mannequins

Resuscitation mannequins are no recent invention and commercially available products for training medical personnel and laypeople in CPR have existed for decades. The mannequins are most often human-like dummies, as seen in Fig. 1, that allow for chest compressions and artificial ventilation, as one would perform on a person suffering from sudden cardiac arrest. The project was proposed as; to rethink and develop a new chest concept for resuscitation mannequins to closer resemble the human chest and enable a more realistic chest compression experience for users in training. This was considered a response to the lack of realism found in currently used mannequins [17]. This project was carried out over a period of 9 months. During this period, a total of 84 prototypes was developed for a new mannequin chest concept.

Fig. 1. Example of one commercially available resuscitation mannequin. This uses a linear compression spring mechanism to enable chest compressions.

2.2. Case 2: Displaced Leg Fracture Task Trainer

Advances in emergency care training and patient simulators, various tasks are now being taught using human-like mannequins. The second project was requested to explore the need for a mannequin-based trainer for realignment of a displaced leg fracture and subsequently the requirements for this functionality. Displaced fractures are common as well as challenging to treat for emergency responders, as these fractures could cause circulation issues and potential damages to tissue and vessels. The procedure of realignment and stabilization of fractures are taught both in theory and by using human markers. Human markers (i.e. actors) are used for training in securing and stabilizing the leg by fixing it using splints but does not enable training of the actual repositioning.

Mannequins are products designed to prepare users for procedures and interactions too dangerous or rare to be trained on real patients or human markers. Hence lack of realism, by their ability to include functionality as found in the human body, could leave users insufficiently prepared for interactions with patients. Therefore, in the design of mannequins, it is a desire to approximate the physiological aspects required to perform a given task, but at the same time avoid introducing aspects not found in human patients. Such aspects could interfere with the simulation, sense of immersion, and potentially introduce sources of false learning.

This development project of a new leg for mannequins was carried out over 4 months and resulted in more than 15 conceptual prototypes.

The following subsections show how prototyping has been extensively used to drive the development of the two projects and to identify and explore revealed product opportunities.

2.3. Exploring Opportunities for Case 1

In Case 1, the starting point for the project was to rethink and create a new chest concept for resuscitation mannequins. A chest would have to have the ability to be compressed and recoil as a human chest would do, to enable users to practice routine and motor skills for CPR. Already existing solutions for CPR training varies by concept, but there is a consensus about their lack of realism and simplified characteristics as compared to a human chest. This being the background for the project, the developers aimed to create a concept with functionalities closer resembling the human body, leaving users better prepared for an eventual real encounter of a cardiac arrest patient in need of chest compressions.

Initial steps of the development consisted of simultaneous explorative prototyping and research in order to create rough prototypes of aspects of the human chest to investigate. Identified characteristics where split into two areas of interest; 1: Whether patients ribs fracture during CPR and how this affects the rescuer? 2: How a chest deforms when compressed and how it feels to perform compressions? Generative low-resolution prototyping resulted in three conceptual prototypes attempting to answer the two questions above.

The first prototype, shown in Fig. 2, attempted to simulate ribs fracturing from excessive loading, while the two

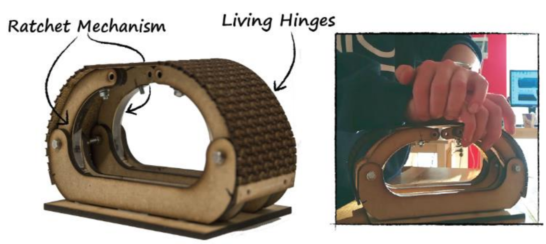


Fig. 2. Rib fracturing model with mechanical features to the left and testing of the prototype shown on the right.

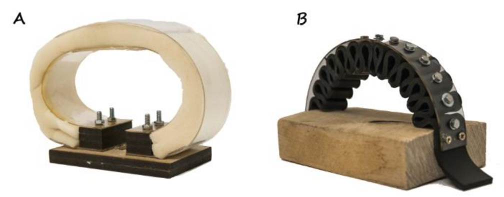


Fig. 3. Spring configurations using (A) foam and (B) rubber for increased resistance and stability.

prototypes, in Fig. 3, were using different spring configurations to simulate the tactility and deformation of a chest. While the questions concerned real-world interactions with patients, the team wanted to expose the prototypes to "users" with prior clinical CPR experience and allow them to test and discuss the characteristics and functionalities of the prototypes.

Experience as in inherited knowledge by the users is, however, not always explicit and articulated. More so, users from the field of medicine possess knowledge from their education, training, and work experience, making the disciplinary knowledge gap between medical personnel and design engineers vast.

Prototyping showed potential in bridging this gap, as the users interacting and testing the prototypes could articulate their experiences by comparing them to the physical characteristics of the artifact. More importantly, this experience and tacit features were made tangible to the development team through the prototypes. Jargon and complex sensory experiences were translated into a physical/technical context that was able to influence future development.

The testing and interaction resulted in new insights and unknown aspects of patient CPR identified as opportunities for the team to investigate. The insights were made explicit as the following points:

- The patient ribs fracture almost every time, and that this is easily sensed. It could be compared to breaking thin branches under a thick carpet as opposed to the brittle clicks provided by the presented prototype.
- Chest compressions are not like compressing on the spring-like prototypes, but more like a hard couch pillow. It becomes harder by the depth of the compression and is considered less responsive than a spring.
- The stiffness of a chest is not constant, as it would reduce in stiffness and responsiveness after many compression cycles.

2.4. Exploring Opportunities for Case 2

Like the previous example, the team in Case 2 (developing a mannequin leg for repositioning training) developed low-resolution prototypes to investigate the context of leg fracture and repositioning. Here, the procedure and interactions when first responders come to aid a patient suffering from a displaced fracture. In this project it was observed how the team used prototyping and physical interaction with prototypes to understand and make their problem tangible. This is exemplified by the prototype, as seen in Fig. 4, that was made to accommodate their initial findings from research, that repositioning is important to relieve pain and ensure circulation to the distal part of the fractured leg. Open design questions were at this point how repositioning a leg is experienced from a rescuer's perspective and what tactile experience and challenges it might impose. In order to explore this interaction, the prototype was strapped to one of the team members legs, as seen in Fig. 5, and was then attempted repositioned by paramedics at the hospital.

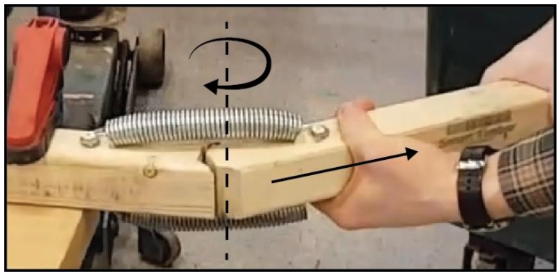


Fig. 4. Broken leg model suspended by springs with arrows indicating the pull and rotate movement.

During realignment, the paramedics pointed out how the procedure is usually very painful, and that the patient must be given sedatives for them to perform it. Swelling and muscle tensioning around the fracture would also constrain the movement, and both sedatives and physical fatigue of the muscles is often necessary to realign the fracture. The paramedics reenacted the procedure and showed how repositioning requires the rescuers stretch the patient's leg by leaning back. Using his or her own body weight, as well as another person holding the patient, could be necessary in order to gradually elongate the muscles and reposition the fracture.

Based on this feedback, simulating tiring and sedated muscles became a new feature to investigate. This had not been identified earlier by the team but was made apparent by users testing and interacting with the rough prototype.

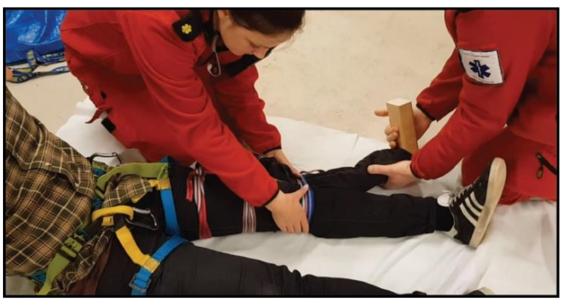


Fig. 5. Paramedics attempting to reposition the broken leg model strapped to one of the team members.

2.5. Generating and Evaluating Concepts for Case 1

From investigating the mannequin chest development, it became evident that the development team used prototyping to generate concepts that could adapt the feedback and insights revealed from the earlier testing and interactions with users. As prototypes were created, they were tested and iterated upon to reveal a potential for answering the identified opportunities. The team prototyped extensively within two domains, namely the chest deformation and characteristics, and rib fractures by haptic and audible response.

The prototyping outcome, in form of prototypes, is illustrated in Fig. 6. In the figure, it is noticeable how different concepts were first evaluated on a rough principle level before being either discarded or further developed through concept iterations. As the team developed prototypes along two distinct paths of interest, each concept had the opportunity to be tested and compared to alternative solutions along that path. Having multiple prototypes to compare, decisions could be made based on relative performance measures.

One example of this prototype evaluation is found along path A in Fig. 6. Concepts A3, A4, A5, and A6, were tested and compared, revealing strengths and weaknesses of the different concepts. As prototype, and concept, potentials were made apparent, the team got empowered to select which concepts to develop further by new prototype iterations. Concepts deemed promising based on the prototype's performance was developed further to investigate the potential and for meeting the targeted form and force characteristics for an adult chest.

In this project it was observed how this iterative and selective approach, discarded unfeasible solutions before landing on one concept for each domain. Here, one was simulating the shape and deformation of the chest when compressed (A6.3), and one was simulating the tactile feeling of ribs being fractured from excessive loading (B9.2). As these prototypes had undergone several rounds of changes and testing, and the team deemed these as good approximations of the functionalities elicited from the medical personnel. As functional prototypes, they were tested by medical personnel to enable feedback and evaluation of the proposed concept and the included functionality. Hence, these could provide answers to if, and how, a product could be realized and the corresponding requirements for the future product.

2.6. Generating and Evaluating Concepts for Case 2

The team investigating repositioning of displaced leg fractures had identified how muscles constraining the fracture played a crucial role in creating a realistic simulator. Hence, investigating the solution space for mimicking the biomechanics of a contracted muscle became a core objective.

During the development of the broken leg simulator, generative design questions enabled widening the solution space and testing multiple alternative concepts through prototyping. Asking "how many ways they could create a linear actuation mechanism constraining a fracture" resulted in the generation of low-resolution prototypes to be tested. The prototypes investigated different physical principles for constraining a simulated leg, and how these principles could be actuated and controlled in order to simulate the elongation of muscles.

Electromagnets, mechanical springs, hydraulics, pneumatics, air-muscles, and muscle-wire were investigated and tested resulting in multiple promising concept proposals. From internal testing, the team noted strengths and weaknesses of their concepts before deciding on which to develop further. The team identified that ease of control for many of their prototypes, compromised the tactile feeling of a muscle as

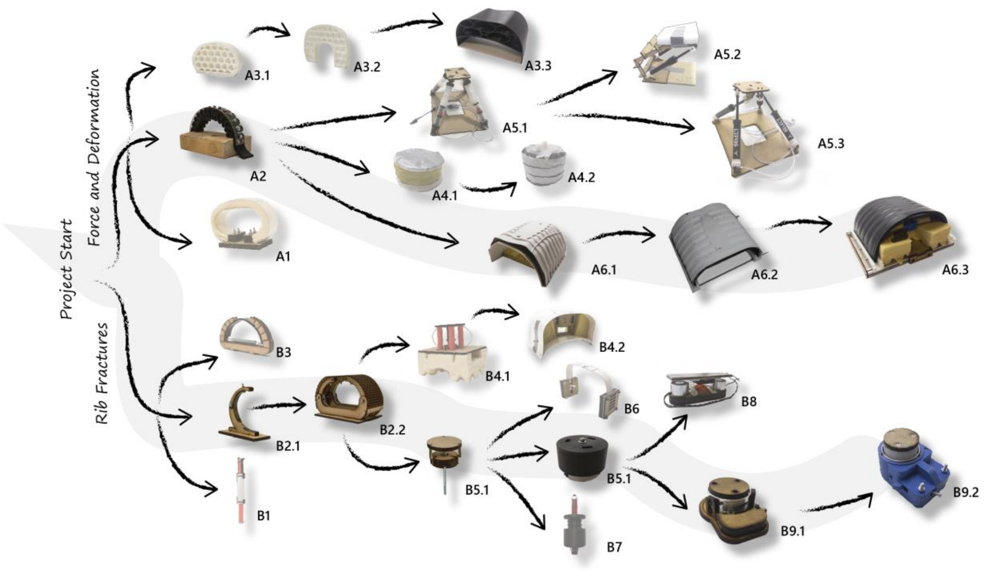


Fig. 6. Retrospective mapping of the most influential prototypes developed throughout the timeline of the mannequin chest project. Path A investigating concepts for chest deformation and tactility and path B concepts for simulating ribs fracturing from compressions.

described by the paramedics. By evaluating the alternative concepts by prototypes, the team decided on moving forward using a pneumatic system. Pneumatic cylinders were evaluated as a robust and controllable principle, which also provided an “organic tactile experience” as the air being compressed in the system allowed for subtle movements.

Investigating how pneumatics could be integrated in a mannequin leg, the team developed a proof-of-concept prototype to experiment with different pressures and connections constraining the leg as seen in Fig. 7. With this prototype, the team tried to answer questions concerning integration of the earlier revealed functionality.

To gain answers to the usability, tactile experience and training procedure, the team further developed the leg model by hiding the mechanisms and replicating a rough look and feel of a human leg, as seen in Fig. 8. This prototype was tested with paramedics to gain feedback on how the proposed concept could aid users in training, and if the captured functionality was accurate.

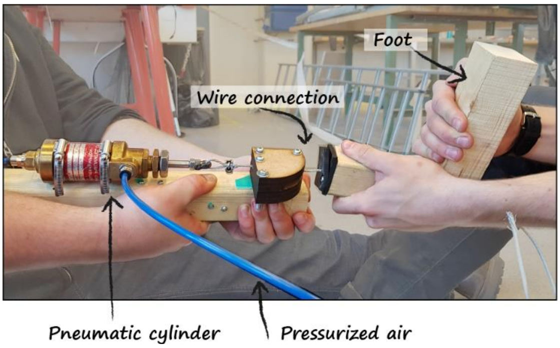


Fig. 7. Proposed concept prototype of a broken leg for mannequins.

Fig. 8. Testing of proposed conceptual prototype with paramedics.

2.7. Selected Concepts and Emerging Requirements

Prototyping was utilized to translate the vision and ideas of the design teams back to users and the physical world and context of medical simulation. By proposing a concept prototype, the teams could gain important answers to if their earlier findings were substantial and accurate for the context of a new product. Hence both development teams utilized higher resolution prototypes to manifest their insights as requirements for future products.

In Case 1, this process consisted of both internal testing, measuring the characteristics of the proposed prototype, and external testing with medical personnel at the hospital. Internal testing and measurements were carried out to quantify prototype characteristics and compare this to the feedback as well as physiology data found in research [18]. These efforts in testing and evaluating the proposed concepts were performed to settle the emerging requirements and manifest the opportunities as features to include in a product. The learning from this process provided suggestions to incremental design changes, as well as affirming the elicited functionality.

In Case 2, the team integrated their proposed concept with an existing simulator enabling paramedics to attempt repositioning on a full-scale mannequin, as shown in Fig. 8. This enabled a realistic scenario for them to reenact the procedure and give feedback to the functionality and tactile experience of performing the procedure. In this process, the emerging requirements from prior testing and concept generation was made apparent and confirmed. For example, the slight movement and play of the pneumatic cylinder was considered a good approximation of the tactility of the tense muscles constraining the fractured leg.

The results from this testing, confirmed the elicited functionalities in both projects. Additionally, it provided new insights for the teams to bring forward in the continuation of the projects. Based on how the presented prototypes performed and their evaluation from medical personnel, the teams could establish and communicate requirements for the future products to be realized.

3. Discussion

In the two presented cases, prototypes enabled a discussion with expert users on needed functionality and aspects important keep on the radar for the development teams. It is, however, worth questioning if similar insights would have been accessible by investing enough resources on upfront research. This would have required looking into, e.g. analytical simulations of the human body, research on biomechanical behavior of human physiology and in-depth interviews with stakeholders. While using a systematic method of establishing upfront requirements could have led to meaningful specifications and functionalities to include, using prototypes quickly made these insights, not only available but also tangible. Prototyping enabled eliciting sensory experiences from trained medical personnel and provided a common understanding of how this was either represented or lacking in the presented prototypes. As the identified functionalities were described and reenacted by using the prototypes, it is not evident that this tacit knowledge could have been accessed through interviews and research alone.

The prototyping carried out by the two teams lead to the generation of multiple concepts and prototypes to be tested and evaluated in parallel. This was made possible by fast low-resolution prototyping in both projects. The identified functionalities and tacit features were attempted realized as multiple conceptual prototypes providing the teams with critical answers informing the development.

Concept generation through generative design questions was proven useful in covering a wider area of the solution space. Hence, having a better chance of finding a suitable concept for the specific design challenge. Further, the generation and testing of multiple concepts and ideas by prototyping avoided prematurely fixating on solutions. This is especially important when approximating aspects of the human body, as designing by the inspiration of physiology and copying human attributes could become a fixating element.

By being able to test often and adapt concepts as requirements emerged and shifted, the development teams could do informed decisions and quickly launch "proof-of-concept" prototypes to gain feedback. The identified functionalities for the two new products could hereby be tested and evaluated before being deemed ready for further development. This is a clear benefit of extensive prototyping as gaining answers fast and aligning development to fit users' needs and specifications is vital for eventually launching a successful product.

The examples from the presented cases have shown the importance of prototyping when moving into and exploring a new product context. However, it is worth noting the limitations of only relying on prototypes and prototype driven methods. Prototyping is but one tool in the toolbox of design engineers and is complementary rather than opposing to other working modes in the early stage of product development. As requirements and product plans are being solidified, new questions arise for product developers to address. Hence, this would require different prototyping strategies, as well as the utilization of diverse engineering tools to gain answers.

We propose this extensive use of prototyping as one way of accommodating the uncertainty of the pre-requirement phase of projects and using prototyping for learning to elicit and explore emerging requirements for new products.

4. Conclusion

The main contribution of this paper, and answer to the research question ("*How can prototypes be used to explore and establish informed requirements as opposed to using prototypes for meeting set requirements?*") has been to give two concrete case examples of how to drive development and establish informed requirements using prototyping.

By studying two case examples on prototype-driven development, it has been identified how prototyping activities for learning are important for eliciting and exploring functionalities and corresponding requirements for new products. In this context, prototyping has been observed to enable design teams to explore product potentials, communicate with users, and doing informed decisions by generation and evaluation of concepts. This paper has shown how prototype-driven development could be done to accommodate the uncertainty before requirements are made fixed or tangible. By this, prototyping is proposed as a complementary tool to be utilized for exploring and establishing informed requirements in the pre-requirement phase of product development projects.

Acknowledgements

This research is supported by the Research Council of Norway through its user-driven research (BIA) funding scheme, project numbers 236739/O30 and 290404.

References

[1] Thomke S, Reinertsen D. Agile Product Development: Managing development flexibility in uncertain environments. California Management Review 1998;41:8–30. doi:10.2307/41165973.

[2] Kriesi C, Blindheim J, Bjelland Ø, Steinert M. Creating Dynamic Requirements through Iteratively Prototyping Critical Functionalities. Procedia CIRP 2016;50:790–5. doi:10.1016/j.procir.2016.04.122.

[3] Jensen MB, Elverum CW, Steinert M. Eliciting unknown unknowns with prototypes: Introducing prototrials and prototrial-driven cultures. Design Studies 2017;49:1–31. doi:10.1016/j.destud.2016.12.002.

[4] Sutcliffe A, Sawyer P. Requirements elicitation: Towards the unknown unknowns. 2013 21st IEEE International Requirements Engineering Conference (RE), 2013, p. 92–104. doi:10.1109/RE.2013.6636709.

[5] Wall MB, Ulrich KT, Flowers WC. Evaluating prototyping technologies for product design. Research in Engineering Design 1992;3:163–77. doi:10.1007/BF01580518.

[6] Camburn BA, Dunlap BU, Kuhr R, Viswanathan VK, Linsey JS, Jensen DD, et al. Methods for Prototyping Strategies in Conceptual Phases of Design: Framework and Experimental Assessment. Volume 5: 25th International Conference on Design Theory and Methodology; ASME 2013 Power Transmission and Gearing Conference, Portland, Oregon, USA: ASME; 2013, p. V005T06A033. doi:10.1115/DETC2013-13072.

[7] Schrage M. The Culture(s) of prototyping. Design Management Journal (Former Series) 1993;4:55–65. doi:10.1111/j.1948-7169.1993.tb00128.x.

[8] Gerstenberg A, Sjöman H, Reime T, Abrahamsson P, Steinert M. A Simultaneous, Multidisciplinary Development and Design Journey – Reflections on Prototyping, Springer, Cham; 2015, p. 409–16. doi:10.1007/978-3-319-24589-8_33.

[9] Jensen LS, Özkil AG, Mortensen NH. Prototypes in engineering design: Definitions and strategies. DS 84: Proceedings of the DESIGN 2016 14th International Design Conference, 2016.

[10] Lauff C, Kotys-Schwartz D, E. Rentschler M. What is a prototype? what are the roles of prototypes in companies? Journal of Mechanical Design 2018;140. doi:10.1115/1.4039340.

[11] Lim Y-K, Stolterman E, Tenenberg J. The anatomy of prototypes: prototypes as filters, prototypes as manifestations of design ideas. ACM Trans Comput-Hum Interact 2008;15:7:1–7:27. doi:10.1145/1375761.1375762.

[12] Leifer LJ, Steinert M. Dancing with ambiguity: Causality behavior, design thinking, and triple-loop-learning. Information Knowledge Systems Management 2011;10:151–73. doi:10.3233/IKS-2012-0191.

[13] Erichsen JAB, Pedersen AL, Steinert M, Welo T. Learning in product development: proposed industry experiment using reflective prototyping. Procedia CIRP 2016;50:454–9. doi:10.1016/j.procir.2016.04.142.

[14] Houde S, Hill C. What do prototypes prototype? Handbook of Human-Computer Interaction, Elsevier; 1997, p. 367–81. doi:10.1016/B978-044481862-1.50082-0.

[15] Eris Ö. Asking generative design questions: a fundamental cognitive mechanism in design thinking. DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design, Stockholm, 2003, p. 587-588 (exec. summ.), full paper no. DS31_1640FP.

[16] Viswanathan V, Linsey J. Design fixation in physical modeling: An investigation on the role of sunk cost 2011:119–30. doi:10.1115/DETC2011-47862.

[17] Baubin MA, Gilly H, Posch A, Schinnerl A, Kroesen GA. Compression characteristics of CPR manikins. Resuscitation 1995;30:117–26. doi:10.1016/0300-9572(95)00874-S.

[18] Tomlinson AE, Nysaether J, Kramer-Johansen J, Steen PA, Dorph E. Compression force–depth relationship during out-of-hospital cardiopulmonary resuscitation. Resuscitation 2007;72:364–70. doi:10.1016/j.resuscitation.2006.07.017.

C2

On Prototyping Methods to Leverage Non-Rigid Materials

in the Early Stages of Engineering Design

Proceedings of the Design Society: DESIGN2020 Conference

Marius Auflem, Hans Hagenes Bøe, Jørgen Falck Erichsen, Martin Steinert (2020),

INTERNATIONAL DESIGN CONFERENCE – DESIGN 2020
https://doi.org/10.1017/dsd.2020.120

ON PROTOTYPING METHODS TO LEVERAGE NON-RIGID MATERIALS IN THE EARLY STAGES OF ENGINEERING DESIGN

M. Auflem [1,✉], H. H. Boe [2], J. F. Erichsen [1] and M. Steinert [1]

[1] Norwegian University of Science and Technology, Norway, [2] CERN Idea Square, Switzerland
✉ mariuauf@stud.ntnu.no

Abstract

Prototyping has been shown important to facilitate learning, inform decisions and to communicate ideas in engineering design. However, it is not evident which methods, tools and materials to use, as prototyping is practised differently across development contexts, and stages. In the early stages of design, different choices in prototyping methods, tools and materials all affect prototyping outcome. This paper is focused on prototyping methods in the context of early stages of design and attempts to highlight identified strengths and limitations of using non-rigid materials for prototyping.

Keywords: prototyping, design tools, early design phase, case study

1. Introduction

In engineering design research, prototyping has been shown important to facilitate learning, inform decisions and to communicate ideas (Lauff et al., 2018). However, it is not evident which methods, tools and materials to use, as prototyping is practised differently across contexts, disciplines and stages of development. The design engineer has to prioritise various trade-offs in order to effectively create focused prototypes with purpose and intention (Bryan-Kinns and Hamilton, 2002; Houde and Hill, 1997; Lim et al., 2008; Menold, 2017).

Although there are many definitions of prototypes in engineering design research, as discussed by Jensen et al. (2016). this paper uses the definition from Schrage (1993), who states that prototypes are created to answer questions. Consequently, all objects could serve as prototypes if they convey some attribute of the conceptual idea to the designer (Buchenau and Suri, 2000). Subsequently, prototyping is defined as a learning activity (Leifer and Steinert, 2011), both cognitive and physical, where the goal is to generate tangible insights in form of prototypes (Auflem et al., 2019).

In the early stages of engineering design, different choices in prototyping methods, tools and materials all affect prototyping outcome. This paper is focused on the use of materials in the context of early stages of engineering design and attempts to highlight identified strengths and limitations of prototyping methods using non-rigid materials. Flexible and semi-flexible materials (further referred to as non-rigid materials), have enabled product opportunities within various fields and disciplines. However, the authors argue that in product development, and especially in the early stages of design, there is a lack of research, tools and, examples on how to leverage these non-rigid materials effectively.

In this paper non-rigid materials will for all intents and purposes be considered materials allowing elastic deformations even when subjected to large deflections. While non-rigid materials are not new to the field of engineering, availability of materials and manufacturing technologies have made them a relevant topic in research and recent developments. Medical technology, soft robotics, biomimicry and compliant mechanisms are some areas where non-rigid materials have had a substantial impact and introduced new functional opportunities.

In engineering design, complex material properties, deformations and/or geometries could all make design and dimensional tasks challenging. Physical computation, as described by Foehr et al. (2015), is shown to be one way of addressing such complex material characteristics and concepts. Here, non-linearity as a result of large deformations caused numerical computations to diverge- and crash- by having too many unknown parameters. A rapid prototype could in this case be used to reduce the complexity of such computational task by solving, or informing, one or several of the unknown parameters by physical experimentation. Kriesi et al. (2016) presents a similar approach in exploring the complex nature of injection molding processes by prototyping. By rapidly creating molds for a simple desktop injection moulding machine, they could iteratively test and learn about functional requirements for mould-designs, without the need of tedious and complex simulations. Other examples involve physical modelling of manufacturing techniques using plasticine as a proxy material for metal. Common for these cases is how prototyping complex materials and/or conceptual models could enable fast learning without tedious simulations or numerical models.

As prototypes are created and evaluated trough testing, working in iterative cycles could enable concepts to adapt as insights get revealed to the designers. This could reduce the risk of costly rework later in the development. In this process, the creation of prototypes is a crucial activity that facilitates generation of new ideas (ideation), divergence, as well as making these ideas tangible and testable in order to inform further design iterations (Dow et al., 2011). This generative role of prototyping activities is important for designers to diverge, exploring multiple design alternatives and avoid prematurely fixating on concepts or designs (Viswanathan and Linsey, 2011). Hence, rapid and iterative prototyping is considered a crucial tool when exploring uncharted solution spaces in the early stages of engineering design (Camburn et al., 2013).

From these principles, this paper will present similar approaches to physical modelling, and evaluate methods of iterative prototyping to generate fast tangible insights. This will be done to address the challenges of doing concept generation, exploring solution space, and more specifically to leverage non-rigid materials in the early stages of engineering design. In this context the paper will focus on specific design activities, namely prototyping methods to explore and leverage non-rigid materials, as an alternative to evaluating prototyping outcome, purpose and the attributes of prototypes (Jensen et al., 2015). With prototyping being defined as cycles of designing, building and testing of prototypes, the method used to create prototypes is of interest to better understand the activities in design projects.

Further, how the various methods utilized, affect design outcome (prototype attributes), and performance, (build time, facilitating design iterations, and design freedom). By researching such prototyping methods, the authors seek to aid designers in doing strategic decisions when selecting approach to create non-rigid prototypes in the early design stages.

The methods considered in this paper will be illustrated through individual cases, gathered from different development projects. Hence, it is of interest how the different methods affect both prototyping activities and outcome. The presented cases are all projects in early design stages focused towards concept generation and evaluation using prototypes. By this, the paper aims to exemplify various prototyping methods, utilizing tools and technologies to leverage non-rigid materials. These are methods that could help generate fast tangible insights through the designing, building, and testing of physical non-rigid prototypes.

2. Prototyping methods in early design

In this section, four prototyping methods for generating non-rigid concepts are illustrated through four project cases. Common for these projects is the development context of early design and potentials for exploring non-rigid concept proposals. As prototypes are created with varying intent and utilized differently across projects, the success criteria for each prototype vary subsequently. Hence, this paper attempts to evaluate the prototyping methods used, rather than the physical artefacts created. By this,

https://doi.org/10.1017/dsd.2020.120 Published online by Cambridge University Press

revealed insights could support designers in effectively deploy methods to generate prototypes with non-rigid abilities. To discuss the presented prototyping methods, the following five parameters are used to describe how the prototyping activity and outcome in terms of prototypes are affected by the method utilized.

- Speed
- Modifications
- Design freedom
- Fidelity
- Resolution

The two initial parameters, speed and modifications, are considering how well a given method enables rapid and iterative working modes in early design. This is important in obtaining fast insights, trough design-build-test cycles when exploring new solution spaces (Leifer and Steinert, 2011). Speed is here used to describe the time per prototype iteration. In the context of tangible insights, time per prototype iteration is of interest because it can help describe if and how a method enables fast learning. Modifications discusses how a method facilitates continuous alterations or if a prototype is hard to change or adapt to revealed insights.

Design freedom evaluates the geometrical, material, and size constraints of any given method. Hence, how workable a method is, and if it can easily adapt to new design ideas, allow for design alterations, or for complex design ideas to be realized. Fidelity discusses the obtainable degree of closeness to eventual design with a given method while resolution discusses the obtainable amount of detail with a given method (Houde and Hill, 1997).

2.1. Cases

The following sections describe four different methods of fabricating non-rigid prototypes trough individual case examples. These cases ranges from low-resolution concept generation for flexible animatronic robots, to creating a proof-of-concept prototype to be tested with ultrasound imaging equipment. The methods considered in these cases have only utilized one, or a limited number, of materials and machining/fabrication possibilities. While all the cases concern different materials and fabrication methods, they are not intended as a representative selection for other methods. However, the presented methods are four distinct different alternatives, that visualize a span of tools and technologies that could help leveraging non-rigid materials in prototyping.

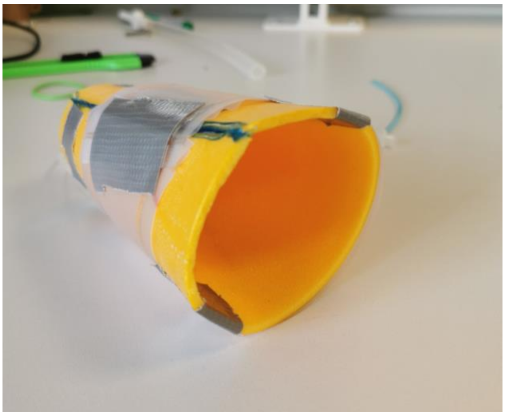
2.2. Soft prototyping

Soft Prototyping is in engineering design known to ease both creative efforts, team engagement and in avoiding sunk cost effects (Dow et al., 2009). Hence, it is often the initial steps of exploratory prototyping activities in new projects. This method enables the generation of low resolution and low fidelity prototypes, that can be easily modified, tested and discarded based on fast insights. Foam, paper, rubber, and fabric are some commonly available materials that are used in this way of prototyping as they can be easily modified using simple hand-tools. Similarly, ease of use makes glue, tape, needles and ropes important tools to have in the toolbox as it enables rapid assembly and continuous alterations of prototypes.

In the context of non-rigid materials, the project case for soft prototyping is the conceptual design of a soft actuator for animatronic robots. This project had specific demands as it required a cylindrical tube to change its shape in order to simulate a human airway. This meant having control over deformation, not introducing rigid elements, as well as having the opportunity to occlude the cylindrical airway shape as seen in Figure 1.

In this project, physical alterations and iterating different functional principles, was vital in order to visualize and generate solutions. Using simple materials such as flexible foam and plastic film, the designers could realize ideas and test them without spending time on Computer-Aided Design (CAD) or machining. In addition, rapid design changes and iterations enabled large throughput of potential concepts. While prototype resolution and fidelity remained low, the designers could quickly test out principles for both the actuation as well as the geometrical constraints of the simulated airway.

https://doi.org/10.1017/dsd.2020.120 Published online by Cambridge University Press



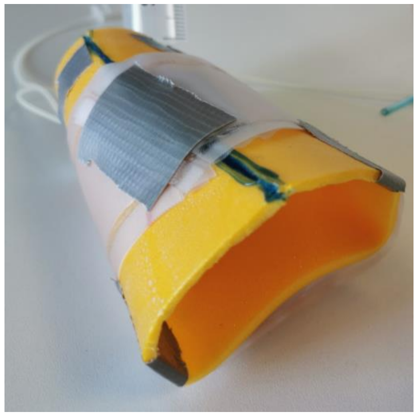


Figure 1. Cylindrical airway shape built using soft prototyping

This case showed the importance of exploratory activities and that also flexible/non-rigid concepts and materials should be considered in early design. While this case attempted to mimic the flexible nature of a human airway, the actuation and geometrical constraints made this a complex design challenge that was hard to visualize. Trough rough physical modelling and having the opportunity to test often the designers could rapidly generate a functional concept that otherwise could have been hard to come up with.

2.3. Prototyping with living hinges

Creating 'flexure patterns' or 'living hinges' is a method utilized to change an object's geometry to enable flexible properties. By cutting or removing material in specific patterns, the mechanical abilities of plate geometries is altered making it possible for rigid materials, such as fibreboard, to bend and flex. Swatches, or premade patterns, can be a time-saver when applying this method. Based on cut orientation and density of the pattern, it is possible to control both axis and radius of deformation. While this method enables rapid generation of flexible concepts using machines such as a laser cutter, design space is constricted to the use of plate geometries. Additionally, both resolution and fidelity are governed by manual assembly of the cut parts.

The project case illustrating this method of prototyping is the conceptual prototyping of a new chest for resuscitation mannequins. The project was challenged with generating new concept ideas for a training mannequin by mimicking the human chest. This required concepts to have similar shape and deformation as a chest would, which inspired the creation of multiple concepts whereas two can be seen in Figure 2. The designers used simple 2D drawings to run the laser cutting operation, which meant little time was spent in virtual prototyping environment. In addition, the speed of fabrication, made it possible for the design team to create multiple design alternatives, hence, adapt quickly to new insights. As seen in Figure 2, the design experimented with different cut patterns that enabled one or two degrees of freedom for the compression mechanism.

Figure 2. Prototypes built using living hinges

https://doi.org/10.1017/dsd.2020.120 Published online by Cambridge University Press

Using living hinges as a means of creating flexible prototypes is trough the case example shown to be fast and allowing quick design iterations. The method of fabrication benefitted the speed and accuracy of using a laser cutter and having customizable cut patterns enabling this method to create new flexible parts in a matter of minutes. However, both fidelity, resolution and design freedom are for this method governed by the need of a cut pattern to achieve flexible behaviour, rather than leveraging non-rigid material properties. This is further restricting the method to using plate geometries, reducing the design freedom yielded by this method.

2.4. Prototyping with fused filament fabrication in flexible materials

Fused Filament Fabrication (FFF), is an additive manufacturing method that uses a layer by layer construction technique of a part modelled in a virtual CAD environment. FFF uses melting of polymer materials on a bed, to build geometrical structures from the bottom up. This method enables the designer to make parts with high fidelity, high resolution and little constraints regarding the geometrical design space, as both layer height and support structures can be optimised to the part geometry. Thermoplastic Polyurethane (TPU) is one example of a flexible FFF material that allows the FFF process to manufacture parts with both complex geometry and flexible properties. As flexible FFF prototypes are melted together, design alterations or rework is highly limited. Therefore, in order to change prototype attributes, changes first must be done virtually before a new part can be manufactured.
The case illustrating this prototyping method, is the development of a spring mechanism for a personal drug mixing container. For this case example, there was a need to reduce mechanical complexity as part count and assembly made the concept prone to jamming and backlash. By drawing on the advantages of FFF with flexible materials it was possible to make a conceptual model, as seen in Figure 3. of high fidelity and resolution, fitting the assembly constraints.

Figure 3. Prototypes built by using fused filament fabrication

In the case it was observed that by utilizing this method of prototyping, designers can simplify a mechanical system that enables easy handling, durability, and robustness. This method of prototyping takes place in both the virtual and physical design space. Therefore, CAD is needed before a physical model is available. An interesting aspect of this project is how the flexible spring structure was tested in the context of the drug container and shown to be unstable because of buckling effects. This could in some cases make the mechanism self-lock and hinder the retraction of the spring. This insight would be hard to obtain without the physical modelling and testing, which highlights the importance of this method as an iterative process. In order to adapt to this insight, the designers had to generate new models. Models that first had to be changed or generated in virtual prototyping environment before creating a new physical model. The virtual rework was done in minutes, but the manufacturing needed hours. The speed of this method is therefore rather low because it does not allow for rapid adjustments and testing. However, the fidelity and resolution it can produce is high.

https://doi.org/10.1017/dsd.2020.120 Published online by Cambridge University Press

2.5. Prototyping using moulding with flexible materials

Using moulding as a means of prototyping flexible materials is a high fidelity and resolution way of prototyping parts for testing. The workflow consists of creating a virtual prototype of the intended part or the mould for the part. Generally, using moulding as a method for prototyping is tedious as both moulds and corresponding parts needs to be fabricated. The outcome is however, of high fidelity, and the design freedom the method yields is acceptable as it allows complex outer geometry and surface finish but is restricted concerning internal structures and geometries.

The case example used to demonstrate this prototyping method, a challenge of fabricating and testing a phantom heart in a composite silicone material is considered. This phantom heart was to be used for research on heart flow and ultrasound imaging. To create this phantom, a contrast material was added the silicone to make the part visible on the ultrasonic imaging equipment. For the phantom to achieve flow conditions as found in real heart, the physical model had to replicate key anatomical features of a human heart. High resolution and fidelity were hereby needed to fulfil flow conditions, shape resemblance, and deformation resemblance to a real heart. An additional point to make is that the phantom needed to withstand the internal pressure present as the closed liquid system where to be displayed in air rather than dispersed in liquid. In this case, the heart was first prototyped virtually to create the mould negatives using fused filament fabrication, as shown in Figure 4, and was then moulded using silicone, as shown in Figure 5.

Figure 4. Mould negatives built using rigid materials that were used for moulding flexible phantom hearts with flexible material

https://doi.org/10.1017/dsd.2020.120 Published online by Cambridge University Press

Figure 5. A human heart represented by a virtual prototype (left and centre) and a physical phantom (right) built using moulding with flexible materials

This method of prototyping is relatively slow, as it requires several sequential steps to produce a testable prototype. The production of the mould negatives with FFF (both in CAD and the fabrication itself) is time consuming, and when coupled with the silicone curing time, the per-prototype iteration time until physical testing can start is long. While this case is illustrating mould fabrication using FFF, other methods such as clay casting, manual sculpting and subtractive manufacturing might be more suitable in other cases. As mould fabrication is a means of prototyping moulds, it is in the context of this paper worth noting that efforts in prototyping the moulds affect both attributes of the flexible porotypes as well as activity performance in terms of time and ability to alter designs. One example is how the resolution and fidelity of moulded parts are governed by the surface finish and geometrical design freedom of the moulds, giving FFF advantages of being little constrained in terms of geometry, but subjected to rough surfaces as a result of fabrication technique.

The design freedom of prototyping with moulding is moderate as it can enable any outer geometry but is clearly limited to internal structures and holes making demoulding a challenge. Further it is worth noting that silicone, as one alternative material, only adheres to silicones making it hard to do later rework, or interface using non-silicone materials. Modifications to the prototype itself after the casting is done is therefore both challenging and time consuming. This results in a slow and rigid method considering adapting to insights, as design alterations often would require new mould designs, mould fabrication/alteration, curing time, and additional required assembly.

3. Discussion

The presented cases have been investigated, insights have been gathered and evaluated as subjective measurements on the method's effect on design activity and output. This evaluation is presented in Table 1. and subsequently visualised in Figure 6. While the evaluation has elements from established prototype descriptions in literature, it is not intended to be prescriptive or generalisable to other prototyping evaluations, but rather a tool to guide the discussion in this article. This evaluation has been conducted by four coders (from the same research team) that have rated the four project cases, and agreed on the values ranging from 1 to 3 presented in Table 1. As this rating is only used for discussing and visualizing the findings from the four cases, inter-coder reliability has not been measured. In this section, the authors describe how both prototyping activities and outcome are impacted by the methods employed in the various cases.

Table 1. The four prototyping methods evaluated in terms of speed, modifications, design freedom, fidelity and resolution. Each method is rated from 1, being low, to 3, being high

	Speed	Modifications	Design Freedom	Fidelity	Resolution
Soft Materials	3	3	2	1	1
Living Hinges	3	2	1	2	2
Flexible FFF Materials	2	1	3	3	3
Flexible Moulding	1	1	2	3	3

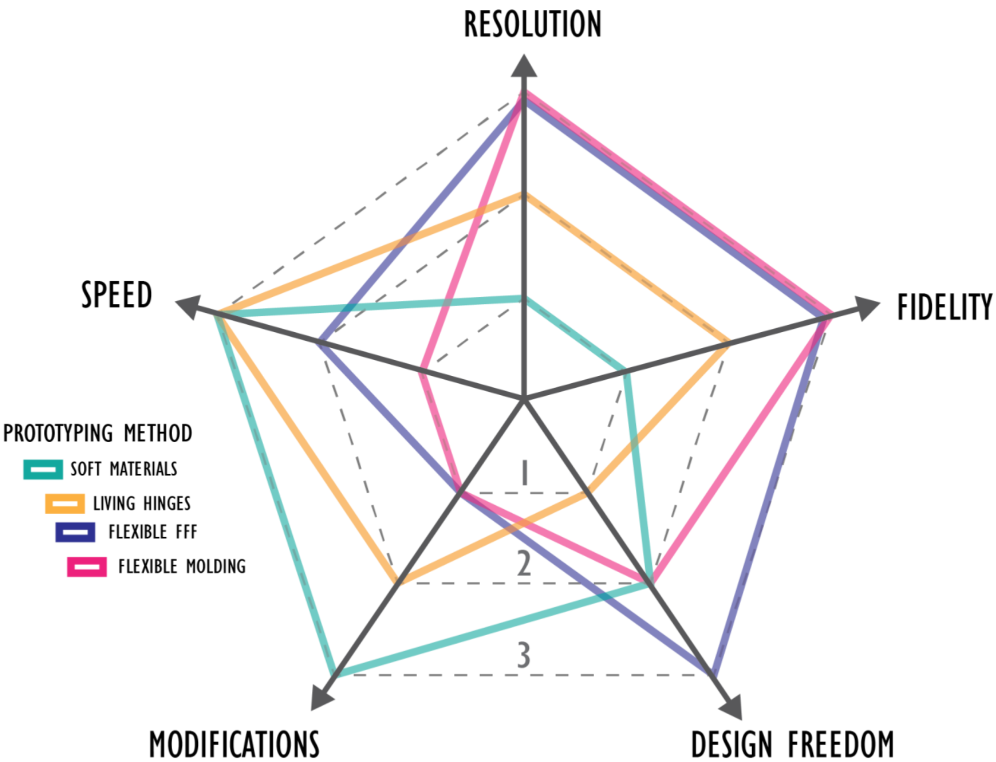


Figure 6. Visualisation of the four prototyping methods evaluated in terms of speed, modifications, design freedom, fidelity and resolution

From inspecting Figure 6, each of the four prototyping methods have distinct differences. Moulding and FFF with flexible materials are both considered methods enabling high fidelity and resolution outcome in terms of accurate and refined prototypes. Because of time and constraints regarding design of moulds and moulding processes, FFF is considered faster and allowing more design freedom than moulding. These methods are however slow and allowing less modifications than both soft prototyping and prototyping using living hinges. The speed of using soft prototyping and living hinges as prototyping approach is high, as these methods require minimal work in CAD environments and not relying on slow machining operations. This makes soft prototyping and using living hinges important methods in gaining insights fast, as learning in this context is a result of creating and testing of prototypes. Further adapting to new insights, soft prototyping is better suited when experimenting and altering designs, while concepts using living hinges require more effort to alter.

https://doi.org/10.1017/dsd.2020.120 Published online by Cambridge University Press

In the early stages of design, prototyping with soft materials and living hinges enable the exploration of new solution spaces. They enable rapid creation of physical prototypes that can be used to test and evaluate non-rigid functions for conceptual designs. The transition to prototyping methods such as FFF and moulding with flexible materials should be done when the required resolution and fidelity is higher. However, these methods require more time and resources for each testable iteration, and should if possible, be delayed to later stages in design. However, when prototyping, the added complexity from using non-rigid materials and creating non-rigid concepts must be acknowledged, possibly creating integration problems at the later stages of design. In addition to opening a potential solution space, flexible mechanical properties make non-rigid concepts harder to predict and replicate. This also applies to the context in which such concepts are intended and how they interface with both users and functional objects. This strengthens the argument for prototyping both early and often to effectively identify and address the uncertainty associated with non-rigid materials, concepts, and context of use in engineering design projects.

Selection of cases and subsequently methods of prototyping with non-rigid materials have in this paper been focused towards exemplifying individual approaches and evaluating them to highlight advantages and limitations for each method. While these methods are not intended as a representation of all methods utilizing non-rigid materials, the selection are highlighting core differences and how they affect prototyping activities and outcome. Since the presented methods are gathered from individual cases, the comparison of methods is limited in terms of workload put into each project, and success criterion and intention for each prototype. However, the cases examples show the applicability of using non-rigid materials and corresponding prototype approaches for leveraging such materials in contexts ranging from low-resolution experiments, to refined models intended for intricate test scenarios.

This paper has evaluated methods of prototyping in order to highlight some of the trade-offs designers must consider when choosing approach for exploring non-rigid concept ideas. With these considerations, selecting appropriate method for the concept to explore, can ensure faster iterations, and being more adaptable to revealed insights. Then designers can decide to invest more time and resources in refined prototypes, when uncertain elements have been revealed trough testing and iterating using more rapid prototyping techniques. This is especially important in the context of non-rigid concepts and in order to leverage non-rigid materials, as predicting mechanical behavior, test conditions, and time and complexity of fabrication all introduce uncertain elements in early design stages.

4. Conclusion

In this paper, four methods of prototyping with non-rigid materials have been presented and illustrated trough cases gathered from individual engineering design projects. From these four cases, insights on the utility and implications of using complex non-rigid materials in engineering design practise has been discussed with respect to five specific dimensions; speed, modifications, design freedom, fidelity and resolution. Lastly, implications and normative recommendations for applying these prototyping methods with non-rigid materials in engineering design projects have been discussed.

Acknowledgement

This research is supported by the Research Council of Norway through its industrial Ph.D. funding scheme, project number 290404.

References

Auflem, M., Erichsen, J.F. and Steinert, M. (2019), "Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators", *Procedia CIRP*, Vol. 84, pp. 572-578. https://doi.org/10.1016/j.procir.2019.04.202

Bryan-Kinns, N. and Hamilton, F. (2002), "One for All and All for One?: Case Studies of Using Prototypes in Commercial Projects", *Proceedings of the Second Nordic Conference on Human-Computer Interaction*, ACM, New York, NY, USA, pp. 91-100.

Buchenau, M. and Suri, J.F. (2000), *Experience Prototyping*, ACM, New York, NY, USA, pp. 424-433.

Camburn, B.A. et al. (2013), "Methods for Prototyping Strategies in Conceptual Phases of Design: Framework and Experimental Assessment", *Volume 5: 25th International Conference on Design Theory and*

Methodology; ASME 2013 Power Transmission and Gearing Conference, presented at the ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, ASME, Portland, Oregon, USA, p. V005T06A033.
Dow, S.P. et al. (2011), "Parallel prototyping leads to better design results, more divergence, and increased self-efficacy", *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 17 No. 4, pp. 18:1-18:24.
Dow, S.P., Heddleston, K. and Klemmer, S.R. (2009), "The efficacy of prototyping under time constraints", *Proceedings of the Seventh ACM Conference on Creativity and Cognition*, ACM, pp. 165-174.
Foehr, A.G.C., Stuecheli, M. and Meboldt, M. (2015), "Efficient design evaluation through the combination of numerical and physical computations", *DS 80-6 Proceedings of the 20th International Conference on Engineering Design (ICED 15) Vol 6: Design Methods and Tools - Part 2*, Milan, Italy (accessed 27-30.07.15).
Houde, S. and Hill, C. (1997), "What do Prototypes Prototype?", In: *Handbook of Human-Computer Interaction*, Elsevier, pp. 367-381.
Jensen, L.S., Özkil, A.G. and Mortensen, N.H. (2016), "Prototypes in engineering design: Definitions and strategies", *14th International Design Conference International Design Conference*, Design Society, pp. 821-830.
Jensen, M.B., Balters, S. and Steinert, M. (2015), "Measuring Prototypes - A Standarized Quantitative Description of Prototypes and their Outcome for Data Collection and Analysis", *Proceedings of the 20th International Conference on Engineering Design (ICED15), presented at the International Conference on Engineering Design*, Milan.
Kriesi, C. et al. (2016), "Creating Dynamic Requirements through Iteratively Prototyping Critical Functionalities", *Procedia CIRP*, Vol. 50 No. Supplement C, pp. 790-795.
Lauff, C., Kotys-Schwartz, D. and Rentschler, M.E. (2018), "What is a Prototype? What are the Roles of Prototypes in Companies?" *Journal of Mechanical Design*, Vol. 140. https://doi.org/10.1115/1.4039340
Leifer, L.J. and Steinert, M. (2011), "Dancing with ambiguity: Causality behavior, design thinking, and triple-loop-learning", *Information Knowledge Systems Management*, Vol. 10 No. 1-4, pp. 151-173.
Lim, Y.-K., Stolterman, E. and Tenenberg, J. (2008), "The Anatomy of Prototypes: Prototypes As Filters, Prototypes As Manifestations of Design Ideas", *ACM Trans. Comput.-Hum. Interact.*, Vol. 15 No. 2, pp. 7:1-7:27.
Menold, J.D. (2017). "Prototype for X (PFX): A Prototyping Framework to Support Product Design", Available at: https://etda.libraries.psu.edu/catalog/13760jdm5407 (accessed 7 May 2019).
Schrage, M. (1993), "The Culture(s) of prototyping", *Design Management Journal (Former Series)*, Vol. 4 No. 1, pp. 55-65.
Viswanathan, V. and Linsey, J. (2011). "Design Fixation in Physical Modeling: An Investigation on the Role of Sunk Cost", pp. 119-130.

C3

Facing the FACS-Using AI to Evaluate and Control

Facial Action Units in Humanoid Robot Face Development

Frontiers in Robotics and AI

Marius Auflem, Sampsa Kohtala, Malte F Jung, Martin Steinert (2022),

frontiers | Frontiers in Robotics and AI

ORIGINAL RESEARCH
published: 14 June 2022
doi: 10.3389/frobt.2022.887645

Facing the FACS—Using AI to Evaluate and Control Facial Action Units in Humanoid Robot Face Development

Marius Auflem[1]*, *Sampsa Kohtala*[1], *Malte Jung*[2] and *Martin Steinert*[1]

[1]*TrollLABS, Department of Mechanical and Industrial Engineering, Faculty of Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway,* [2]*Robots in Groups Lab, Department of Information Science, Cornell University, Ithaca, NY, United States*

This paper presents a new approach for evaluating and controlling expressive humanoid robotic faces using open-source computer vision and machine learning methods. Existing research in Human-Robot Interaction lacks flexible and simple tools that are scalable for evaluating and controlling various robotic faces; thus, our goal is to demonstrate the use of readily available AI-based solutions to support the process. We use a newly developed humanoid robot prototype intended for medical training applications as a case example. The approach automatically captures the robot's facial action units through a webcam during random motion, which are components traditionally used to describe facial muscle movements in humans. Instead of manipulating the actuators individually or training the robot to express specific emotions, we propose using action units as a means for controlling the robotic face, which enables a multitude of ways to generate dynamic motion, expressions, and behavior. The range of action units achieved by the robot is thus analyzed to discover its expressive capabilities and limitations and to develop a control model by correlating action units to actuation parameters. Because the approach is not dependent on specific facial attributes or actuation capabilities, it can be used for different designs and continuously inform the development process. In healthcare training applications, our goal is to establish a prerequisite of expressive capabilities of humanoid robots bounded by industrial and medical design constraints. Furthermore, to mediate human interpretation and thus enable decision-making based on observed cognitive, emotional, and expressive cues, our approach aims to find the minimum viable expressive capabilities of the robot without having to optimize for realism. The results from our case example demonstrate the flexibility and efficiency of the presented AI-based solutions to support the development of humanoid facial robots.

Keywords: humanoid robots, artificial intelligence, medical simulation, robot development, facial action units, facial expression

Edited by:
Gabriel Skantze, Royal Institute of Technology, Sweden

Reviewed by:
Jonas Beskow, Royal Institute of Technology, Sweden
Christian Werner Becker-Asano, Stuttgart Media University, Germany

***Correspondence:**
Marius Auflem
mariuauf@stud.ntnu.no

Specialty section:
This article was submitted to Human-Robot Interaction, a section of the journal Frontiers in Robotics and AI

Received: *01 March 2022*
Accepted: *11 May 2022*
Published: *14 June 2022*

Citation:
Auflem M, Kohtala S, Jung M and Steinert M (2022) Facing the FACS—Using AI to Evaluate and Control Facial Action Units in Humanoid Robot Face Development. Front. Robot. AI 9:887645. doi: 10.3389/frobt.2022.887645

1 INTRODUCTION

Humanoid robots with expressive attributes that encourage social interactions with humans are an important topic for various fields of research and industrial contexts (Breazeal, 2009). With recent technological advancements pushing the boundaries for complex behavior and humanlike appearance, several robots designed to look and behave like humans have been developed (Becker-Asano and Ishiguro, 2011; Ameca 2021; Cominelli et al., 2021; Sophia 2021). New materials, accessible electronics, rapid prototyping, and artificial intelligence (AI) have all been key enabling factors for the emergence of these uncannily realistic (humanlike) robots (Oh et al., 2006). Not only are they approaching a realistic visual resemblance to humans (Mori et al., 2012), but by movements and simulated cognition, robots are enabling eerily realistic interactions with people (Pan et al., 2020). Hence, how the robot looks, behaves, and reacts are important aspects to consider when designing solutions for human-robot interaction applications (Cameron et al., 2018; Ghazali et al., 2018). In this context the face of the robot is particularly important for non-articulate responses like body language, expressions, and sudden reactions. The synergetic effects of realistic appearance and complex humanlike behavior, i.e., gaze, expressions, and motor abilities, have been identified as essential factors (Minato et al., 2004). Hence, novel robots with expressive capabilities have facilitated research on mimicking, synthesizing, and modelling of robotic face movements (Wu et al., 2009; Magtanong et al., 2012; Mazzei et al., 2012; Meghdari et al., 2016). Furthermore, researchers aim at providing insights on how we evaluate, recognize, respond, react, and interact with such social and emotional humanlike robots (Hofree et al., 2018; Hortensius et al., 2018; Jung and Hinds, 2018; Tian et al., 2021).

However, while advanced expressive robots enable us to explore ways to achieve humanlike face movements, there is a lack of tools and methods supporting such robots' (early-stage and ongoing) development. Specifically, accessible, fast, and easy-to-use tools and methods aiding in prototype evaluation and control of new humanoid robots. Furthermore, these tools should not be limited to specific hardware architecture and should moreover, provide objective feedback on obtainable face movement to the designers. For example, characterizing the relations between actuator input and resulting face movement could be critical to understanding (and improving) humanoid robots' design. Such tools may enable simplified control of these robots by using human face parameters, such as facial action units (AUs), to create a variety of custom facial responses and expressions (Ekman and Friesen, 1978). For evaluating generated expressions or movement, automatic visual inspection leveraging human face tracking software and AI applications could be purposeful to mitigate designer (or user) biases. Additionally, this could speed up learning the potentials and limitations of hardware prototypes, given that different use-cases yield different design constraints and needs for future robots. Hence, resources that inform development of humanoid robots are essential as these will become custom in a variety of industrial contexts.

This paper presents the use of open-source computer vision and machine learning (ML) methods as tools for supporting the development and evaluation of robotic faces. The approach of utilizing these tools is showcased in a development project of a new humanoid robot with facial movement capabilities intended for healthcare learning applications. The development project has utilized a highly iterative approach for designing, building, and testing prototypes. This is an effective way of dealing with ambiguity caused by complex design problems before functional requirements have been established. However, a challenge with prototyping a robotic face is the unavoidable subjective effects of designing for social and communicative attributes such as face appearance, movement, and subsequently, expressions. As both user and designer biases are evident, a flexible method for rapidly evaluating prototypes and informing design decisions is needed. AI is a broad field that includes computer vision and machine learning algorithms that can make decisions on a par with humans. Hence, we present a method for utilizing open and readily available AI-based software solutions without requiring specialized hardware or human-in-the-loop for obtaining facial action units from random motor movements to create a control model for the robot. Furthermore, the generated data are used to inform the current robot design by using AU correlations, both between other AUs and against motor actuation modes. The presence or absence of AUs in the obtained data is used as a performance measurement for the prototype capabilities. The proposed method is also applicable for any type of mechanically actuated robotic face resembling a human, regardless of the number of control units or face characteristics.

This is a proof-of-concept intended to showcase the applicability of using available AI tools for design of humanoid robots, and the advantages and limitations of using such tools during development. Furthermore, we want to give an outlook and highlight the potential benefits of using this method in development of flexible and customizable humanoid robots for healthcare learning applications through a development case example. To summarize, the aim of our method is to 1) gain objective and actionable insights for early-stage robotic face development, 2) rapidly generate AU data and train a control model tailored to the specific robot face appearance and actuation capabilities, and 3) to objectively evaluate the current robot design and control model using various datasets.

2 BACKGROUND

2.1 Robot Motion and Behavior

State-of-the-art robots designed to look and behave like humans are complex (and expensive) equipment, encompassing delicate hardware to achieve many degrees of freedom for facial movement (Faraj et al., 2020). Expressive motions and behavior are important as they can communicate the internal state of the robot and convey information about its' affect, fatigue, intent, style, and personality (Venture and Kulić, 2019). Expressive robots are often developed by inspiration from the anatomy of real humans using biomimicry (Hanson et al., 2002;

Pioggia et al., 2004; Hashimoto et al., 2006) or by mechanical modeling of target output movements predetermined from human face capabilities (Magtanong et al., 2012). Subsequently, control and evaluation of such robots becomes challenging utilizing manual operation and hand-coded motion sequences (Venture and Kulić, 2019). Moreover, due to the complex system interactions and dynamic behavior, such robots may need to be analyzed after being fully developed (Ishihara et al., 2021). Therefore, exploring new and intuitive ways to control (and evaluate) humanoid robotic faces has been a topic of interest. For example, to control using machine vision software and AI to recognize specific human expressions and mirroring, recreating, or reacting using a robotic agent (Kim et al., 2006; Silva et al., 2016; Todo, 2018). Similarly, AI applications have been utilized to analyze robots' facial capabilities and automatically learn various expressions (Wu et al., 2009; Mazzei et al., 2012; Meghdari et al., 2016; Chen et al., 2021; Rawal et al., 2022). For automated control of robotic faces, using AUs is valuable as it becomes a transferal unit of facial movement, representing both the human action and the robot's actuation capabilities (Lin et al., 2011; Lazzeri et al., 2015; Faraj et al., 2020). However, limitations of these approaches include the sequential development of robot and control systems, thus restricting rapid design cycles and performing simultaneous and cross-disciplinary improvements. Furthermore, the tools and methods deployed are often restricted to a set of static expressions or are limited to specific hardware or appearance, making them less suited in the early, conceptual, and prototype-driven development of humanoid robots. The novelty of our approach compared to existing solutions is the possibility of effortlessly capturing AUs from different robot designs with varying degrees of freedom and using this information to support the development process and to rapidly create control models for expression synthesis. While Wu et al. (2009) used the correlations between AUs and servos of a high degree of freedom robot face to create a linear mapping between the two, we highlight the importance of additionally using AU to AU correlation analysis to gain valuable information and support the development of expressive robots. We also show how modern and scalable ML algorithms can quickly approximate both linear and non-linear AU to servo mappings with limited degrees of freedom.

2.2 Case of a Medical Simulator Face With Expressive Capabilities

Simulation-based medical training and education is an area where humanlike robots already play an important role. In this context, the robot, often referred to as the manikin, portrays a patient that needs care and treatment. Having evolved from static and limited anatomical chassis', manikins have gained a range of simulated physiological and cognitive abilities enabled by remote-controlled operation, and autonomous or semi-autonomous control systems (Cooper and Taqueti, 2008). These manikins have excided far beyond their initial use-cases of psychomotor skills training and routine practice for medical students. However, the non-articulate communicative aspects of such robots remain limited, often having a generic and static appearance incapable of performing facial movements to render expressions, communicate, react, or simulate important medical cues (Lalitharatne et al., 2020). For training scenarios where medical simulators (i.e., robots) are used instead of real patients, the simulators should accommodate multimodal tasks, such as combining data acquisition, interventions, and clinical assessment (Pourebadi and Riek, 2018). This would enable the simulated patient's facial movements and behavior to be observed and used for evaluating medical conditions, cognitive abilities, and emotional states to pose a diagnosis. Furthermore, the standardized appearance of simulators is limiting, as it is important to capture various patient characteristics to reflect the diversity found in the general population. Age, gender, ethnicity, and cultural traits should therefore be adequately captured by robots' appearance and behavior (Hortensius et al., 2018) to enable nuanced and ecological valid training scenarios and improve medical simulation by ensuring inclusivity and important training variance (Conigliaro et al., 2020).

For a robot to simulate a human patient, clinical cues such as pain response, altered cognitive state, and emotional gestures need to be adequately captured for learners to recognize and perform the required actions for treatment (Moosaei et al., 2017). Hence the goal is to trigger the appropriate responses from users by the robots' actions in simulated scenarios. This poses the challenge of determining sufficient facial movement capabilities for the different use-cases of the robot. Furthermore, trade-offs concerning scaling potential, robustness, and integration in existing equipment for clinical simulation training needs to be addressed. Since humanoid robots for healthcare learning applications require several anatomical features to enable clinical interventions and facilitate training and routine practice, the available design space is constrained. A multifaceted design problem is therefore inevitable, where both non-verbal communication and physical interventions are required to ensure ecologically valid training-scenarios. In addition, the facial movements of robots may introduce uncanny effects, aversion, misunderstandings, and expectation gaps (Kwon et al., 2016). To approach these challenges, there is a need for exploring and characterizing the capabilities of robots by common parameters such as AUs. Hence, we ought to explore the minimum viable expressive capabilities, and simultaneously uncover the expressive potential the robot can achieve given contextual design constraints. We have developed a humanlike robotic face prototype with facial actuation capabilities to highlight these challenges with potential solutions by exploring AI tools to evaluate and inform the current design. Using the robot as a sandbox we have generated a control model utilizing intuitive and high-level instructions by AU parameters instead of manual control and pre-programmed sequences.

2.3 The Prototype

The robotic face prototype consists of a silicone-rubber skin with embedded skull-interface connectors, and a rigid skull chassis for

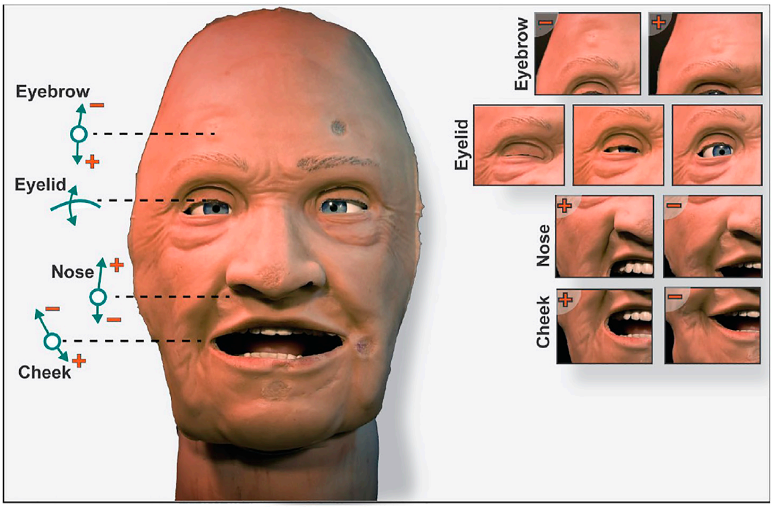


FIGURE 1 | The robotic face prototype with indicated actuation points, direction of travel, and achieved face movements.

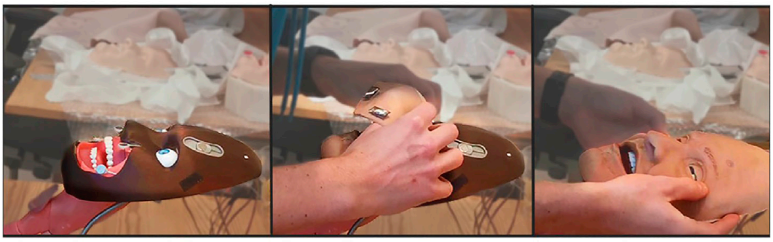


FIGURE 2 | Modular face design for changing the visual appearance using interchangeable face skins.

mounting the actuators and providing structural support. The prototype also has eyelids that can open and close, with static eyeballs and a non-articulated jaw. An anatomically accurate (upper) airway with teeth and tongue is also a part of the assembly. Furthermore, the prototype has six individual actuation points located at the root of the nose, corner of the mouth (cheek), and eyebrows. The location of actuation points is set to accommodate anatomical artefacts such as the airway, eyes, and tactile landmarks to enable clinical interventions. This naturally limits available design space, and thus servo motors for actuation are connected through individual push and pull cable arrangements. The actuation is done by six inexpensive 9 g RC servos that are controlled by a 18-channel Pololu Mini Maestro servo motor controller (Pololu - Mini Maestro, 2021). Additionally, two servos are mounted behind the eyeballs with a bar linkage to actuate the eyelids. The prototype with indicated connection points, actuation modes, and direction of servo movement (positive and negative) are shown in **Figure 1**.

To both alter the appearance and accommodate rapid and parallel design iterations, the robotic face and skull is designed with easily interchangeable outer skins, as seen in **Figure 2**. The current iteration of the facial skin portrays a geriatric male, where the proportions of the skull, relative distances between landmarks, and skull geometry remain fixed. Even though different characteristics pose facial anthropometric differences (Zhuang et al., 2010), the current use of (static) simulators with interchangeable skin appearances suggests the generalized hardware would enable a sufficiently wide design space to portray broad span of patients with similar anatomy. The skin is made from a highly flexible and low-density silicone rubber and is molded with variable thicknesses to simulate the tactility of facial tissue and muscles. The face to skull interface consists of mechanical snap connections embedded in the silicone, either interfacing the wire endpoints connected to the actuators, or as fixed anchoring to the skull. With this setup, various designs for the face skins can be explored, where flexible materials, appearance, geometry, and connector designs can be tested iteratively. Furthermore, alterations to the skull assembly can be tested with readymade face skins by changing actuator positions, connector positions, and structural geometry. This enables rapid design iterations and new prototypes to be generated and tested fast to address technical obstacles encountered through the development process.

Another important aspect of being able to easily swap face skins (apart from accommodating rapid design changes) is for the robot to portray various patients corresponding to the vast differences found in the human population, and thus increase training variance, inclusivity, and realism. This is especially important considering the robot is intended for healthcare learning applications (Conigliaro et al., 2020). As a result, the design objectives become more complex as the range of motion and connections between the face skin and skull should be tailored to the attached skin appearance. Since different appearances yield different face geometries, and thus mechanical properties of the artificial skin, it is not evident how each skin appearance would perform. Furthermore, as each appearance portrays a persons' age, gender, cultural, ethnical, and other visual attributes, control of the robot, and subsequentially evaluating the output, becomes a challenge given both designer and user biases.

The conceptual prototype was developed to test assumptions and elicit technical requirements. The development has been highly iterative, and thus several prototypes have been built to answer technical questions, as well as being a manifestation of the idea that can be presented to users to gain important feedback (Auflem et al., 2019). However, the challenge of the subjective matter of facial cues and expressions is obtaining actionable and objective data to measure the performance of prototypes by multimodal evaluation (Moosaei and Riek, 2013; Ege et al., 2020). This is particularly challenging when evaluating expressive robots due to the resolution and fidelity of the presented prototype being perceived differently, especially when users are unaware of the current state of development. Furthermore, clear limitations such as the absence of a complete head, fully covered skin, and missing anatomical landmarks not addressed by the current prototype may influence the feedback (Houde and Hill, 1997). This makes the evaluation process both challenging, tedious, and costly. Alternative tools are, therefore, needed in the early development stages to quickly improve the design of robotic faces.

To rapidly evaluate the prototype with the current appearance, we propose using a facial behavior analysis toolkit for capturing human face attributes. We seek to obtain data on the possible numbers and intensities of AUs the robot can generate to inform the hardware design. Using this data could also accommodate the many control methods and input data we want to explore for the robot. This flexibility could be beneficial since it is not evident how we want to control the robot in simulated scenarios. For example, control instructions could be given based on face expression coding using FACS or tracking an operator for real-time reenactment (Nygaard et al., 2018). Also, in a medical context, we could create simulation sequences using video of real-life events and patient assessments. Therefore, since structured data is available on human face movements using AUs, we want to use AUs as framework for controlling and evaluating the prototype (Lalitharatne et al., 2020).

2.4 Open-Source Computer Vision and Machine Learning Methods

The facial behavior analysis toolkit OpenFace 2.0 (Baltrusaitis et al., 2018) consists of state-of-the-art computer vision algorithms for automatically detecting and estimating facial landmarks, head pose, eye-gaze, and AUs. The toolkit has been used for understanding and recognizing mental states and social signals in human subjects within numerous fields. With its successful utilization on human subjects, and because the algorithms are trained and validated on actual humans, we believe it can be used for improving the way robotic faces are developed and controlled. More specifically, if we can detect the AUs of the robotic face and map them to its actuation units, we can synthesize more humanlike facial expressions while alleviating the potential designer bias. The dynamic AU recognition framework in OpenFace 2.0 also employs a person-specific normalization step (Baltrušaitis et al., 2015), making it adaptable for individual faces instead of relying on generalization.

The behavior of the robotic face can be modeled by applying various ML methods using facial expression analysis for the input and actuation parameters of the mechanical face as output. Scikit-learn (Pedregosa et al., 2011) is an open-source ML library for Python that is simple to use and provides efficient tools for predictive data analysis. By randomly moving the face-servos and capturing the resulting AUs through OpenFace 2.0, the robotic face can autonomously learn its facial expressions using a generic Python application. The intensity of an expression can then be adjusted on a continuous range by applying regression analysis to enable a more objective way of controlling a robotic face. Combinations of AUs can thus be used to estimate facial expressions instead of manually adjusting the servo angles and subjectively assessing the resulting expression. The same ML methods can also be applied to different robot designs since a model can be trained for each actuator with the same AUs as input, where essential and redundant AUs are weighted accordingly through the optimization algorithm. Furthermore, the methods are not restricted by the number of actuators since the automatic AU capturing approach allows the creation of large sample sizes, although using many redundant actuators may cause some of them to influence each other in opposite directions, thus worsening the training data. A method to quantify the relevance of each actuator is therefore beneficial before creating the control models.

The ability of a robot to show specific emotional expressions can be further evaluated using Residual Masking Network (RMN) by Pham et al. (2021); a state-of-the-art ML model for facial emotion recognition. RMN has achieved the highest classification accuracy of 74.14% on the widely used Facial Emotion Recognition (FER-2013) dataset (Goodfellow et al., 2013), which includes 35,887 images of facial expression of humans in seven categories: anger, disgust, fear, happiness, sadness, surprise, and neutral. By using ML-based facial emotion

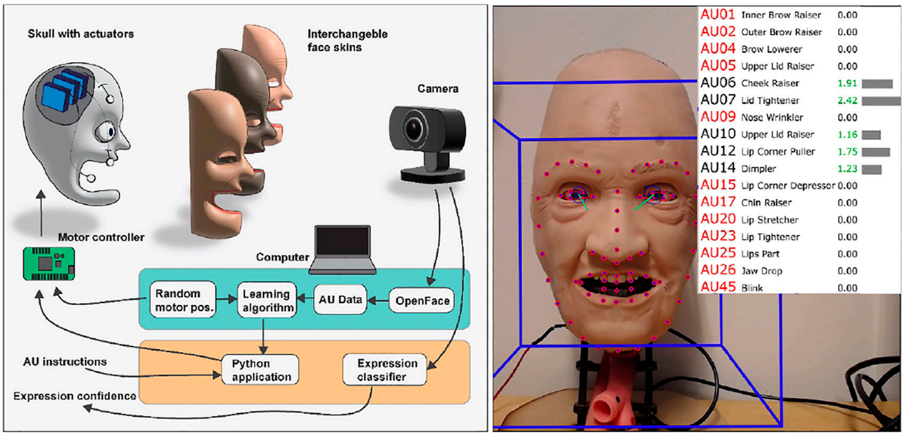


FIGURE 3 | Pipeline for controlling and analyzing the robotic face, with an example frame captured and analyzed with OpenFace 2.0.

recognition, we can evaluate the robot faster than finding experts or using multiple people through surveys.

3 MATERIALS AND METHODS

3.1 Experiment Setup and Pipeline

The pipeline for the proposed method is illustrated in **Figure 3**. It consists of two main parts: sending servo control instructions to the robotic face to change its appearance and capturing the characteristics of the resulting face using a camera and OpenFace 2.0. A Logitech C930e webcam was used for capturing frames of 1280 × 960 pixels to be analyzed with OpenFace 2.0 for predicting AUs. Both the webcam and Maestro servo controller were connected to a laptop through USB connections.

3.2 Data Collection

The range of motion for each servo was set manually through the Maestro control center and then normalized to a range between zero and one. The 17 AUs have intensity values ranging from 0 to 5, with 0 meaning the AU is not present, 1 representing presence at minimum intensity, and 5 representing presence at maximum intensity. Datasets containing AUs and servo positions were created by randomly moving the nose, eyebrow, and cheek-servos symmetrically along the vertical axis while recording the face through the webcam to extract AUs at a rate of approximately 20FPS. Due to the framerate of the webcam being affected by several factors including processing power and lighting (exposure), the framerate fluctuated slightly during real-time analysis. The AU estimations also fluctuated accordingly during static facial expressions. Therefore, based on a few initial tests, we found that capturing every 35th frame and taking the average AU values of the seven preceding frames reduced the variance and improved robustness while allowing the face servos to settle for each random position. The data collection and analysis sequence are illustrated in **Figure 4**.

Since OpenFace uses person-specific normalization, which assumes that a neutral facial expression is present in most of the analyzed frames in a video sequence, we need to find the frequency of randomly moving the servos and returning them to a neutral position to get the most consistent AUs. The effect of AU normalization was therefore analyzed by adjusting the frequency of returning the face to a neutral position (servos at 0.5) and capturing the combined mean of the resulting AUs to observe the overall variance. Based on an appropriate frequency of moving the face to a neutral position, a final dataset was captured where 500 frames were extracted containing random servo positions and the corresponding AUs. The datasets consist of the 17 AUs shown in **Figure 3** as independent variables and the three symmetrical servo positions shown in **Figure 1** as the dependent variables, excluding the eyelid.

3.3 Correlation Analysis

The Pearson correlation coefficients between every pair of variables were calculated using the “pandas.DataFrame.corr” module in Python and subsequently visualized using a correlation matrix. In addition to showing how the servos relate linearly to AUs, the correlation matrix can also indicate if AUs correlate to each other. If, for example, two AUs have a strong correlation, we can deduce that the current servo configuration cannot distinguish the respective AUs, thus providing valuable feedback to the development process. Redundant actuators can be discovered by detecting a corresponding lack of affected AUs, which can be disconnected to reiterate the process and analyze its effect. In addition to analyzing the current robot design, the correlation matrix is a tool for the development process to continuously discover possible improvements to the design and make sure the training data for the control model is appropriate.

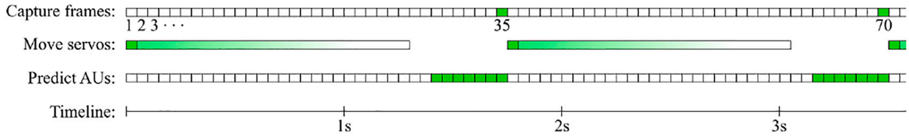


FIGURE 4 | Illustration of the data capturing and analysis procedure. Each square represents one iteration, where green denotes the stages where data are recorded.

TABLE 1 | Models and parameter values used for the grid search training procedure.

Model	Hyperparameters	Parameter values
Linear regression	—	—
Ridge regression	Alpha	0.01, 0.1, 1, 10, 100
SVR	C	0.1, 1, 10, 100, 1,000
	Gamma	1, 0.1, 0.01, 0.001
	Kernel	Rbf, linear, sigmoid, poly
MLP	Hidden layers	1, 10, 25, 50
	Activation function	Identity, logistic, tanh, relu
	Solver	Lbfgs, sgd, adam
	Alpha	0.00005, 0.0005
	Learning rate	Adaptive
	Max iterations	5,000
	Random state	Numpy.random.RandomState (42)

3.4 Modeling and Evaluation

We trained several regression models for predicting the servo positions with AUs as input. A supervised learning approach can be used since the input and output samples are collected directly through the pipeline. A few common supervised ML methods, namely linear regression (LR), ridge regression (RR), support vector regression (SVR), and multilayer perceptron (MLP) were used through the scikit-learn library (version 0.23.1) to test how different linear (with and without regularization) and non-linear learning algorithms affects performance. Exploring multiple ML methods without requiring considerable time is feasible if the number of samples is within a few thousand. The dataset was split into a training set (80%) and a test set (20%). Grid search with 5-fold cross-validation, provided by the GridSearchCV module from the scikit-learn library, was used when training models containing hyperparameters, as shown in **Table 1**. A random state value of 42 was used for the MLP model to preserve reproducible results. The best model for each servo was selected based on the lowest root-mean-square error (RMSE) and then evaluated with the test set.

Learning curves are also presented to assess the effect of training set size and to discover potential bias and variance in the data. The number of training samples was incremented by 15 and subsequently trained using the optimal parameters found from GridSearchCV and evaluated with 100 validation samples, where the RMSE for both sets was reported in the learning curves.

The servo positions for six expressions were then estimated using the best models by maximizing the relevant AUs based on FACS and setting the others to zero. The resulting facial expression of the robot was subsequently captured with the webcam and evaluated using RMN. Although the RMN model uses a softmax function for its output and consequently returns a confidence score for each of the seven output categories, we only report the top two predictions. Additionally, to demonstrate an alternative method for dynamically controlling the robot, we capture the AUs from a person in real time as input for the models to predict the servo positions, i.e., using mimicry to control the robotic face. This presented pipeline and experiments can be seen in a video added as **Supplementary Material**.

4 RESULTS

4.1 Effect of Action Unit Normalization and Correlation Analysis

The distribution of the average AU intensities based on the frequency of setting the robotic face to a neutral position is shown in **Figure 5**. As expected, displaying the neutral face constantly (100% of the frames) results in the lowest AU values with the least variation. The variation increases when introducing random positions in 10%–25% of the frames (90%–75% neutral faces), while the AU values increase substantially when only 66.7%–12.5% of the frames contain a neutral face position. A trade-off between speed and the effect of person-specific normalization was made when capturing the final dataset, where 75% of the frames contained neutral positions. Thus, we captured a total of 2,000 samples, for which 500 contained the random servo positions included in the final dataset. The distribution of intensities for each AU for the final dataset is shown in **Figure 6**.

The correlation matrix in **Figure 7** shows the Pearson correlation coefficients between all AUs and servos. Every correlation coefficient higher than 0.1 in absolute value is statistically significant ($p < 0.05$). Furthermore, each AU has a moderate to strong correlation with at least one of the servos, with AUs 5, 14, 20, 25, and 26 having the weakest (absolute values between 0.49–0.59), and AUs 1, 2, 9, 10, 15, and 23 having the strongest (absolute values above 0.8). While the strong correlation between AUs 1 and 2 is expected since they represent the eyebrows, the correlation for AUs 7, 9, and 10 with AUs 17 and 20 is less expected and may indicate a limitation in the current servo configuration.

4.2 Regression Analysis and Evaluation

The results on the test set after training and validating each model are shown in **Table 2**. Each method achieves low RMSE, with

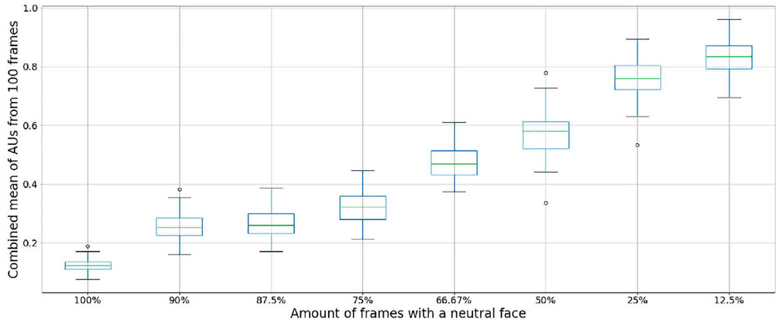


FIGURE 5 | Distribution of combined AU intensity values based on the frequency of a normal face analyzed with OpenFace 2.0.

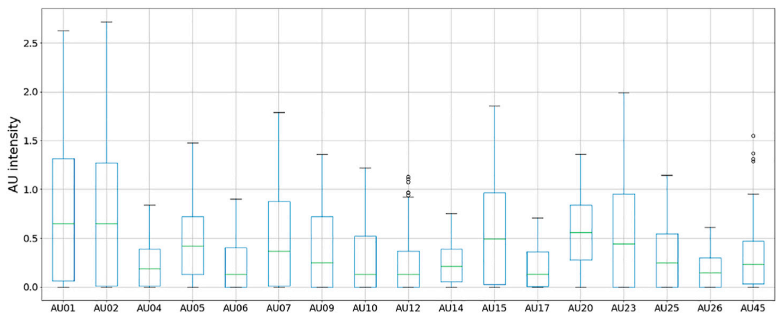


FIGURE 6 | Distribution of AU intensities captured from the final 500 random face-servo positions.

MLP having the lowest for each servo. Given the large number of independent variables, it is expected that SVR and MLP can capture some of the non-linear effects of the input and thus achieve lower errors.

Using the parameters found for the best models through cross-validation and grid search, we retrained each model using increments of 15 training samples, resulting in the learning curves shown in **Figure 8**. Here we can observe that relatively few training samples were needed to converge to an optimal solution for LR, RR, and SVR models. Low variance is observed by the comparable training and validation accuracy, indicating that no more than ~60 training samples are needed. The models do not appear to be affected by high bias given the low errors for both datasets, demonstrating that using AUs as features for predicting the servo positions is feasible and does not result in overfitting.

5 INTERPRETATION OF RESULTS AND EXPERIMENTAL EVALUATION

Based on the data we have generated from using OpenFace 2.0 and the results from training a control model for the robot, this section presents how we can gain valuable information and insights using the proposed tools when developing humanoid robots. **Sections 5.2**, **5.3** demonstrate how we can use the control model to generate facial expressions and dynamic motion through real-time reenactment.

5.1 Evaluating Robot Face Movement Capabilities Using Generated Action Unit Data

To obtain actionable data for improving the current design, we reviewed the AU intensities and respective correlations between variables seen in **Figures 6**, **7**. The distribution of AU intensities shown in **Figure 6**, indicates that only subtle expressions are obtainable as none of the captured AUs reach high levels ranging from 0 to 5. However, while the overall intensities of the recorded AUs are low, we can distinguish between the recognized output movements that are obtainable and the ones that remain idle given the current setup. We can further investigate the movement capabilities by looking at the correlation matrix in **Figure 7**, which shows that each servo has several strong correlations with action units. Based on these observations we can identify, and

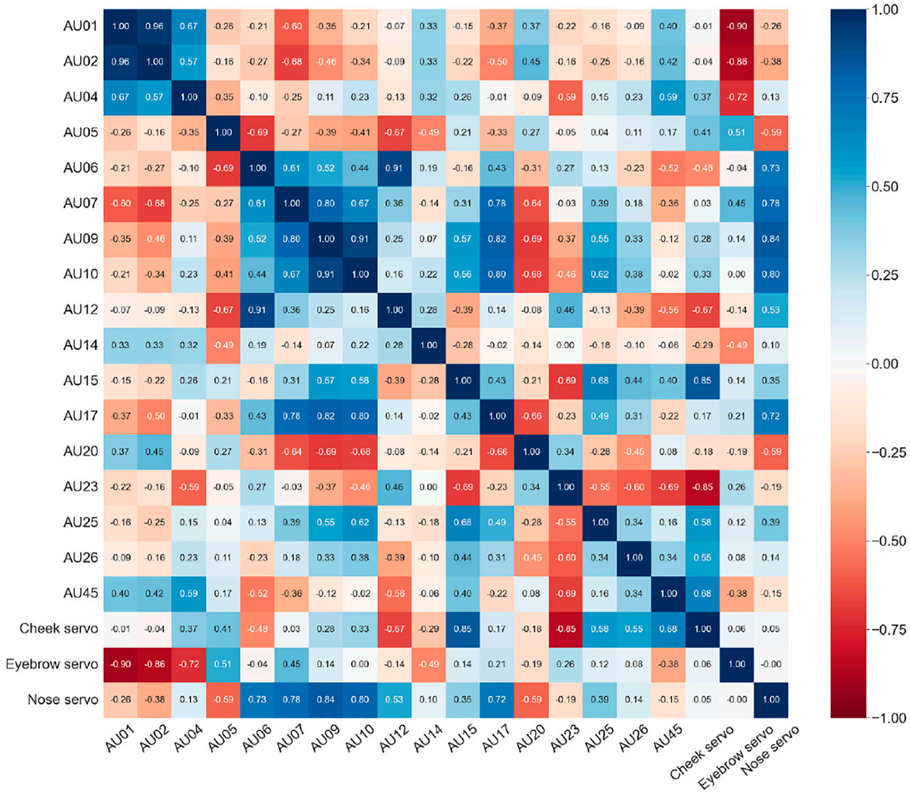


FIGURE 7 | Correlation matrix showing the Pearson correlation coefficients between AUs and servos. The blue to red gradient indicates the strength of positive and negative correlations, respectively.

TABLE 2 | RMSE on test set for best models. Lowest values are highlighted in bold.

	LR	RR	SVR	MLP
Cheek servo	0.054887	0.054788	0.054162	**0.040817**
Eyebrow servo	0.073310	0.073149	0.071086	**0.067251**
Nose servo	0.084295	0.083979	0.062116	**0.059183**

investigate, desirable AUs with weak intensities for improving the range of motion and location of actuation points. Furthermore, we can identify desirable correlations, such as servo to AU correlations, and inseparable AUs limiting isolated face movement given the current design.

Figure 7 shows a strong correlation between AU01, AU02, and the eyebrow servo. This is expected as they correspond to the inner and outer brow raiser, respectively. Their correlation with the eyebrow servo means that moving the servo upwards (in the negative direction) is identified as raising of the brows. However, differential control of these AUs is not possible, meaning that both the outer and inner brow will move simultaneously. Looking at the other action unit related to brow movements, we find that AU04 also correlates negatively with the eyebrow servos, which may seem counterintuitive since an increase in the eyebrow servo (**Figure 1**) should result in an increase in AU04 (brow lowerer). However, AU04, also known as the corrugator and depressor supercilli (Bartlett et al., 2005), represents the constriction of the area between the eyebrows in addition to brow lowering, which can happen simultaneously with AU01 and AU02. This means brow lowering is obtainable, although with a constrained path and weaker correlation indicating that a greater travel distance is desired to fully achieve this actuation mode. This is also highlighted in **Figure 6**, where the obtainable intensity of AU04 is significantly lower than that of AU01 and AU02, indicating that both travel and location of the brow connection point should be re-evaluated to enhance the brow lowering capability. Hence, we can improve the balance between the observed action units through physical design changes to create a model with more realistic brow movements.

In the mouth and nose area, we can observe a correlation between the action units AU09 (nose wrinkler) and AU10 (upper

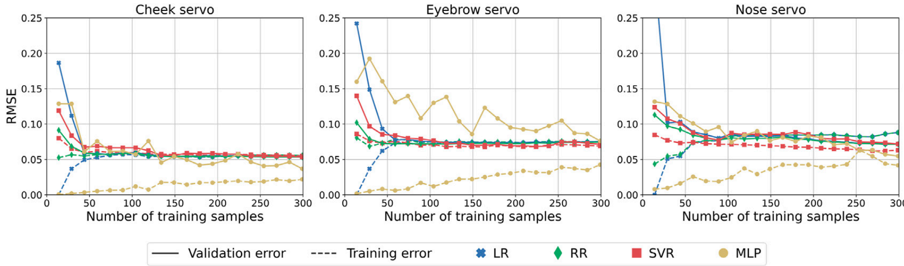


FIGURE 8 | Learning curves showing the effect of training samples on validation and training error.

lip raiser), and between AU06 (cheek raiser) and AU12 (lip corner puller). These AUs also correlate respectively to the nose and cheek servos. This indicates some of the dependent movements that are inseparable given the current design, caused by the artificial facial skin being a single deformable structure. Without additional anchoring of the face skin, deformations will be transferred throughout the elastic material and thus cause deformations ahead of the interface point along the line of travel in the face. While these connected movements are also found naturally in expressions by human faces, additional anchoring or actuation points would be required to separate them and enhance the movement capabilities of these AUs. Other correlations such as AU07 (lid tightener) and AU17 (chin raiser) with both the nose servo and to other action units, is also worth investigating. Since the eyelids are not actuated in the presented results, the intensity of AU07 is surprisingly high, suggesting additional anchoring around the eyes could be required. AU17 on the other hand, has a low intensity as seen in **Figure 6**, indicating that despite correlation with other AUs, the observed movement might be negligible. These findings suggest that either additional anchoring or additional actuation points could be beneficial to mitigate inseparable movements considering lower lips, jaw, and lower eyelid, and thus achieving a broader span of individual face actions. These are tradeoffs that should be evaluated alongside the required movement capabilities, as additional actuation modes would require additional actuators.

We can also identify which AUs are least represented in **Figure 6**, and review their correlation to the servos and other AUs in **Figure 7**. Here we highlight AU06 (cheek raiser), AU10 (upper lip raiser), AU12 (lip corner puller), AU14 (dimpler), AU17 (chin raiser), AU25 (lips part), and AU26 (jaw drop). These action units concern areas of the face located distant to the actuation points, indicating that we can only obtain movement in these areas by targeting nearby action units with a stronger presence in the current prototype. This is however not the case for AU25 and AU26, that shows modest intensities with few and weak correlations. Because the prototype has a fixed jaw and no lower lip actuation, these findings match our assumption that little movement should be obtained in these areas. These findings show that utilizing action unit data to describe and evaluate robotic faces to obtain actionable design input is purposeful.

5.2 Predicting Servo Positions Using Action Unit Instructions and Residual Masking Network

Six robot expressions were generated by predicting each servo position using the MLP models, with the relevant AUs maximized for the input. An image of each expression was then analyzed with RMN, resulting in the predictions shown in **Table 3**.

By sending instructions based on the FACS, we obtained a performance indicator on how the rendered expressions compare to sampled face expression images in the utilized dataset. These results indicate whether we can control the robot through AU instructions and render expressions that are recognized by ML. Furthermore, we can inspect the least successful expressions and correlate this back to the obtained AU data we have previously reviewed. Amongst the top confidence scores, only the Surprise and Happy expressions were correctly predicted. Disgust and sadness expressed by the robotic face are predicted similarly by RMN, both resulting in Angry and Surprise predictions with low confidences, which may be explained by the few representative AUs while sharing AU15. The limitations of the eyebrow actuation and the eyelids (not being active) could also be critical, since they impact the AUs for both anger, fear, and sadness expressions. The overrepresentation of Surprise predictions from RMN might be due to the underestimated importance of the eyelids, affecting AU5 (upper lid raiser), in combination with the non-articulated jaw affecting AU26 (jaw drop). The static jaw also limits obtainable AUs concerning mouth shapes and movement, which may explain why the RMN is less successful at predicting the negative expressions (where AUs around the mouth is particularly important). This suggest that additional actuation around the mouth and jaw is desirable to achieve a broader span of obtainable emotion expressions.

5.3 Real-Time Reenactment Using Recorded Action Units From Human Actor as Input

The MLP models were further utilized to experiment with real-time and automatic control of the robotic face. Here, we attempted to mimic facial movements by capturing AUs from

TABLE 3 | Generated face expressions with RMN predictions.

Expression	Anger	Disgust	Fear	Happy	Sadness	Surprise
Maximized AUs	4, 7, 23	9, 15	1, 2, 4, 5, 7, 20, 26	6, 12	1, 4, 15	1, 2, 5, 26
Resulting expression						
Top RMN predictions with confidence	Happy: 0.954 Surprise: 0.046	Angry: 0.401 Surprise: 0.368	Surprise: 0.900 Angry: 0.064	Happy: 0.433 Surprise: 0.421	Surprise: 0.565 Angry: 0.333	Surprise: 0.901 Happy: 0.099

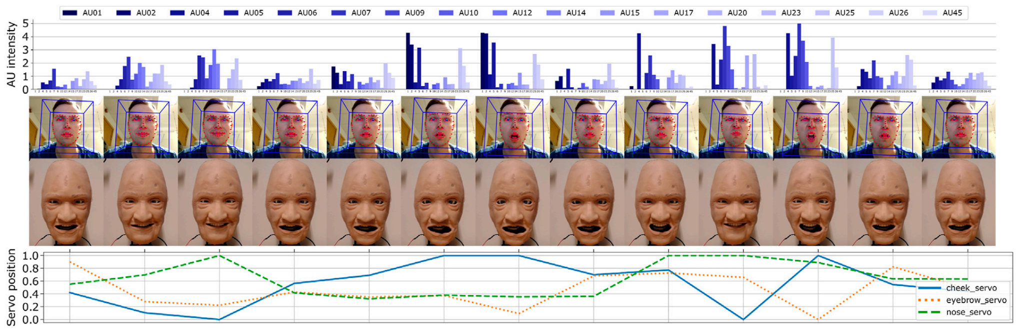


FIGURE 9 | Real-time reenactment using recorded AUs from a human actor to predict the servo positions.

a human actor using the same toolkit as previously described (OpenFace 2.0). By tracking the actors face through a webcam, we obtained and sent a stream of unfiltered AU intensities to the trained MLP models which then predicted and adjusted the servo positions in real time. A few frames from the face-tracking and the corresponding robot actuation are shown in **Figure 9**, including the raw input data (AU intensities) and the predicted output (servo positions). Although we can extract AUs at 20FPS, our application recorded at roughly 12FPS due to storing each captured frame as images, both of the human actor and robot simultaneously, where the robot face was updated at approximately 2FPS due to the delay in moving the servos.

6 DISCUSSION

6.1 Using Open-Source Software for Robotic Face Development

Our results demonstrate the use of open-source computer vision and ML methods as tools for supporting the development and evaluation of robotic faces. With careful consideration of how OpenFace 2.0 is utilized, such as the processing sequence shown in **Figure 4** and the effect of person-specific normalization, it is possible to reliably capture and utilize AUs to better understand the capabilities and limitations of a robotic face. Furthermore, since the approach is not reliant on the robot's specific characteristics or control parameters, it can be applied to any robotic face resembling a human.

We have also shown that AUs can be used to predict the servo positions for objectively controlling the robot, with errors as low as 0.041–0.067 RMSE. However, in our specific case, the learning curves and the correlation matrix reflect some limitations in the captured data, indicating the absence of dispersed and complex combinations of generated AUs. Since our robotic face can only produce a limited set of variability on AU combinations and intensities, the models may show low errors on test data using relatively few training samples while being unable to predict the servo positions appropriately for entirely different combinations of AUs never before seen. Suppose the goal is to simulate specific emotions (combinations of AUs). In that case, it is essential to focus on the correlation between AUs and reduce their interdependency through physical design choices, thus allowing more complex AU combinations to occur. We were nonetheless able to generate facial expressions by combining relevant AUs, as demonstrated in **Table 3**, and using mimicry to control the robot, showing generalizable tendencies of the trained ML models. These results are especially promising when considering the few actuators used in the robot design.

While the RMN model provides a quick and easy way to evaluate the expression of a robotic face, it may not be an accurate representation of how experts or the public perceives it. The model is trained on the FER-2013 dataset with labels corrected by humans, which may introduce biases, as well as having an unbalanced distribution of expressions. Interestingly, the human accuracy was roughly 65% on FER-2013, according to Goodfellow et al. (2013), which is about 9% lower than RMN. Due to various factors such as location, perceived gender, and age affecting people's subjective judgment (Moosaei and Riek, 2013), using an ML-based expression classifier may provide comparable accuracies while being more efficient in terms of time and resources. People should therefore be included in the evaluation stage when the robot is sufficiently developed and can be applied in the context it was created for.

6.2 Outlook, Further Work, and Limitations

Utilizing accessible AI tools has proven valuable in generating and evaluating a functional robotic face prototype to be further tested with users in intended use cases. The objective insights presented were enabled by the speed and flexibility of generating AU data to inform the design, train a specific control model, and evaluate performance using the FER-2013 dataset. Based on the acquired results, we see potential for this method to assist further development steps. Firstly, from the iterative nature of early-stage development, insights and elicited requirements (from users) can be addressed by altering or generating new prototypes. This is accommodated by quickly and cheaply generating a control algorithm and obtaining objective evaluation not restricted by hardware capabilities or appearance traits. Secondly, expanding the portfolio of facial appearances is supported by the flexibility and valuable feedback provided in the AU data. This implies that each appearance can be tested and evaluated fast, and a model can be trained to utilize the actuation capabilities of the given robot face. These points highlight that rapid and objectively informed design iterations are possible, even when addressing complex and multifaceted problems, such as expressive humanoid robots. We showcase this by the results and insights obtained for the current iteration of our robotic face intended for medical simulations.

The presented results in this study indicate that the current robot cannot generate strong AU intensities. This finding is further scrutinized by relating the AU data to the hardware setup and suggesting how to improve the design by looking at the AU and servo correlations and the intensity of AUs in the generated dataset. The weak AU intensities and strong correlations could suggest that in addition to the range of motion, more human artifacts such as particular wrinkles, textures, and distinct landmarks such as marked eyebrows could amplify the obtainable AUs. However, the current performance of the robot is still sufficient to train a control model, as previously discussed, and the robot can mimic a human actor based on unfiltered AUs recorded using a fast and flexible pipeline. These results are promising as ours and other robotic faces can benefit from the presented support tools to inform and speed up the design process. Development tools are also essential as the complexity of the robotic setup increases, where both manual operation and hardware evaluation become more challenging and time-consuming. Thus, the data generated in the form of both descriptive intensities of AUs and correlations between actuation modes and generated output is purposeful when pursuing more complex behaviors and expressions rendered by robotic faces. In addition, our approach can address the non-linear relations between obtained AUs and actuation modes which is essential as increasing actuation points would increase interdependencies and complex behavior since the skin is a single deformable structure. Concerning more complex hardware setups, the generated data and correlations between input and output parameters could be a helpful tool in addressing actuators not correlating to detected action units or even interfering with other actuation modes limiting the expression output of the robot. Our approach should be applied on robot faces with varying degrees of freedom to further validate these potential advantages.

Since OpenFace 2.0 is essentially made for analyzing the facial behavior of humans (and now also of robots), it can additionally be used in the control system of the robot itself to analyze humans during interactions and adjust its non-verbal communication approach accordingly. The utilization of OpenFace 2.0 can also be scrutinized to advance the method further by incorporating facial landmark detection to measure facial deformations or tracking head pose and eye-gaze to analyze gaze behavior, which is essential for improving human-robot interactions (Abubshait and Wykowska, 2020). An approach for automatically finding the most neutral expression of the robot should also be explored to enhance the effect of person-specific normalization and potentially increase the range of AU intensities.

As a proof of concept, the expressive capabilities of the robotic face is tested by manually sending AU instructions to a trained control model. Static emotional expressions are difficult to evaluate, even for humans, so investigating dynamically changing expressions is interesting (Mollahosseini et al., 2019). Hence, the transition between expressions, actuation speed, and mechanical noise are essential parameters. Since AUs can be used as a transferal unit between humans and robots, facial responses from humans can be automatically captured to simplify the dynamic control of the robot. This approach could further be leveraged to generate models to automatically control the speed, onset, and offset of different expressions. Other examples for controlling the robot using AUs include real-time reenactment by operator, obtaining AU data sequence from video samples, or allowing users to generate custom expressions by either saving, mirroring, or manually adjusting AUs performed by the robot. The experiment using an actor for real-time reenactment of face movements suggests these possibilities to be promising. However, even though OpenFace 2.0 can analyze AUs in real time, our current robot design is limited by the actuator control unit moving the servos one at a time. By implementing a controller that can control several servos in parallel and with higher power output, this bottleneck can be reduced or potentially removed to increase response times and enable more nuanced motions during dynamic interaction.

While we in this paper discuss the applicability of using fast and accessible AI-based software to analyze robotic face movements, we acknowledge the importance of human user-interaction and evaluations in this setting (Moosaei and Riek, 2013). Human perception is particularly important for robots to be used in medical training scenarios, as the learning effects of having facial movement capabilities, and users' ability to respond to these, is not possible to deduce in any other way. This is also supported by users' limited ability to evaluate expressions rendered by alternative mediums or agents, suggesting that a physical robot is required to get further design inputs (Hofree et al., 2018). Concluding necessary design alterations, we want to pilot the robotic face in training scenarios and explore interactions and potentials for having expressive capabilities. Furthermore, as medical simulation is performed in teams, understanding the implications expressive cues of robots could pose on team dynamics is crucial. Therefore, utilizing the robotic face and control model to elicit requirements for expressions and facial responses that can enhance medical simulation scenarios is essential for future development.

The tests and insights showcased in this paper are obtained using only the face portrayed in the current prototype, and thus, limitations for using other appearances could be encountered. For example, how realism and fidelity concerning the visual appearance of the prototype influences AU data generated by OpenFace 2.0 and distort the confidence scores from RMN. This brings to question how closely the prototype needs to resemble a generalized human face. Furthermore, to explore visual edge-cases of the robotic face, appearance traits such as proportions, complexion, texturing, hair, tattoos, or scarring could be investigated. Further work would also be required to evaluate the effects of age, gender, and ethnicity, evaluate the robustness of the tools utilized, test for biases in the utilized data, and expand the suite of available appearances for the robotic face. It is also not evident how a generalized hardware setup accommodates the various facial characteristics as this would suggest anthropometric differences concerning size, proportions, and landmark location. However, we believe that a standardized setup would enable a sufficient design space to explore the potentials and limitations of switching the appearance of humanoid robots such as this one. As this is a crucial aspect of ensuring inclusivity and training variance in medical simulation, we see the approach utilized in this paper as effective for enabling faster prototyping iterations when developing humanoid robots.

7 CONCLUSION

We have presented methods utilizing open-source AI tools for supporting the development and evaluation of robotic faces. First, dynamic AUs of the robotic face were automatically captured through OpenFace 2.0 during random movements to find correspondence between facial attributes and the servo configuration. The correlations between AUs and servos provided objective feedback on the possibilities and limitations of the robot design. Next, a control model for the robot was developed by estimating the relationship between AUs and servo positions through regression analysis, enabling facial expressions to be rendered using AUs as input. We then evaluated the simulated expressions using a classifier trained on a large dataset of human facial expressions, providing additional assessment opportunities for the robotic design and control model. The methods have proven to be beneficial during early-stage development to rapidly gain actionable insights, in addition to being low-cost and easy to use.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

MA did the physical prototyping, hardware development, and experimentation using accessible machine learning tools to evaluate face expression capabilities. SK did the data generation, statistics, coding, and training of machine learning models to control the face. MJ and MS contributed to the conception and design of the study. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This research is supported by the Research Council of Norway through its industrial Ph.D Funding Scheme, Project number 290404.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frobt.2022.887645/full#supplementary-material

REFERENCES

Abubshait, A., and Wykowska, A. (2020). Repetitive Robot Behavior Impacts Perception of Intentionality and Gaze-Related Attentional Orienting. *Front. Robot. AI* 7, 150. doi:10.3389/frobt.2020.565825

Ameca (2021). Eng. Arts.. Available at: https://www.engineeredarts.co.uk/robot/ameca/ (Accessed December 13, 2021).

Auflem, M., Erichsen, J. F., and Steinert, M. (2019). Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators. *Procedia CIRP* 84, 572–578. doi:10.1016/j.procir.2019.04.202

Baltrušaitis, T., Mahmoud, M., and Robinson, P. (2015). "Cross-dataset Learning and Person-specific Normalisation for Automatic Action Unit Detection," in 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG) (Ljubljana: IEEE), 1–6. doi:10.1109/FG.2015.7284869

Baltrusaitis, T., Zadeh, A., Lim, Y. C., and Morency, L.-P. (2018). "OpenFace 2.0: Facial Behavior Analysis Toolkit," in 2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018) (Xi'an: IEEE), 59–66. doi:10.1109/FG.2018.00019

Bartlett, M. S., Movellan, J. R., Littlewort, G., Braathen, B., Frank, M. G., and Sejnowski, T. J. (2005). "Toward Automatic Recognition of Spontaneous Facial Actions," in *What Face Reveals Basic Appl. Stud. Spontaneous Expr. Using Facial Action Coding Syst. FACS* (Oxford: Oxford University Press), 393–413. doi:10.1093/acprof:oso/9780195179644.003.0019

Becker-Asano, C., and Ishiguro, H. (2011). "Evaluating Facial Displays of Emotion for the Android Robot Geminoid F," in 2011 IEEE Workshop on Affective Computational Intelligence (WACI) (Paris: IEEE), 1–8. doi:10.1109/WACI.2011.5953147

Breazeal, C. (2009). Role of Expressive Behaviour for Robots that Learn from People. *Phil. Trans. R. Soc. B* 364, 3527–3538. doi:10.1098/rstb.2009.0157

Cameron, D., Millings, A., Fernando, S., Collins, E. C., Moore, R., Sharkey, A., et al. (2018). The Effects of Robot Facial Emotional Expressions and Gender on Child-Robot Interaction in a Field Study. *Connect. Sci.* 30, 343–361. doi:10.1080/09540091.2018.1454889

Tian, L., Carreno-Medrano, P., Allen, A., Sumartojo, S., Mintrom, M., Coronado Zuniga, E., et al. (2021). "Redesigning Human-Robot Interaction in Response to Robot Failures: A Participatory Design Methodology," in Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems CHI EA '21 (New York, NY: Association for Computing Machinery), 1–8. doi:10.1145/3411763.3443440

Chen, B., Hu, Y., Li, L., Cummings, S., and Lipson, H. (2021). "Smile Like You Mean It: Driving Animatronic Robotic Face With Learned Models," in 2021 IEEE International Conference on Robotics and Automation (ICRA), 2739–2746. doi:10.1109/ICRA48506.2021.9560797

Cominelli, L., Hoegen, G., and De Rossi, D. (2021). Abel: Integrating Humanoid Body, Emotions, and Time Perception to Investigate Social Interaction and Human Cognition. *Appl. Sci.* 11, 1070. doi:10.3390/app11031070

Conigliaro, R. L., Peterson, K. D., and Stratton, T. D. (2020). Lack of Diversity in Simulation Technology: An Educational Limitation? *Simul. Healthc.* 15, 112–114. doi:10.1097/SIH.0000000000000405

Cooper, J. B., and Taqueti, V. R. (2008). A Brief History of the Development of Mannequin Simulators for Clinical Education and Training. *Postgrad. Med. J.* 84, 563–570. doi:10.1136/qshc.2004.009886

Ege, D. N., Lilleløkken, O., Auflem, M., and Steinert, M. (2020). Dealing with Ecological Validity and User Needs when Developing Simulation Based Training Equipment - Case Study of a Medical Palpation Task Trainer. *Procedia CIRP* 91, 722–727. doi:10.1016/j.procir.2020.03.116

Ekman, P., and Friesen, W. V. (1978). *Facial Action Coding Systems.* Palo Alto, CA: Consulting Psychologists Press.

Faraj, Z., Selamet, M., Morales, C., Torres, P., Hossain, M., Chen, B., et al. (2021). Facially Expressive Humanoid Robotic Face. *HardwareX* 9, e00117. doi:10.1016/j.ohx.2020.e00117

Ghazali, A. S., Ham, J., Barakova, E. I., and Markopoulos, P. (2018). Effects of Robot Facial Characteristics and Gender in Persuasive Human-Robot Interaction. *Front. Robot. AI* 5, 73. doi:10.3389/frobt.2018.00073

Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., et al. (2013). "Challenges in Representation Learning: A Report on Three Machine Learning Contests," in *Neural Information Processing Lecture Notes in Computer Science*. Editors M. Lee, A. Hirose, Z.-G. Hou, and R. M. Kil (Berlin, Heidelberg: Springer), 117–124. doi:10.1007/978-3-642-42051-1_16

Hanson, D., Pioggia, G., Dinelli, S., Francesco, F. D., Francesconi, R., and Rossi, D. D. (2002). "Identity Emulation (IE): Bio-Inspired Facial Expression Interfaces for Emotive Robots," in AAAI Mob. Robot Compet. (Edmonton, AB: AAAI Press), 11.

Hashimoto, T., Hiramatsu, S., and Kobayashi, H. (2006). "Development of Face Robot for Emotional Communication between Human and Robot," in 2006 International Conference on Mechatronics and Automation (Luoyang: IEEE), 25–30. doi:10.1109/ICMA.2006.257429

Hofree, G., Ruvolo, P., Reinert, A., Bartlett, M. S., and Winkielman, P. (2018). Behind the Robot's Smiles and Frowns: In Social Context, People Do Not Mirror Android's Expressions but React to Their Informational Value. *Front. Neurorobot.* 12, 14. doi:10.3389/fnbot.2018.00014

Hortensius, R., Hekele, F., and Cross, E. S. (2018). The Perception of Emotion in Artificial Agents. *IEEE Trans. Cogn. Dev. Syst.* 10, 852–864. doi:10.1109/TCDS.2018.2826921

Houde, S., and Hill, C. (1997). "What Do Prototypes Prototype?," in *Handbook of Human-Computer Interaction* (Cupertino, CA: Elsevier), 367–381. doi:10.1016/B978-044481862-1.50082-0

Ishihara, H., Iwanaga, S., and Asada, M. (2021). Comparison between the Facial Flow Lines of Androids and Humans. *Front. Robot. AI* 8, 29. doi:10.3389/frobt.2021.540193

Jung, M., and Hinds, P. (2018). Robots in the Wild: A Time for More Robust Theories of Human-Robot Interaction. *J. Hum.-Robot Interact.* 7, 1–5. doi:10.1145/3208975

Kim, D., Jung, S., An, K., Lee, H., and Chung, M. (2006). "Development of a Facial Expression Imitation System," in 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, 3107–3112. doi:10.1109/IROS.2006.282329

Kwon, M., Jung, M. F., and Knepper, R. A. (2016). "Human Expectations of Social Robots," in 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 463–464. doi:10.1109/HRI.2016.7451807

Lalitharatne, T. D., Tan, Y., Leong, F., He, L., Van Zalk, N., De Lusignan, S., et al. (2020). Facial Expression Rendering in Medical Training Simulators: Current Status and Future Directions. *IEEE Access* 8, 215874–215891. doi:10.1109/ACCESS.2020.3041173

Lazzeri, N., Mazzei, D., Greco, A., Rotesi, A., Lanatà, A., and De Rossi, D. E. (2015). Can a Humanoid Face Be Expressive? A Psychophysiological Investigation. *Front. Bioeng. Biotechnol.* 3, 64. doi:10.3389/fbioe.2015.00064

Lin, C.-Y., Huang, C.-C., and Cheng, L.-C. (2011). "A Small Number Actuator Mechanism Design for Anthropomorphic Face Robot," in 2011 IEEE Int. Conf. Robot. Biomim. (Karon Beach: IEEE). doi:10.1109/ROBIO.2011.6181357

Magtanong, E., Yamaguchi, A., Takemura, K., Takamatsu, J., and Ogasawara, T. (2012). "Inverse Kinematics Solver for Android Faces with Elastic Skin," in *Latest Advances in Robot Kinematics.* Editors J. Lenarcic and M. Husty (Dordrecht: Springer Netherlands), 181–188. doi:10.1007/978-94-007-4620-6_23

Mazzei, D., Lazzeri, N., Hanson, D., and De Rossi, D. (2012). "HEFES: An Hybrid Engine for Facial Expressions Synthesis to Control Human-like Androids and Avatars," in 2012 4th IEEE RAS EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob), 195–200. doi:10.1109/BioRob.2012.6290687

Meghdari, A., Shouraki, S. B., Siamy, A., and Shariati, A. (2016). "The Real-Time Facial Imitation by a Social Humanoid Robot," in 2016 4th International Conference on Robotics and Mechatronics (ICROM) (Tehran: IEEE), 524–529. doi:10.1109/ICRoM.2016.7886797

Minato, T., Shimada, M., Ishiguro, H., and Itakura, S. (2004). "Development of an Android Robot for Studying Human-Robot Interaction," in *Innovations in Applied Artificial Intelligence.* Editors B. Orchard, C. Yang, and M. Ali (Berlin, Heidelberg: Springer), 424–434. Lecture Notes in Computer Science. doi:10.1007/978-3-540-24677-0_44

Mollahosseini, A., Hasani, B., and Mahoor, M. H. (2019). AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. *IEEE Trans. Affect. Comput.* 10, 18–31. doi:10.1109/TAFFC.2017.2740923

Moosaei, M., Das, S. K., Popa, D. O., and Riek, L. D. (2017). "Using Facially Expressive Robots to Calibrate Clinical Pain Perception," in Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction HRI '17 (New York, NY, USA: Association for Computing Machinery), 32–41. doi:10.1145/2909824.3020216

Moosaei, M., and Riek, L. D. (2013). "Evaluating Facial Expression Synthesis on Robots," in *Proceedings of the HRI Workshop onApplications for Emotional Robots* (Tokyo: ACM), 6.

Mori, M., MacDorman, K., and Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robot. Autom. Mag.* 19, 98–100. doi:10.1109/MRA.2012.2192811

Nygaard, T., Kriesi, C., Sjöman, H., and Steinert, M. (2018). "From the Eyes of the Patient: Real Time Gaze Control of Medical Training Mannequins," in Proceedings of the 10th Nordic Conference on Human-Computer Interaction NordiCHI '18 (New York, NY, USA: Association for Computing Machinery), 932–935. doi:10.1145/3240167.3240228

Oh, J.-H., Hanson, D., Kim, W.-S., Han, Y., Kim, J.-Y., and Park, I.-W. (2006). "Design of Android Type Humanoid Robot Albert HUBO," in 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (Beijing: IEEE), 1428–1433. doi:10.1109/IROS.2006.281935

Pan, M. K. X. J., Choi, S., Kennedy, J., McIntosh, K., Zamora, D. C., Niemeyer, G., et al. (2020). "Realistic and Interactive Robot Gaze," in 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (Las Vegas, NV: IEEE), 11072–11078. doi:10.1109/IROS45743.2020.9341297

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830. https://www.jmlr.org/papers/v12/. https://www.jmlr.org/papers/v12/pedregosa11a.bib

Pham, L., Vu, T. H., and Tran, T. A. (2021). "Facial Expression Recognition Using Residual Masking Network," in 2020 25th International Conference on Pattern Recognition (ICPR) (Milan: IEEE), 4513–4519. doi:10.1109/ICPR48806.2021.9411919

Pioggia, G., Ahluwalia, A., Carpi, F., Marchetti, A., Ferro, M., Rocchia, W., et al. (2004). FACE: Facial Automaton for Conveying Emotions. *Appl. Bionics Biomechanics* 1, 91–100. doi:10.1533/abib.2004.1.2.91

Pololu - Mini Maestro (2021). Mini Maestro 18-Channel USB Servo Controller (Assembled). Available at: https://www.pololu.com/product/1354 (Accessed December 13, 2021).

Pourebadi, M., and Riek, L. D. (2018). "Expressive Robotic Patient Simulators for Clinical Education," in ACMIEEE Int Conf Hum. Robot Interact. Robots Learn. Workshop (Chicago, IL: ACM), 5.

Rawal, N., Koert, D., Turan, C., Kersting, K., Peters, J., and Stock-Homburg, R. (2022). , 8, 730317. doi:10.3389/frobt.2021.730317ExGenNet: Learning to Generate Robotic Facial Expression Using Facial Expression Recognition*Front. Robot. AI*

Silva, V., Soares, F., and Esteves, J. S. (2016). "Mirroring Emotion System - On-Line Synthesizing Facial Expressions on a Robot Face," in 2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT) (Lisbon: IEEE), 213–218. doi:10.1109/ICUMT.2016.7765359

Sophia (2021). Hanson Robot. Available at: https://www.hansonrobotics.com/sophia/(Accessed December 13, 2021).

Todo, T. (2018). "SEER: Simulative Emotional Expression Robot," in ACM SIGGRAPH 2018 Emerging Technologies SIGGRAPH '18 (New York, NY, USA: Association for Computing Machinery), 1–2. doi:10.1145/3214907.3214921

Venture, G., and Kulić, D. (2019). Robot Expressive Motions: A Survey of Generation and Evaluation Methods. *J. Hum.-Robot Interact.* 8, 1–17. doi:10.1145/3344286

Wu, T., Butko, N. J., Ruvulo, P., Bartlett, M. S., and Movellan, J. R. (2009). "Learning to Make Facial Expressions," in IEEE 8th International Conference on Development and Learning (Shanghai: IEEE), 1–6. doi:10.1109/DEVLRN.2009.5175536

Zhuang, Z., Landsittel, D., Benson, S., Roberge, R., and Shaffer, R. (2010). Facial Anthropometric Differences Among Gender, Ethnicity, and Age Groups. *Ann. Occup. Hyg.* 54, 391–402. doi:10.1093/annhyg/meq007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

C4

Dealing with Ecological Validity and User Needs when Developing Simulation Based Training Equipment – Case Study of a Medical Palpation Task Trainer

Procedia CIRP 91, 722-727: CIRP Design 2020 Conference

Daniel Nygård Ege, Oscar Lilleløkken, Marius Auflem, Martin Steinert (2020),

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 91 (2020) 722–727

www.elsevier.com/locate/procedia

30th CIRP Design 2020 (CIRP Design 2020)

Dealing with Ecological Validity and User Needs when Developing Simulation Based Training Equipment – Case Study of a Medical Palpation Task Trainer

Daniel Nygård Ege[a]*, Oscar Lilleløkken[a], Marius Auflem[a], Martin Steinert[a]

[a] *NTNU Trondheim, Richard Birkerlandsvei 2B, Trondheim 7491, Norway*

* Corresponding author. Tel.: +47 465 03 255. *E-mail address:* danieneg@stud.ntnu.no

Abstract

Simulation-based training offers a safe and repeatable environment where users can increase their skill level. Development of products facilitating such training is, however, faced with high uncertainty and ambiguous requirements. This complexity is a result of the intended use-cases, required product functionality and required approximation of clinical realism. In order to explore such uncertain aspects of simulation-based training equipment, two interesting factors are considered; user needs and ecological validity. This paper presents a case project where a medical palpation task trainer has been developed. In this case, a medical diagnostics procedure has been enabled through specific product functionality. As the complexity associated with simulators makes it difficult to elicit and fixate requirements, this paper highlights the benefits of using prototypes during the early stages of development. By presenting prototypes and testing these with both experienced and novice users, developers could better understand complex use cases and product interactions. Through case examples it is shown how prototyping extensively in various domains and levels of abstraction could help elicit product requirements as well as enlighten unknown problems and corresponding opportunities. In order to create products facilitating effective training, training scenarios must be made sufficiently realistic, i.e. ensuring ecological validity. To accomplish this, a simulator concept must encompass realistic tasks and physical attributes. Further, the potential of the conceptual prototype is discussed where self-directed learning and learning algorithms are of interest. Future research will, therefore, be focused on improved medical training and consequently training devices.

Keywords: Prototyping; Prototype-Driven Development; User-Centerd Design; Ecological Validity

1. Introduction

Developing products to enable simulation and simulation-based training can be associated with a lot of ambiguity and uncertainty, given complex user scenarios and attributes. Products encompassing human computer interaction in the context of simulation, offer challenges such as recreating complex scenarios and realism in physical functions and features [1]. Developers may aim to recreate the real world for their users to be better prepared for real world scenarios, enabled by training in a low risk environment, such as in driving- and flight simulators. Hence, medical task trainers and the development of these are also associated with this product context. Such products may encompass vastly different success criterions, but the necessity of eliciting requirements is key for all of them in the early development phase. The complexity associated with simulators, and their interaction, makes it difficult to elicit and fixate requirements. In order to create products facilitating effective training, training scenarios must be made sufficiently realistic. To accomplish this, the simulator has to encompass realistic tasks and physical attributes[2]. In order to map out uncertain aspects regarding

10.1016/j.procir.2020.03.116

the development of such products, *user needs* and *ecological validity are* two interesting factors to consider.

User needs can be uncovered and evaluated through conversation and prototype testing with expert users[3,4]. By presenting a case project, this paper exemplifies how a user-centered and prototype-driven development framework is used to elicit requirements. With the goal of designing a medical task trainer to impact how palpation is trained in medicine[5], two known challenges in early stage product development are addressed. Firstly, the user-centered design challenge, of identifying and addressing a user's needs. This could be done by eliciting users pain points and creating a conceptual solution for this problem. Secondly it concerns the complexity and ambiguity facing design teams and exemplifies how prototyping answers design questions through providing tangible insights. These are made tangible through both designing and building prototypes, but also through the interactions between users and conceptual prototypes [6].

Kushniruk et al. (2013) argues that human behavior is dependent on their environment, and therefore ecological validity is required to ensure that tasks/training objectives are applicable to the real world. The parameters encompassed in ecological validity are tasks being performed, users involved, the training scenario and training environment. In the context of this paper, ecological validity is considered a prototype's ability to facilitate training of tasks, by presenting functionality and attributes appropriate for the user to perform them. These attributes can be physical, like the look and feel of a prototype, and intangible, e.g. feedback provided to the user. Ecological validity can be measured by how closely a simulated scenario or task compares to its real counterpart. Ensuring a sufficient degree of realism is necessary to maximize learning outcomes and avoid false learning.

1.1. Prototypes and Prototyping

In this paper, prototypes are considered tools to enable learning [4]. Learning as an activity, where insights are created and made tangible through the creation and interaction with prototypes. Prototypes are, by this, merely the physical outcome from prototyping activities, but for the designer they serve as physical embodiments of concept potentials. Prototyping is considered a critical activity in the early phase of product development, as it is a powerful tool to facilitate learning, communicate and inform decisions [7,8]. In this phase, the goal of designers is to address the uncertain aspects, unknown unknowns, identify product potentials, and eliciting functional requirements for a given concept [9,10].

However, determining critical functions and consecutively the functional requirements is not evident in early stage product development. Aligning with users' needs and adapting to feedback could be challenging as a result of disciplinary and/or organizational boundaries. Prototypes are in this context important tools in crossing these boundaries and facilitate communication [11]. Here, the designers should be reflective in their prototyping activities and bear in mind all the features of any given artifact acting as the prototype [12]. Prototypes should be realized with focused purposes in order to test core assumptions and test out specific aspects of eventual designs [13,14].

1.2. Aim and Scope

The scope of this paper is to exemplify how prototypes can aid development of products encompassing complex attributes and user scenarios. This is done by considering the ecological validity of conceptual prototypes, and eliciting requirements in a user-centered and prototype-driven framework. The paper presents a case project where a medical palpation task trainer has been developed and generalizes the learning outcome of the project. Furthermore, the paper aims to show how product potentials and thereby functional requirements could be identified through user interaction and prototyping. Further it will be shown how users' tacit knowledge, could be made explicit and tangible to designers, when evaluating concepts with expert users.

2. Development of a Conceptual Task Trainer

In the case project presented below the aim has been to explore solutions for an abdominal examination task trainer. Medical task trainers are equipment enabling users to practice technical skills in low risk, repeatable environments [15], and could enable training of novel and/or challenging medical tasks and recognizing conditions and symptoms too rare or dangerous to perform on real patients. As with other projects aiming to recreate real world procedures, ecological validity and user needs are essential parameters to consider when developing a conceptual prototype.

A product potential for such a task trainer was revealed, as students in training practiced on each other, without any additional training aids available. Practicing on a healthy stomach makes it hard to prepare for a real examination, as several symptoms are not possible to simulate. There are also great variations between tests. This became the initial starting point for the project proposed by a medical training device company. Without fixed predeterminations, the development team utilized a user-centered design approach to evaluate a product potential as well as generate a conceptual prototype.

2.1. Exploration of the Use Case

When working with complex problems, such as when designing simulator-based training equipment, defining and evaluating functionalities can be challenging. By providing feedback on prototypes, expert users can aid developers in mapping out needed functionalities, as well as identifying requirements for such functionalities. In the context of this project, expert users are trained medical personnel. Even if trained professionals are not the primary users, their tacit knowledge and experience is vital, and their interaction with prototypes are a key part of iterative prototyping. Contact with expert users is especially valuable in the medical field, given a large chasm between professions, knowledge bases and expertise in the field [16].

Early consultation with expert users gave insight into how an abdominal examination is conducted. Diagnosis of the abdomen requires different symptoms to be identified that may be divided into three main categories: the tactile feeling of the stomach, location of pain, and bowel sounds. When exploring ways to replicate the users sensory experience and interaction with the product, prototypes replicating its physical attributes were used.

The diseases commonly identified by palpation, and which would be most relevant for the abdominal task trainer to simulate were described by expert users. Most diseases are identified by the location of abdominal pain, and the detection of enlarged organs. As shown in Table 1, six diseases and corresponding symptoms [5] were selected for simulation.

Table 1. Most common abdominal diseases with typical symptoms

Disease	Typical symptoms
Appendicitis	Pain in lower right quadrant of stomach
Diverticulitis	Pain in lower left quadrant of stomach
Peritonitis	Pain all over stomach, rebound tenderness
Inflamed gallbladder	Pain in upper right quadrant, gallbladder detectable by touch under ribcage
Inflamed liver	Liver detectable by touch below ribcage
Pancreatitis	Upper abdominal pain

Required functionalities was mapped out based on the diseases chosen for simulation, as shown in Table 2. The medical symptoms were translated into technical functionalities in a product-oriented context; a realistic tactile experience and visual attributes of the abdomen, being able to localize pain, and having the ability to simulate inflamed organs. In the following sections, these functionalities are explored by prototyping and described in further detail.

Table 2. Sub- functionalities prioritized for each function by expert users

Functionality	Sub-functionality
Accurate tactile experience	Firmness of stomach
Pain response	Visible indication of pain
	Indication of pain intensity
Physical resemblance	Has the shape of a stomach
	Includes hipbones and ribcage
	Other parts of appearance must not compromise immersion
Inflamed organ simulation	Possibility to simulate inflamed organs
	Turn on/off inflamed organs, according to different training scenarios
	Realistic tactile experience
Selection of illness	Illness selectable by user

2.2. Prototyping Physical Attributes

Quantifying attributes such as tactile feeling and visual resemblance is challenging and requires expert users opinions to be recreated in a task trainer. Prototypes with different attributes were created so users could test and provide feedback.

2.2.1. Tactile Experience

Aiming to rapidly develop a testable prototype which feels like a stomach, a large plastic container was filled with foam and water filled condoms, as depicted in Fig. 1. A PVC sheet was stretched over to resemble skin, and resulted in a soft, deformable stomach, replicating interaction with the future task trainer. An expert user tested the prototype and raised a concern regarding the uneven feel of it.

Fig. 1. Expert user testing a simple prototype.

Several other materials were proposed to resemble the stomach, some shown in Fig 2., and was internally tested. When a viable possibility was uncovered, a new prototype was tested with an expert user to verify it as a viable solution.

a b

Fig. 2. (a) Ballistic gel with different mix ratios was used to mimic the feeling of a stomach; (b) A multi-layered silicone model offered robustness as well as being soft to the touch.

A shortcoming in using fellow students as markers when training abdominal examinations, is the inability to simulate change in tactility of the abdomen. This might happen as a result of inflamed organs, that could indicate a serious disease. Some organs swell and become detectable by touch when inflamed. An interview with an expert user revealed that the feel of some inflamed organs was only provided through literature. Simulating this in a task trainer enables practice in identifying such symptoms, which is not possible when practicing palpation on a healthy stomach.

An example of an organ detectable by touch when inflamed is the gallbladder, described to feel like a soft lump about the size of a plum when inflamed. With a description as vague as this, the need for prototyping and iterating was apparent to simulate it. Internal tests were done to match the description and validated by expert users.

Fig. 3. Simulating an inflamed gallbladder was done by inflating a balloon with water.

2.2.2. Visual Resemblance

In the field of simulation-based training equipment, it is believed that increasing fidelity and visual resemblance does not necessarily increase learning outcome [17]. It can be argued that expert users with a lot of experience do not require a high-fidelity task trainer to practice. However, when teaching students, being novice users, an unnatural looking simulator could ruin immersion and lead to less learning outcome.

Steps were taken to make an immersive simulation. The prototype depicted in Fig. 1 was implemented in a CPR training mannequin in order to enhance realism. Adding a ribcage and head to the container resulted in a more realistic task trainer, as shown in Fig. 4.

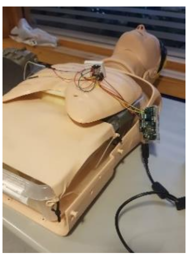


Fig. 4. Integrating a prototype in a CPR mannequin increased realism and immersion.

The iterative process of making a more visually realistic prototypes was done in order to increase immersion of the task trainer. This allowed for not only testing of physical attributes, but facilitates a more realistic, ecologically valid training scenario.

The final prototype was integrated in a medical simulation mannequin, as shown in Fig. 5. Besides demonstrating future integration opportunities in a mannequin, an increased natural look was achieved by having realistic proportions to a human torso, as well as encompassing anatomical landmarks.

Fig. 5. Prototype integrated in medical simulation mannequin.

2.3. Prototyping Functionality

Prototypes are useful in eliciting expert user's tacit knowledge, as described earlier. This knowledge is important when designing prototypes that look and feel realistic and provide an opening for discussion. However, prototype feedback does not ascertain in what degree a prototype is ecologically valid. To establish this, designers must consider how a task can be broken down into tangible objectives, that could be recreated and enabled by a physical model. Hence, necessary input and output parameters must be considered. For instance, when designing a flight simulator, input parameters such as flight controls and switches must be present, while output such as visuals must be recreated in a sufficient way, in order to be ecologically valid.

By breaking down a medical procedure into input and output parameters, distinct prototype functionalities can be established. In the context of this paper, where the procedure at hand is abdominal examination, inputs and outputs are determined by disease symptoms. The procedure is broken down, or translated, into functionality and corresponding requirements for the task trainer. The main input parameters were found to be pain localization and intensity. Pain feedback was determined to be the main output parameter.

2.3.1. Pain Feedback

According to expert users, locating pain and understanding the severity is a necessity when diagnosing an abdominal illness. Detecting touch in a soft material was considered an unknown, and prototyping was fundamental for discovering solutions. A way to display pain resembling what happens in real life was selected through dialogue with users.

Implementing a simple sensor system controlled by a microcontroller, allowed the expert user to detect the location of a force sensor under a PVC skin by palpation, as depicted in Fig. 6. Pain intensity was represented by numbers between 1 and 10 and printed on a separate screen next to the prototype in real-time. A limited number of diseases are possible to simulate by using one sensor, but the concept of using force sensors for both locating pain and detecting pain intensity in a soft material was verified.

Fig. 6. A force sensor offered pain and location feedback in the soft stomach.

Adding sensors to the stomach increased the number of simulatable diseases by delivering sufficient position and force measurements. Having the possibility of differentiating between palpation in each of the abdominal quadrants, the expert user was able to diagnose the task trainer for diseases such as diverticulitis and appendicitis. An LCD screen was implemented to display pain intensity, and printed messages such as “I’m feeling fine” and “this hurts”, to simulate a patient’s reaction to pain while being examined.

The expert user requested an easier visible display distinguishing between levels of pain. By mimicking the numeric pain scale often used in hospitals, a colored LED strip was introduced, replacing the screen. This increase in realism put the simulation closer to what users experience in real life, thus facilitating ecological validity.

2.3.2. Geometrical aspects

The palpatable part of the stomach was designed to fit into a human simulator mannequin. By basing the design on human anatomy, ecological validity could be ensured thru having palpatable landmarks. Landmarks such as hip bones and rib cages were added through iterations, in order to capture all required anatomical attributes. The silicone abdomen integrated in the final conceptual prototype is shown in Fig. 7., with outer dimensions at 345mm in height and 290mm in width. The figure also shows the placement of sensors.

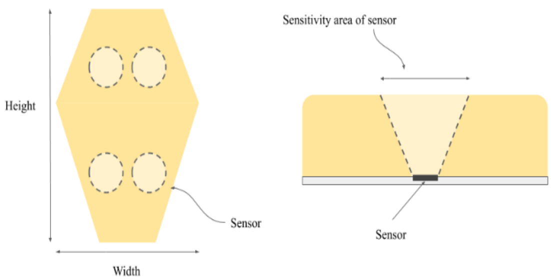


Fig. 7. Illustration of abdomen with sensor placement.

Force applied to the task trainer propagates through silicone, and sensors register it in a greater area than the area of the sensor itself. This effect is also illustrated in Fig. 7. The sensitivity area of the sensors was measured to be approximately 78.5cm^2, while the area of the sensors is 16cm^2. The sensors were placed below 4.5cm of silicone, and approximately 15cm apart.

2.4. Testing the Ecological Validity of the Conceptual Prototype

The final iteration of the task trainer prototype is an independent system, encompassing the critical functions identified through interactions with expert users. Physically resembling a real patient, and with the ability to provide the user with pain feedback the prototype can simulate relevant diseases. To increase the learning outcome, data from practice sessions can be collected to provide each user with feedback on their performance. A conceptual prototype was tested by students and expert users, as depicted in Fig. 8 and Fig. 9. Both were able to diagnose the task trainer for diverticulitis by palpation, and feedback on the experiences were positive. Students were engaged in the test scenario and expert users confirmed the realism of the case. The company associated with the project was also positive to the physical results of the project, and pleased about the positive feedback from users.

As described earlier, ecological validity can be measured based on a prototype’s ability to facilitate training of tasks, by presenting functionality and attributes appropriate for the user to perform them. For the conceptual prototype developed in this project, ecological validity was tested by medical professionals and students. An expert user confirmed the realism of the task trainer and that it had the necessary functionality to facilitate learning. Students examining the task trainer managed to determine the correct diagnosis, which could indicate that appropriate attributes are present.

Fig. 8. An expert user diagnoses the task trainer to confirm realism

Fig. 9. A student practices palpation of the abdomen

3. Discussion and Concluding Remarks

Through a specific case project, this paper gives an exemplification of how user-centered development could leverage prototypes. The case concerns the development of simulation-based training equipment, and more specifically a medical palpation task trainer concept. By medical products mimicking the human physiology, trainees can attempt procedures in both a safe and repeatable training environment. How extensive and to what degree of realism any physical simulator needs to approximate in order to successfully simulate a given task is, however, not evident.

In order to address this challenge, the described project utilized prototyping probes, to elicit functional requirements, as well as to pilot conceptual ideas as training equipment with users. Through these efforts it became apparent that the ecological validity of any evaluated concept needed to satisfy users (and expert users) requirements. This meant, ensuring both task abilities, simulating relevant and needed conditions, tactile and visual resemblance, as well as performance feedback communicated to the trainees.

Insights gathered shows the importance of intentional and reflective prototyping and prototype testing [12]. Specific tests should be performed to explore specific dimensions of the conceptual idea. In the project case, the two levels of abstraction, mainly, prototype attributes and a concept's role in the training curricula needed to be acknowledged. As some prototypes explored specific dimensions of a simulator concept, these were less suited to test and evaluate the intended role of the task trainer. However, both these levels needed to be mapped out in order to ensure an ecological valid concept proposal that had been verified trough several tests with both novice and experienced users. This demonstrates the importance of prototyping continuously, intentionally, and together with users in order to address the challenges concerning early stages of design.

The developed concept is showing potential in training for tactile recognition as well as the motoric skill training. More so the concept is shown to be a good conceptual model to investigate self-directed training concept. We are intending to investigate how pattern-recognition in the data from the simple sensor setup of the prototype could be utilized in providing feedback and debrief information for trainees. Even when only considering the already integrated sensors, position, area coverage, palpation forces, force peaks and timing could inform algorithms to enable not only simple pain point location but also following complete diagnosis scenarios. This is the underlying potential for this prototype system, and for future development.

The main contribution of this paper has been to exemplify how prototypes can aid development of products encompassing complex attributes and user scenarios. This was exemplified by considering the ecological validity of conceptual prototypes, and eliciting requirements in a user-centered and prototype-driven framework with a case project. Determining realistic physical attributes was enabled by discussion with healthcare professionals, through extensive use of prototypes. Furthermore, the paper has showed how user needs and thereby functional requirements could be identified through user interaction and prototyping. The presented case project explains how requirement elicitations was aided by prototypes to make a task trainer with a realistic tactile experience and visual resemblance. Especially in the case presented this was important, as the complexity associated with simulators, and their interaction, makes it difficult to elicit and fixate requirements. Designers investigated not only prototype attributes through user interactions, but also their usability, i.e. the prototype as a concept of training routines and testing of these. By not only investigating user needs, but by being reflective over both physical attributes and the procedure at hand, an ecologically valid task trainer could be developed.

Acknowledgements

This research is supported by the Research Council of Norway through its industrial PhD funding scheme, project number 290404.

References

[1] Rex Hartson H. Human–computer interaction: Interdisciplinary roots and trends. J Syst Softw 1998;43:103–18.
[2] Kushniruk A, Nohr C, Jensen S, Borycki EM. From Usability Testing to Clinical Simulations: Bringing Context into the Design and Evaluation of Usable and Safe Health Information Technologies. Yearb Med Inform 2013;22:78–85.
[3] Sutcliffe A, Sawyer P. Requirements elicitation: Towards the unknown unknowns, 2013, p. 92–104.
[4] Ulrich KT, Eppinger SD. Product design and development. New York: McGraw-Hill/Irwin; 2012.
[5] Jarvis C. Pocket Companion for Physical Examination and Health Assessment. 7th Edition. Missouri: Saunders; 2015.
[6] Leifer L, Steinert M. Dancing with Ambiguity: Causality Behavior, Design Thinking, and Triple-Loop-Learning.Inf-Knowl-SystManag 2011;10:151–73.
[7] Lauff C, Kotys-Schwartz D, Rentschler M. What is a Prototype? What are the Roles of Prototypes in Companies? J Mech Des 2018;140.
[8] Jensen LS, Özkil AG, Mortensen NH. Prototypes In Engineering Design: Definitions and Strategies. 84 Proc Des 2016 14th Int Des Conf 2016:821–30.
[9] Kriesi C, Blindheim J, Bjelland Ø, Steinert M. Creating Dynamic Requirements through Iteratively Prototyping Critical Functionalities. Procedia CIRP 2016;50:790–5.
[10] Auflem M, Erichsen JF, Steinert M. Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators. 572-578 2019.
[11] Buchenau M, Suri JF. Experience prototyping. Proc. Conf. Des. Interact. Syst. Process. Pract. Methods Tech. - 00, New York City, New York, United States: ACM Press; 2000, p. 424–33.
[12] Houde S, Hill C. What do Prototypes Prototype? In: Helander MG, Landauer TK, Prabhu PV, editors. Handb. Hum.-Comput. Interact. Second Ed., Amsterdam: North-Holland; 1997, p. 367–81.
[13] Menold JD. Prototype for X (PFX): A Prototyping Framework to Support Product Design 2017.
[14] Lim Y-K, Stolterman E, Tenenberg J. The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. ACM Trans Comput-Hum Interact TOCHI 2008;15:7:1–7:27.
[15] Maran NJ, Glavin RJ. Low- to high-fidelity simulation - a continuum of medical education? Med Educ 2003;37 Suppl 1:22–8.
[16] Zenios S, Denend L, Brinton TJ, Kumar UN. Biodesign: The Process of Innovating Medical Technologies. Cambridge University Press; 2010.
[17] Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA. Reconsidering Fidelity in Simulation-Based Training. Acad Med 2014;89:387.

C5

Proposing a Novel Approach for Testing Complex Medical Task Trainer Prototypes

DS 101: Proceedings of NordDesign

Oscar Lilleløkken, Daniel Nygård Ege, Marius Auflem, Martin Steinert (2020),

NordDesign 2020
August 11-14, 2020
Kgs. Lyngby, Denmark

Proposing a novel approach for testing complex medical task trainer prototypes

Oscar Lilleløkken[1], Daniel Nygård Ege[1], Marius Auflem[1], Martin Steinert[1]

[1]Norwegian University of Science and Technology

oscarli@stud.ntnu.no

1 Introduction

In the early stages of engineering design, efforts are put into mitigating risk and preventing costly rework later in the design process. This early stage is often referred to as the Fuzzy Front End and is described by Kim and Wilemon (2002), as the period from when a product opportunity is first made apparent until it is deemed ready for further development. As this stage is early, unmatured, and highly ambiguous, risk mitigation concerns exploring multiple designs, by not being bound to one solution that might not work out (Sobek Ii et al., 1999). Prototypes are in this context important tools, not only to evaluate design ideas, but to map out and explore a solution space (Auflem et al., 2019; Leifer & Steinert, 2011). Limiting time and resources invested, multiple prototypes can be created while still informing designers about which concepts perform well, and which concepts are prone to fail (Leifer & Steinert, 2011). Such prototype-driven development methods show the importance of both divergent and convergent thinking (Eris, O., 2003). Divergent thinking enables designers to diverge from the concrete facts of the problem and generate multiple solutions that could help solve it. In the divergent space generative design questions are used to help idea generation. In the convergent space the aim is to converge on a fact-based foundation from which a solution can be identified.

While various stages of product design all face uncertain elements, this paper aims to address the lack of knowledge and insights needed to inform further development of concepts generated in the Fuzzy Front End. More specifically, it concerns user-centred design challenges within the field of medical training equipment, how they may be tackled using structured prototyping methods and how proposed solutions may be validated through user interactions.

1.1 Wayfaring

Wayfaring is a design framework encouraging freedom to explore and enlighten both problem and solution space in parallel as described by Steinert & Leifer (2012). Wayfaring suggests that iterative cycles of designing, building and testing of prototypes should be used to address uncertainty of facing new product opportunities. In this context, prototypes are used as learning probes to uncover and explore a design space and test concepts (Gerstenberg et al., 2015). Each probe investigates one or more ideas that in turn inform and steer further development. This approach allows flexibility, as well as the opportunity of finding a novel idea and/or uncovering a unique product potential in the early stages of design.

Two important concepts in this phase of development are known- and unknown- unknowns. While the known unknowns can be revealed through information-gathering such as consulting with experts, the unknown unknowns are hidden and cannot be revealed through inductive studies (Ramasesh & Browning, 2014). The unknown unknowns could, however, be uncovered by extensive prototyping as diverging within a broad solution space one is prone to encounter new challenges not yet on the radar. These unknown unknowns are important to uncover as they can have major unforeseen implications for a project. By Wayfaring, an understanding of product requirements may develop in a dynamic and emergent manner, and unknown unknowns be revealed as the project progresses (Kriesi et al., 2016).

1.2 Mixed method research approach

Mixed method research combines both quantitative and qualitative data, with the assumption that combining methodologies with different strengths and limitations can uncover more information from each test subject (Abowitz Deborah A. & Toole T. Michael, 2010). The reliability of the results is also increased, as it enables triangulation of the obtained data. Abowitz and Toole (2010) defines triangulation, in the context of research, as "using multiple research methods or measures to test the same hypothesis or finding", exemplified by using both open-ended and closed-ended questions in a survey to measure the attributes in question. Dybvik (2018) describes the difference between quantitative and qualitative research;

- Qualitative research collects data from individual test subjects, and the subject's personal experience interacting with the experiment or concerning the research issue at hand.

- Quantitative research collects empirical, numerical data from experiments or test subjects. An important difference is that data is collected, not by the researchers themselves, but through a research tool, e.g. questionnaires and sensors.

As the strengths in quantitative and qualitative research compensate for the weaknesses of the other, this combination is valuable (Abowitz and Toole, 2010). Though mixed method research can provide more data with higher reliability, Dybvik (2018) calls attention to the possible drawbacks of this approach compared to a single method. Most important is the added cost and time required to carry out the research, as well as the need for researchers with knowledge of different research methods.

1.3 Aim and scope

The scope of this paper is to explore and structure a way of testing conceptual prototypes of medical task trainers by using a mixed method approach, in order to guide further development. By presenting a case project where a medical task trainer has been developed, the use of this novel approach is exemplified. The aim of this paper is to discuss how to test conceptual prototypes to assess multiple functionalities and features in order to ensure that the required ecological validity of a medical task trainer is met. The paper will also show how the same approach can help developers make informed decisions for further design iterations when dealing with complex user-interactions, use-cases or other independent uncertainties. Furthermore, the implications for a novel approach to testing leveraging a mixed method research is discussed.

2 A novel approach to guide development by testing prototypes using a mixed method approach

User-testing and testing users' interaction with prototypes is complicated, as limiting sources of error and elements affecting users' feedback is challenging. During prototype testing, it is often necessary to interview and ask specific questions to obtain useful data for further development. In this context attention should be brought to the features that are being tested and whether the designers are collecting the necessary information. For example, will asking broad, open-ended, questions lead to new questions, rather than answer the ones crucial for further development? Houde & Hill (1997) emphasize the intent for each prototype, and that designers should be aware and reflective of what each prototype is testing. This paper argues that the same mindset should be applied when designing a *test* of a prototype - it should be tailored to the prototype at hand, and open design-questions that need answering. The novel approach presented in this paper elicits user feedback from testing by combining a Likert-scale questionnaire, sensor data and a post-test interview.

For many projects, a sufficient test of a conceptual prototype could be achieved by simplifying the prototype to only test individual features or functionalities at one time. Designers also have the option of simplifying and focusing prototype tests by using an approximated environment. As described by Vestad & Steinert (2019) prototyping the test environments themselves, and conducting proxy tests, can be useful to gain meaningful insights and understand how the prototype performs. However, in many cases considerable simplifications are not viable. Oversimplification may degrade the realism of a trial to a level where the observations made are no longer transferable to the actual situation the designer is interested in. This is especially relevant when prototyping products involved in complex real-world interactions and scenarios. The development of medical training equipment is one example. These products often require the simulation of human anatomy, and the interaction of the user with that simulation. In order to assess the perceived realism of anatomical features, a certain degree of fidelity is required for the test to be sufficient. If elements which affect the prototypes' physical attributes are neglected the entirety of the prototypes' realism is reduced. If the prototype cannot be made simpler, the test must be made more comprehensive.

In a novel approach to testing, high numbers of unconnected and uncertain elements are addressed by combining subjective and objective observations. Combining quantitative and qualitative data collected from the user is suggested to maximize usable information collected from each test subject. As the strengths in quantitative and qualitative research compensate for the weaknesses of the other, combining them can give a better chance of obtaining useful information (Abowitz & Toole, 2010). Quantitative data is collected from sensors in the prototype, and subjective opinions are collected and quantified through a post-test questionnaire. The questionnaire includes a Likert scale (Allen & Seaman, 2007) evaluation of prototype features. A Likert scale is a rating scale where the test subject rate in what degree it agrees with a statement regarding the product. Thus, whether the individual functionalities uncovered during development are present in the final conceptual prototype, can be evaluated. By using the Likert scale questionnaire, the subjective opinions of all test participants are also combined and presented as quantifiable data. The conceptual prototypes should also include sensors, so that objective data from the interaction between prototype and user can be collected. Qualitative data can also be collected as opinions from the users, either in written form or through interviews, or as observations made by the developers during testing. Thus, in-depth explanation of why certain features was ranked more poorly than others can be collected. An overview of the new approach is shown in Fig. 1.

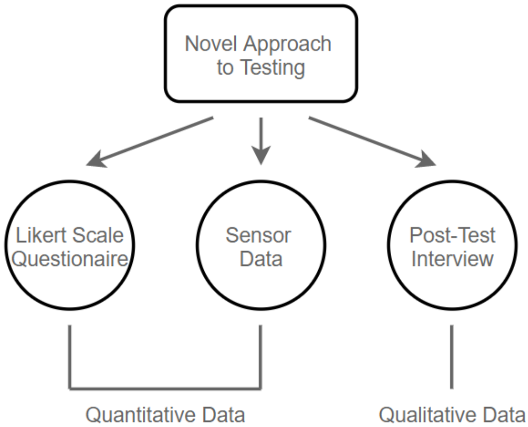


Fig.1. Illustration of a proposed novel approach for testing complex prototypes.

3 Case project- developing an abdominal examination task trainer

The case project concerns a product development project carried out in 2019 in Trondheim, Norway, by the authors to develop a tool to assist in abdominal palpation training. This project was initiated at TrollLABS, a product development laboratory at MTP, NTNU, in collaboration with a major European manufacturer of medical training equipment, as limited aids were available in medical education for the training using abdominal palpation to diagnose patients. This situation required users to train on human markers. While human markers enable high realism and close to real- world patient interactions; simulating diseases, pains, and altered tactile response is not possible. This project therefore wanted to explore the potentials of creating a "simulated patient", i.e., an abdominal task trainer allowing users to train their abdominal palpation diagnostic skills. One prototype task trainer developed over the course of this project is depicted in Fig. 2.

Fig 2. Conceptual prototype of an abdominal palpation task trainer.

This prototype had the overarching goal of fulfilling training goals from medical curricula, and obtain ecological validity, i.e. the required level of realism for users to benefit. Hence, multiple functions and design questions were attempted to be answered in parallel and represented in one conceptual model. To facilitate freedom to explore design space and work iteratively, the Wayfaring approach was chosen by the designers for this project. The workflow is presented in Fig. 3. The first step was to uncover what functionality and physical attributes an abdominal examination task trainer requires to meet the curriculum training goals. Through interviews and user interactions with prototypes, requirements were uncovered, and included both visual and tactile realism as well as ways of determining where pain occurs and its intensity. Physical and visual realism are both highly subjective experiences and a threshold for when something becomes "realistic" is very hard to pin down. Various ways of addressing these ambiguous qualities were explored in parallel in order to examine as many solutions as possible. Prototypes were used to address the uncertainties in how to achieve requirements. Iterations of different solutions was done in parallel, until the best solution emerged. Further iteration was completed before all solutions were collected and combined in a conceptual prototype. The development process of this conceptual prototype is described in further detail by Ege et al. (in press).

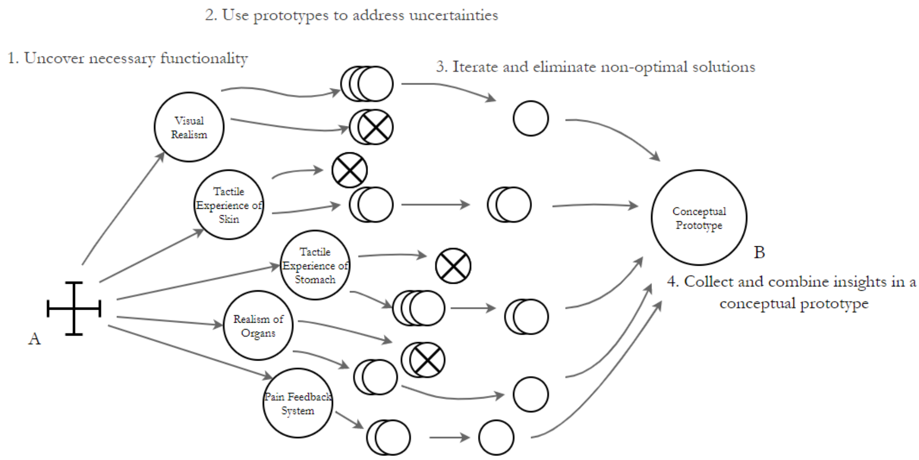


Fig 3 illustrates how parallel paths of prototypes with single functions are joined in a conceptual prototype.

The conceptual prototype developed consists of a soft stomach with an embedded sensor system. LEDs display pain intensity, while sensors embedded in the stomach measure applied pressure. The stomach is modelled to both look and feel anatomically correct and is integrated into a commercially available medical mannequin simulator. The prototype can simulate symptoms for 10 common abdominal diseases, including appendicitis and diverticulitis, as well as simulating inflamed organs.

The present case shows that in practice, the Wayfaring model which suggests development is like a winding path, can in fact be a collection of parallel paths taken in different directions of interest to answer design questions. Fig. 3 shows how parallel prototyping paths with single functions are joined in a conceptual prototype, combining insights and functionality. Throughout the project, prototypes were used to answer questions (Schrage, 2004). As insights and concepts are collected in one final prototype, the question of how well it incorporates all insights arise. Many of the features explored in the project are ambiguous, such as how something feels. Designers therefore relied on the opinions of expert users (in this context trained medical personnel), to prototype and recreate such attributes. Furthermore, a challenge that emerged throughout development was how to join the parallel prototyping paths. After a while, several loose ends had to be tied together, and compromises had to be made. E.g. in the presented project, incorporating the simulation of inflamed organs led to integration challenges because of a lack of space. This meant an optimal solution might have to be altered.

For the task trainer to accomplish its goal, i.e. facilitating the user to reach training goals, the right ecological validity must be present in the task trainer. Ecological validity is a measure of whether a test environment matches the real-world environment where the actual procedure is being performed (Kushniruk et al., 2013). It is important to note that only ecological validity in features important for creating effective training is in focus. Maximum ecological validity of all aspect of the task trainer is not expedient. Body heat, for example, was attempted included in the task trainer, but eventually abandoned as it added little to the learning outcome, while increasing complexity significantly. Each of the functionalities explored separately in development are meant to contribute to the ecological validity of the task trainer. To what degree the final conceptual prototype encompasses all the attributes is difficult to establish, as it is based on the subjective input of few expert users. It is also difficult to predict how the features will affect each other. The test scenario is therefore complex and difficult to structure. Adding the subjective opinions of testers increases the difficulty of designing valid tests of such products.

This was the backdrop designers faced when aiming to test the developed conceptual prototype to establish its applicability to the intended educational setting, and if acceptable ecological validity was achieved. It was apparent that the designers needed a structured way for testing conceptual prototypes containing multiple unconnected unknowns. The approach presented in the next section can be generalized for how to test complex medical task trainer prototypes. Further, it can aid developers in answering questions to finish the Fuzzy Front End, by structuring both qualitative and quantitative data.

3.1 Novel approach applied for the case project

An approach to testing tailored to the conceptual prototype, as discussed previously, is presented below, and shows how the test approach can be used in practice. It simplifies test answers and helps structure results to address independent unknowns and inform further development phases. The novel approach can be applied to the abdominal examination trainer to verify that all requirements were adequately present in the conceptual prototype. The Likert scale questionnaire addressing functionality and physical attributes individually are shown in Fig. 4. Educational applicability is also covered in the questionnaire. A post-test interview allows participants to provide in-depth answers to support the questionnaire. This way follow-up questions could be asked regarding factors rated low in the questionnaire which could be valuable for development.

Using a scale from 1 (strongly disagree) to 5 (strongly agree), in what degree do you agree with the following statements concerning the palpation simulator

1 - Strongly disagree, **2** - Disagree, **3** - Neither agree nor disagree, **4** - Agree, **5** - Strongly agree

Statement		Rating				
		1	2	3	4	5
1	The simulator is visually realistic					
2	The skin feels realistic					
3	The stomach feels realistic					
4	The organs feel realistic					
5	The simulator is useful in training and education					
6	The simulator provided sufficient pain feedback					

Fig. 4. Abdominal task trainer post-test questionnaire.

The prototype includes four sensors located in each of the stomach's quadrants, shown in Fig. 5. The sensors track the relative applied pressure registered in each of the stomach quadrants over time. Analysing sensor data provides information on what areas of the stomach were covered by the test subject, and whether the correct amount of force was used. This provides a better understanding of how well users perform and can be used as a measure of the training quality.

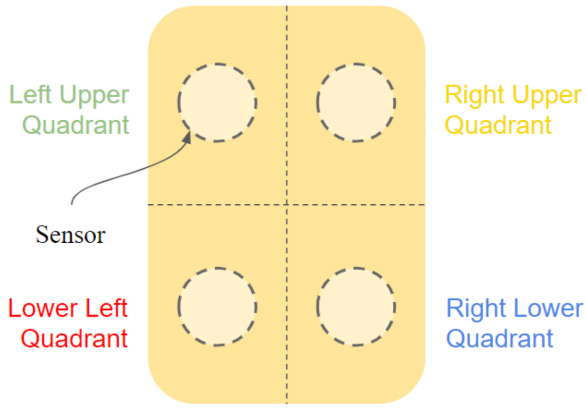


Fig. 5. Sensors placed in each quadrant of the stomach in the conceptual prototype.

Fig. 6. shows the data collected when a test subject was using the task trainer for different training cases. The X- and Y-axis show the time in milliseconds and relative pressure registered from each sensor respectively. The cyan line indicates what training case the user is working on, each increment shows the beginning of a new case. The cases and their correlating illness are shown below the X-axis.

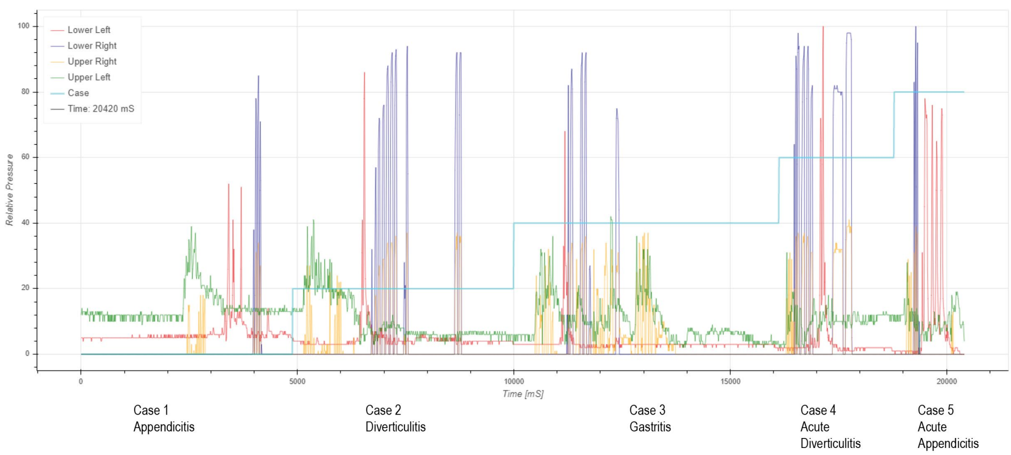


Fig. 6. The relative pressure registered by each sensor (Y-axis) over time (X-axis)

4 Discussion

In Wayfaring it is not self-evident how to perform effective tests as prototype resolution increases and as multiple questions are attempted answered within the same conceptual model.

However, performing these tests is still crucial as they might be solving known unknown challenges, validating known knowns or uncover unknown unknowns. It is therefore necessary to have enough tests in order to assure that previous iterations (and insights) are well represented and utilized in a concluding model. Prototypes exploring complex user-interactions and functionality, such as a medical task trainer, should not be tested in an overly simplified environment, test scenario or by leaving out critical product functionality. To confirm that the required ecological validity is present in a medical task trainer concept, a novel approach for testing of prototypes is proposed. Exemplified through a case project, the ambiguity of ecological validity has been addressed by utilizing a mixed method research approach and capturing both qualitative and quantitative data from user-testing.

In the case project, utilizing a mixed method research approach for prototype testing has been useful in order to inform future decisions and further development. The quantitative side of the applied method structures prototype feedback and illustrates how well a prototype performs based on results from all test subjects. Morse (2005), however, raises a concern of quantifying qualitative data by using forced-choice questionnaires. An alternative to giving a relative vote based on experience is to deploy multiple prototypes and hence use a ranked voting system. In terms of prototype resolution and complexity this could become costly – and time consuming, instead calling for a more comprehensive test. The qualitative side therefore helps explain and elaborate on the obtained results in the questionnaires. Further than merely investigating planned questions, unstructured interviews also allow following up on insight obtained during the actual test-procedure otherwise not captured.

Accommodating and leveraging sensor data is also important when performing complex prototype tests. In the context of this paper, sensor-integration in the prototype could enable both capturing and analysing how users interacted with the prototype. While these sensors were intended to generate feedback to users in training, collecting the sensor data could prove an important source of both insights and quantitative counterparts to the qualitative test results. This also calls for designers to be proactive when designing prototypes, and their corresponding tests, as sensors (redundant sensors) that might not seem valuable at the time could provide invaluable or supporting datapoints when analysing the results.

Traditionally in early-stage new product development the focus has been put on qualitative feedback collected from a limited number of key users. By making the data analysis more manageable through more structured data collection the novel approach proposed here enables collection of data from a greater number of users. An interesting aspect of collecting data in this fashion, is how to bring the qualitative and quantitative data from the test phase into the next phase of development. As the quantitative data can be used to prioritize the areas that need improvement, the qualitative feedback can shed light on why, and possibly how the improvement could be achieved. This allows features disrupting the ecological validity to be identified and solutions found based on feedback from the user.

Ambiguity and unknowns characterize the Fuzzy Front End, and whether the proposed test approach can mitigate some uncertainty from this development phase is worth consideration. By introducing objective and quantifiable data, unknowns and ambiguities can be identified and addressed in a structured and manageable way. Furthermore, it is interesting to see if this novel approach could aid projects in moving forward from the Fuzzy Front End. Eventually, when the needs have been elicited and functionality is identified, there is a need for measuring the prototype's performance and characteristics. Hopefully, the test approach can be used as a decision-making tool for determining when the early stage development phase is over and more traditional project management and development tools can take over. Can a new approach to testing, tailored to the prototype or test in question, decide when this point has been reached? This question cannot yet be answered, but future test projects will hopefully shed more light on these issues.

A final consideration when applying the approach proposed is that by increasing prototype fidelity and capturing multiple design questions in a combined conceptual model, a more comprehensive test scenario would be required. Thus, the fidelity of the test should match the fidelity of the prototype in question. When prototyping in the Fuzzy Front End, extensive testing will grind the project to a halt, and time and effort spent testing complex prototypes is meaningless if not sufficient data is collected. Hence, whether an approach, such as described in this paper, is expedient or not is yet to be determined. The disadvantages described by Dybvik (2018), i.e. greater time consumption and increased cost, should therefore be carefully considered and compared to an alternative test regiment without mixed method testing. Pilot tests show promising results and have provided improvements to the proposed approach. It has however only been tested on a limited number of projects.

5 Conclusion

The main contribution of this paper is a tool to help developers of medical task trainers ensure necessary ecological validity for learning objectives. This tool is a novel approach to testing able to verify if all independently developed functionalities are present, when combined in a conceptual prototype. The case project presented in this paper exemplifies the need for such a method. The same approach has also been proposed as a way of informing decisions when faced with complex user interaction or ambiguous product requirements, by utilizing a mixed method research method. The advantages and implications of combining quantitative and qualitative data was also discussed. More research is necessary in order to verify that the proposed approach to testing provides adequate results. The community is therefore asked to use this novel approach in their development projects, to obtain more data.

6 Acknowledgements

This research is supported by the Research Council of Norway through its industrial PhD funding scheme, project number 290404.

7 References

1. Abowitz Deborah A., & Toole T. Michael. (2010). Mixed Method Research: Fundamental Issues of Design, Validity, and Reliability in Construction Research. *Journal of Construction Engineering and Management*, *136*(1), 108–116.
2. Allen, I. E., & Seaman, C. A. (2007). Likert scales and data analyses. Quality progress, 40(7), 64-65.
3. Auflem, M., Erichsen, J. F., & Steinert, M. (2019). Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators. Procedia CIRP, 84, 572-578.

4. Dybvik, H. (2018). An exploration of mixed methods, researching business model phenomena-Investigating the business model phenomena through literature, in-depth case study, statistical testing and user experiments (Master's thesis, NTNU).
5. Ege, D.N., Lilleløkken, O., Auflem, M., & Steinert, M. (in press). Dealing with Ecological Validity and User Needs when Developing Simulation Based Training Equipment – Case Study of a Medical Palpation Task Trainer. *Procedia Cirp*.
6. Eris, O. (2003). Manifestation of divergent-convergent thinking in question asking and decision making processes of design teams: A performance dimension. In Human Behaviour in Design (pp. 142-153). Springer, Berlin, Heidelberg.
7. Gerstenberg, A., Sjöman, H., Reime, T., Abrahamsson, P., & Steinert, M. (2015). A Simultaneous, Multidisciplinary Development and Design Journey – Reflections on Prototyping. In K. Chorianopoulos, M. Divitini, J. Baalsrud Hauge, L. Jaccheri, & R. Malaka (Eds.), *Entertainment Computing—ICEC 2015* (pp. 409–416). Springer International Publishing.
8. Houde, S., & Hill, C. (1997). What do Prototypes Prototype? In M. G. Helander, T. K. Landauer, & P. V. Prabhu (Eds.), *Handbook of Human-Computer Interaction (Second Edition)* (pp. 367–381). North-Holland.
9. Kim, J., & Wilemon, D. (2002). Focusing the fuzzy front–end in new product development. *R&D Management*, *32*(4), 269–279.
10. Kriesi, C., Blindheim, J., Bjelland, Ø., & Steinert, M. (2016). Creating Dynamic Requirements through Iteratively Prototyping Critical Functionalities. *Procedia CIRP*, *50*, 790–795.
11. Kushniruk, A., Nohr, C., Jensen, S., & Borycki, E. M. (2013). From Usability Testing to Clinical Simulations: Bringing Context into the Design and Evaluation of Usable and Safe Health Information Technologies: Contribution of the IMIA Human Factors Engineering for Healthcare Informatics Working Group. *Yearbook of Medical Informatics*, *22*(01), 78–85.
12. Leifer, L., & Steinert, M. (2011). Dancing with Ambiguity: Causality Behavior, Design Thinking, and Triple-Loop-Learning. *Information-Knowledge-Systems Management*, *10*, 151–173.
13. Morse, J. M. (2005). Evolving Trends in Qualitative Research: Advances in Mixed-Method Design. *Qualitative Health Research*, *15*(5), 583–585.
14. Ramasesh, R. V., & Browning, T. R. (2014). A conceptual framework for tackling knowable unknown unknowns in project management. *Journal of Operations Management*, *32*(4), 190–204.
15. Schrage, M. (2004). Never go to a client meeting without a prototype [software prototyping]. IEEE Software, 21(2), 42-45.
16. Sobek II, D. K., Ward, A. C., & Liker, J. K. (1999). Toyota's principles of set-based concurrent engineering. MIT Sloan Management Review, 40(2), 67.
17. Steinert, M., & Leifer, L. (2012). 'Finding One's Way': Re-Discovering a Hunter-Gatherer Model based on Wayfaring. *International Journal of Engineering Education*, *28*, 251–252.
18. Vestad, H., & Steinert, M. (2019). Creating your Own Tools: Prototyping Environments for Prototype Testing. *Procedia CIRP*, *84*, 707–712.

C6

User Involvement in Early-Stage Design of Medical Training Devices

– Case of a Palpation Task Trainer Prototype

Design for Health

Daniel Nygård Ege, Marius Auflem, Oscar Lilleløkken, Martin Steinert (2021),

DESIGN FOR HEALTH
2021, VOL. 5, NO. 2, 214–232
https://doi.org/10.1080/24735132.2021.1980270

Routledge
Taylor & Francis Group

CASE STUDY

OPEN ACCESS Check for updates

User involvement in early-stage design of medical training devices – case of a palpation task trainer prototype

Daniel Nygård Ege, M. Auflem, O. Lilleløkken and M. Steinert

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway

ABSTRACT

Design of new medical training equipment is demanding as unavoidable complexity and ambiguity facing designers must be addressed in the early stages of the development. In this paper, a novel concept for abdominal palpation training is presented and used to exemplify challenges and approaches for designing new medical training equipment. Concluding the initial development of the palpation training concept, experienced medical personnel evaluated a conceptual prototype. A Likert scale questionnaire, a clinical assessment by participants, and recorded sensor data were used to evaluate prototype functionalities, perceived tactile- and visual realism, and usability of the concept in medical training. Obtained results are used to discuss insights for further development of the concept. Further, the paper discusses observations from user interactions and considerations regarding fidelity of medical training equipment prototypes. Moreover, it highlights the benefits of utilizing mixed-method research to identify areas of improvement for conceptual prototypes even before initiating industrial development efforts. By this, designers can ensure that user needs, product requirements, and sufficient fidelity are collectively captured in new medical training equipment in the early design stages.

ARTICLE HISTORY
Received 25 March 2021
Accepted 2 September 2021

KEYWORDS
Medical simulation; abdominal palpation; simulation-based training equipment; product development; user involvement; early-stage design

Background

In medical education, repetition and applying theoretical knowledge through skill training are essential for various procedures, including abdominal palpation (Cooper and Taqueti 2004). As it relies on tacit knowledge and recognizing distinct symptoms and tactile sensations, abdominal palpation is a fundamental skill that requires hands-on training (Aubin, Gagnon, and Morin

CONTACT Daniel Nygård Ege danieneg@stud.ntnu.no Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Postboks 8900, Trondheim NO-7491, Norway

2014). In the procedure, practitioners use fingers and palms to localize areas with abnormal tactility (due to inflammation, anomalies, trapped fluids, and others) or areas causing pain to the patient. Abdominal palpation training (APT) is thus crucial for teaching primary diagnostics and for students to practice psychomotor skills and procedural algorithms in a safe and repeatable environment. This calls for solutions enabling realistic physical interactions, as the tactile sensations and techniques are hard to formalize and describe by theory alone.

APT is generally performed by fellow students acting as the patient to prevent risk and discomfort for actual patients, limiting the trainee's objective feedback. Furthermore, such training does not capture the tactility of abdominal symptoms and corresponding illnesses. There are examples of research and development efforts attempting to provide APT equipment for education, proposing physical (Yakubo et al. 2008; Arita et al. 2019; ABSIM 2020; Laerdal 2021), virtual (Dinsmore et al. 1997), and hybrid patient simulation solutions (Ribeiro et al. 2016; Escobar-Castillejos et al. 2016; Ullrich and Kuhlen 2012). However, these solutions are generally expensive, non-continuously available, or lacking critical functionality such as objective performance feedback or tactile/physical feedback. Another potential issue for the current training solutions is the finite set of study cases to practice on, and not being able to tailor these to the various disciplinary and contextual norms of APT.

Given the current APT challenges, a leading provider of medical training equipment proposed a new development project to explore and conceptualize a solution for simulation-based APT. This project developed a concept through iterative prototyping and evaluated it using a multi-faceted test framework for the user-centred design of medical training equipment. Development and testing revealed the need for an easy-to-use solution that could provide users with objective feedback, enabling a large (and expandable) set of study cases and facilitate realistic tactile interactions. The resulting conceptual prototype is a low-cost implementation of readily available hardware, software, and materials that could be developed further as an open hardware solution in the future.

As the design of equipment for medical training can impact the performance of healthcare professionals and thus patient safety (Leape et al. 1991; Lawton et al. 2012; Zhang et al. 2003), it is essential to ensure sufficient clinical realism through user involvement in the development process (Money et al. 2011; Blandford, Furniss, and Vincent 2014). However, sufficient realism can be difficult to manage, given a large chasm between designers and healthcare professionals (Zenios et al. 2009). Although similarities between healthcare and design exist (Rowe, Knox, and Harvey 2020), generally, differences in methodology and procedures are perceived as challenging for designers (Groeneveld et al. 2018).

This paper presents how structured prototype testing with expert users can address the disciplinary gap between healthcare professionals and designers. Following the initial development of an APT concept, experienced medical personnel assessed a prototype and conducted a diagnosis for a set of different simulated conditions. The users were then requested to evaluate and provide feedback on the presented concept. This study aimed to answer whether the concept met the needs and requirements of healthcare professionals and if the prototype provided sufficient fidelity and clinical realism to be used in APT. With this, we provide an example of how structured user testing can evaluate the required functionalities and fidelity of prototypes during the development of new medical training equipment.

Development of prototype

In response to the highlighted challenges with APT, Ege et al. (2020) has developed a novel simulator concept to be used in medical education (Figure 1). This development project used an iterative and prototype-driven approach, accommodating the knowledge gap and ambiguity facing design teams in the early stages of product development (Gerstenberg et al. 2015; Motavalli and Nestel 2016). Multiple low-resolution prototypes were designed, built, and tested to elicit and explore functional requirements and necessary fidelity (Ege et al. 2020).

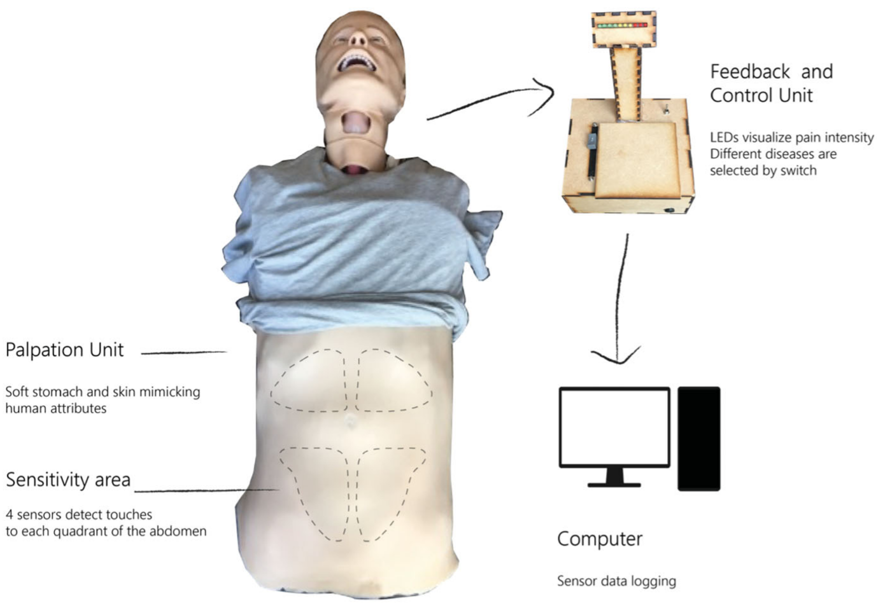


Figure 1. The APT simulator developed by Ege et al. (2020) and assessed by expert users in this paper.

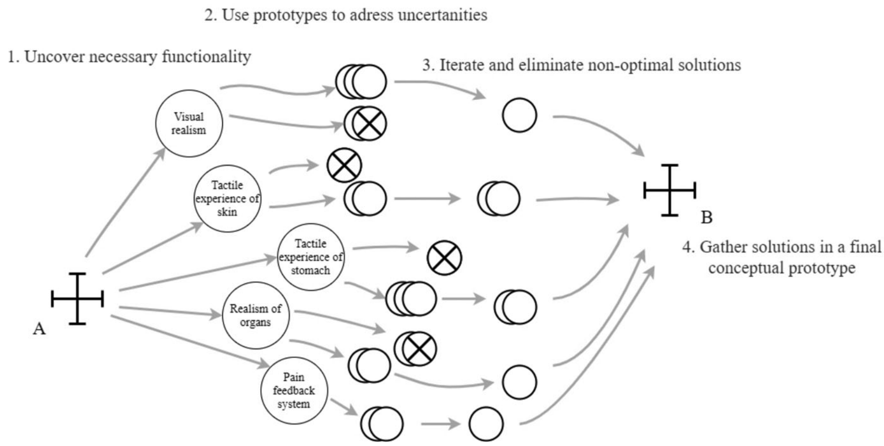


Figure 2. Illustration of prototyping activities guiding development (Lilleløkken et al. 2020).

The simulator consists of a soft stomach with an embedded sensor system. Three layers of different silicone rubbers make up the stomach and skin. It encompasses a force sensor in each quadrant of the stomach, measuring both forces applied when palpated and their location. Sensor readings were correlated to pain levels by having expert users interact with early prototypes and suggest how painful the pressure they applied would feel. The sensor's variable voltage is measured and mapped into ten pain levels, ranging from low amounts of pain to high. An array of ten LEDs light up to visualize pain intensity during an examination of the prototype. Simulated palpable organs, including the liver and gallbladder, can be inflated to indicate inflammation, giving the task trainer the possibility of simulating various diseases and symptoms. A total of 10 common abdominal diseases are currently implemented, including appendicitis, diverticulitis, and gastritis, and could easily be increased or changed based on the need of educators. The task trainer can simulate both mild and acute symptoms of appendicitis and diverticulitis. The intended disease corresponds to typical symptoms such as pain intensity and location and signs of rebound tenderness and enlarged organs. A computer logs sensor data allowing for post-test analysis of trainees. The task trainer is integrated into a modified commercially available medical mannequin, modelled to look and feel realistic with anatomical landmarks such as ribs and hip bones.

Medical professionals were consulted in the earliest phases of design to elicit requirements for a new training device. Figure 2 illustrates how prototyping activities guided development, building on previous findings and requirements. In addition, it shows how prototypes addressed uncertainty by building multiple solutions and testing them to pick the best one before combining all sub-systems in a final prototype.

While quantifying attributes such as tactility and the visual resemblance is challenging, designers captured an appropriate resemblance in a final

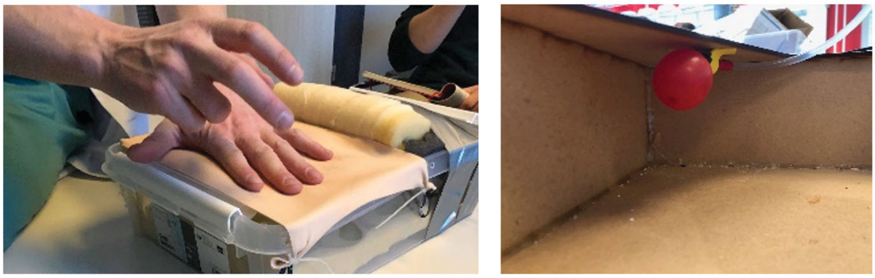


Figure 3. (a, b) Low-fidelity prototypes built early in the project.

prototype by building and testing multiple rapid and low-cost prototypes qualitatively. An initial, rough prototype consisting of foams, waterfilled condoms, and a skin-coloured PVC sheet (Figure 3(a)) was built as an initial idea of what an APT simulator could be. As well as being a valuable discussion point when meeting experts, this prototype yielded important insights challenging to capture using quantitative measures, such as the palpation technique and tactile realism. Likewise, solutions to simulate organs changing tactility due to disease were also prototyped using simple, readily available materials. For example, based on a doctor describing an inflamed gallbladder feeling like a soft plum, several prototypes were made to replicate what the designers believed it should feel like before experts confirmed or invalidated its tactile realism. A prototype aiming to simulate an inflamed gallbladder is showed in Figure 3(b), where a balloon filled with water replicates the tactility the expert described.

Iterative prototyping and continuous user interactions, focussing on hands-on testing and interviews, guided development towards the conceptual prototype presented in this paper. Early designs with increasing complexity and realism are showed in Figure 4, where more comprehensive sensor systems and more realistic anatomical landmarks were included as development progressed.

The development of this task trainer prototype supports that before investing time and resources into building refined models/products, low-resolution prototypes should be rapidly developed and evaluated by users (Ege et al. 2020; Auflem, Erichsen, and Steinert 2019). This could aid both eliciting and solving concrete user needs and mitigate the risk of encountering severe obstacles later in the development (where it is more costly and time-consuming to address these) (Takeuchi and Nonaka 1986; Steinert and Leifer 2012; Thomke and Reinertsen 1998). In order to keep development agile and rapid, qualitative methods such as interviews and hands-on testing proved helpful in this project. User testing requires both reflective and intentional prototyping approaches as users are unaware of which aspects of the design idea the prototype represents (Houde and Hill 1997; Reay et al. 2017).

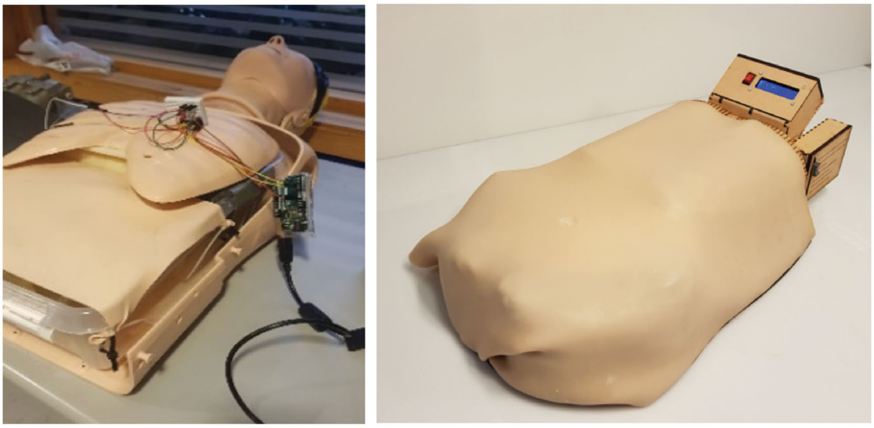


Figure 4. Prototypes with increasing complexity and fidelity.

More so, it is not evident which level of resolution (amount of detail) and fidelity (closeness to eventual design) a prototype requires to sufficiently convey a design idea and be deemed realistic enough to gain feedback. Especially when designing new medical training equipment, products mainly involved in complex real-world interactions, oversimplification may degrade the realism of a test to a level that is no longer valid (Lilleløkken et al. 2020). In early-stage development, qualitative feedback is often collected from a small sample of key users. However, increasingly complex prototypes call for increasingly comprehensive test procedures to accommodate them. Lilleløkken et al. (2020) suggest using a multi-faceted test framework utilizing quantitative methods strategically, as done in this paper.

Method

A mixed-method research approach generated three separate data sets: sensor data, questionnaire answers, and a clinical assessment performed by participants. Both qualitative and quantitative data were gathered to guide further development iterations. This research project was conducted over a 10-week period, where testing was performed with four participant groups of healthcare professionals, constituted by physicians and nurses, ranging from novice to experts.

Participants

Eighteen healthcare professionals participated in this study. Subjects were divided into four groups based on their profession and level of experience. G1 was a gastroenterologist with 40 years of experience. G2 consisted of newly educated physicians doing their first year of hospital training. G3 were general practitioners with an average of 8 years of experience. G4 consisted of experienced nurses, in administrative and educational positions, with an

Table 1. Demographical information (SD in parentheses).

D	Item	G1 ($n = 1$)	G2 ($n = 9$)	G3 ($n = 2$)	G4 ($n = 6$)
1	Profession	Gastroenterologist	Physician in the first year of training	General practitioner	Nurse
2	Experience (# of years [SD])	40	1 (0.66)	8 (9.9)	16.8 (2.85)
3	Gender (#)	1♂	2♂, 7♀	1♂, 1♀	1♂, 5♀
4	Age (mean [SD])	62	28.2 (1.7)	38 (9.9)	43.6 (3.9)
5	Expertise (#)	1/1	8/9	2/2	0/6

average of 16.8 years of experience. Table 1 presents demographic information and how often participants perform abdominal examinations. Item D5, Expertise, describes the number of participants that perform abdominal palpation weekly from each participation group.

Procedure

Data were collected during four separate sessions, one for each participant group, in a similar environment, and a hospital setting. All examinations were done individually in a separate room from other participants. Participants were introduced to the APT simulator before how it works and how it indicates pain was demonstrated. It was then prepared to simulate different diseases. Because of different expertise and availability, the procedure varied between participation groups. G1 and G2 followed Scenario 1, while G3 and G4 followed Scenario 2, as illustrated in Figure 5.

In Scenario 1, participants from G1 and G2 were first presented with a list of 10 common abdominal diseases, shown in Figure 6(a). They were asked to examine the task trainer six times and propose a diagnosis from the list for each examination. Participants were allowed to use as much time as they wanted to gather findings and to determine diagnoses. The order in which diseases were simulated was the same for all participants.

In Scenario 2, participants from G3 and G4 were asked to examine the task trainer twice. For each examination, they proposed a diagnosis. Participants were allowed to use as much time as they wanted to gather findings and to determine diagnoses. The order of the simulated diseases was the same for all the participants in each test group.

Immediately after examinations, all participants answered a 5-point Likert scale questionnaire.

Results

Clinical assessments

After each examination, participants were asked to propose a diagnosis for the task trainer. Table 2 shows how many of the participants gave the correct diagnosis for each of the different diseases that were simulated.

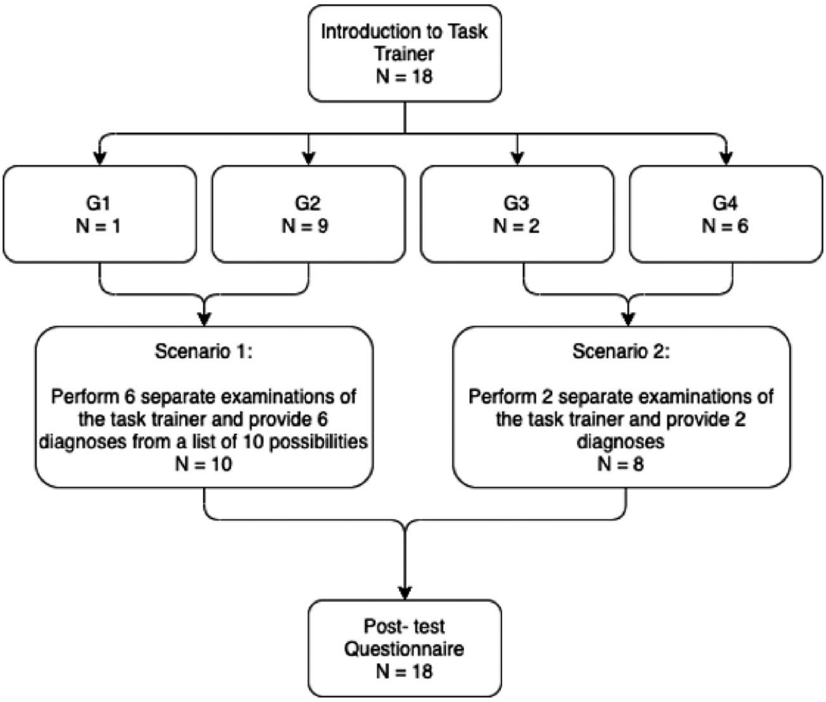


Figure 5. Test Procedure.

a

Disease
Pancreatitis
Intestinal Obstruction
Appendicitis
Acute appendicitis
Gastritis
Diverticulitis
Acute diverticulitis
Inflamed gallbladder
Peritonitis
Inflamed liver

b

Figure 6. (a) List of possible diseases. (b) The task trainer being examined by a physician.

The participants from G1 and G2 who followed the procedure in Scenario 1 could score six correct diagnoses. The participant from G1 scored six out of six correct diagnoses. From G2, three participants managed to provide five correct diagnoses, three provided four correct diagnoses, two provided three correct diagnoses, while the last participant managed to provide two correct diagnoses. The mean accuracy score was 4.1 out of six possible, with a standard deviation of 1.11.

Table 2. Diagnose assessment.

D	Diagnose	G1 ($n = 1$)	G2 ($n = 9$)	G3 ($n = 2$)	G4 ($n = 6$)
1	Appendicitis	100%	78%	100%	100%
2	Diverticulitis	100%	100%	100%	33%
3	Gastritis	100%	33%		
4	Acute Diverticulitis	100%	88%		
5	Acute Appendicitis	100%	88%		
6	Inflamed gallbladder	100%	11%		

Table 3. Questionnaire answers.

		G1		G2		G3		G4	
Q	Item	M	SD	M	SD	M	SD	M	SD
1	The simulator is visually realistic	5	0	3.44	0.68	3	0	3.33	0.75
2	The skin feels realistic	4	0	3	0.67	2.50	2.12	3.5	0.76
3	The stomach feels realistic	3	0	2.56	0.83	2.50	0.71	3.67	0.75
4	The organs feel realistic	3	0	2.11	0.87	2	0	2.83	0.90
5	The simulator is useful in training and education	5	0	3.89	0.87	4	1.41	4.17	0.69
6	The simulator provided sufficient pain feedback	5	0	3.89	0.87	4	0	4	1.00

Questionnaire answers

A questionnaire determined how well participants agreed with six statements regarding the prototype's physical, visual, and conceptual level. A 5-point Likert scale, where 1 = strongly disagree and 5 = strongly agree, was used to rate the level of agreement to each statement.

Table 3 summarizes questionnaire answers for all participant groups. Participants agreed the strongest with the question regarding usefulness in education, while they disagreed most with the realism in the feel of stomach and organs.

On average, G1 agreed most with the statements in the questionnaire ($M = 4.17$, $SD = 0.98$). G4 ($M = 3.58$, $SD = 0.48$) and G2 ($M = 3.15$, $SD = 0.72$) agreed less with the statements, while G3 ($M = 3.0$, $SD = 0.84$) agreed the least to the statement of all participating groups.

Collected sensor data

From a development context, sensor data was not intended for post-test analysis but rather to make the prototype function as intended. However, observed trends in the recorded sensor data show potential for uses exceeding the primary goals. Each sensor registers forces applied over time. Therefore, visualizing readings from all the sensors provides an overview of how the user performed the examination. In the visualization, the forces exerted on each of the sensors are visible, showing the palpation activity (either magnitude or location) in relative proximity to each of the sensors. Further timing, such as rhythm and time to trigger all sensors, can be observed in the raw data visualization. An example of the raw data gathered

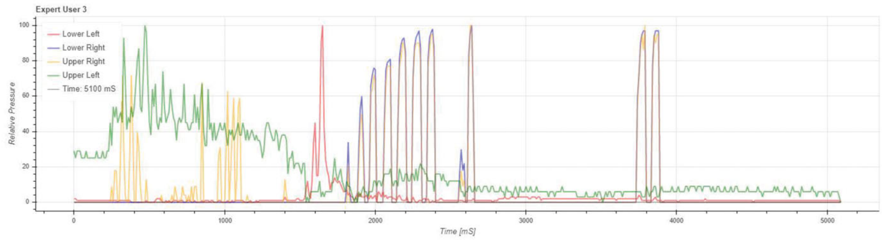


Figure 7. A plot of raw sensor readings from each quadrant of the stomach over time.

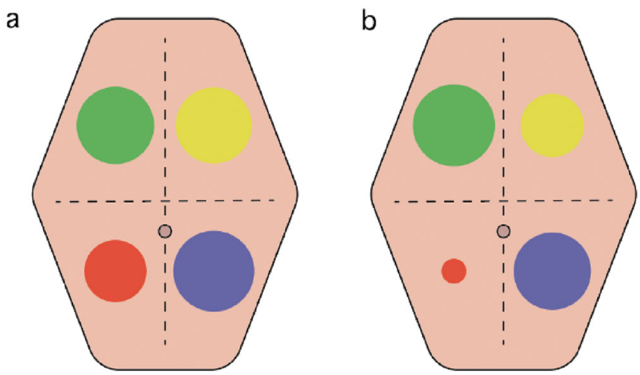


Figure 8. The cumulated force applied to each quadrant of the stomach. (a) shows an expert user's examination. (b) shows a novice user's examination.

from sensors during a test scenario where one disease was examined is depicted in Figure 7.

Cumulating the magnitude of force applied to each quadrant can visualize how evenly an examination is performed and whether the subject focuses equally on each quadrant. In Figure 8, the circles in each quadrant relate to the total force registered from each sensor over a simulated case. Similar-sized circles are positive, while variable sizes show uneven focus during examination. Figure 8(a) shows the plot of an expert user examining the task trainer for appendicitis. Figure 8(b) show a novice users plot doing the same task.

An observation made during testing was that the amount of force applied to the task trainer when the pain was identified increased. Several participants seemed to try to max out the pain scale when detecting pain in a region, possibly to verify that the pain was indeed present.

Discussion

Interpretation of results

This study shows that the task trainer can be used for diagnosis practice and symptom recognition successfully. Results, therefore, indicate that it has the functionality and attributes necessary for facilitating APT. However, it is not possible to ascertain the learning outcome of using the task trainer at this stage and thus its training potential. When comparing participants in the clinical assessment, it is clear that diagnosing appendicitis and diverticulitis yielded high success rates. This might be caused by the fact that these diseases have distinct symptoms different from other diseases. Regardless, the task trainer seems to replicate these distinct symptoms adequately. Also apparent in the obtained results is that high success rates can be seen for both acute and milder symptoms of appendicitis and diverticulitis, showing continuity in answers. This strengthens the hypothesis that simulated symptoms are adequately incorporated in the task trainer.

In the obtained questionnaire answers, participants neither strongly agree nor disagree with most statements. An interesting observation is that all groups disagree the most with the same statement (Q4). Similarly, they agree the most with the same statement (Q5). Participants not strongly agreeing with certain statements is not unexpected, given that the presented task trainer is a prototype and not a finished commercially available product. Nevertheless, the results are still of interest for further development, as they highlight parts of the concept with the most considerable improvement potential. For instance, improving tactile realism of enlarged organs is of interest as the users rated this functionality poorly. Further, participants agreeing to the usefulness of the concept in education is promising, suggesting further development. These insights illustrate why a Likert scale questionnaire was utilized. It proved a convenient way of capturing participants' feedback on the current prototype's physical, visual, and conceptual traits.

Limiting the amount of pain endured by a patient throughout an examination is essential in a clinical examination. By investigating the sensor data obtained during testing, how much force each participant has exerted can be determined, making it possible to detect whether the patient experienced excessive pain. Informing trainees of this during training with a task trainer can make them aware of how much force to apply during palpation. Secondly, the pattern of how trainees divide the abdomen and examine it quadrant by quadrant can also be derived from sensor data. This can be used to determine if a trainee methodically examines each quadrant, one at a time, to make sure the entire stomach is palpated. This way, he/she can be sure of where pain occurs and can determine a diagnose. The examination pattern also allows for comparison between experienced and inexperienced users, which could provide meaningful data for creating self-directed learning scenarios. The amount of time used on each scenario is also evident in sensor data, giving yet another basis for comparison.

How well quadrants have been covered by palpation can be derived by combining the total force applied to each quadrant. This can be used to visualize uneven examination for trainees. It also allows for comparison with an expert and can highlight incorrect techniques. For example, Figure 8(a) shows how an expert pays equal attention to each quadrant. However, in Figure 8(b), a novice user is seen to focus more on some parts of the abdomen, possibly missing out on vital information to determine a correct diagnose. These findings in the obtained data should be investigated further, as they could be valuable as objective performance measures to novice trainees and to give corrective feedback during and post APT scenarios.

Comments on the prototype

The prototype's usefulness as a learning tool was rated highly across all participation groups, indicating its applicability in education. In addition, most users also agreed that the prototype's ability to indicate pain was sufficient. Detection of pain has been described as the most prominent part of abdominal examinations, as the location of pain often is the most telling symptom for common abdominal illnesses. With the ratings of the other assessment points scoring around the middle of the 5-point scale, it is an indication that the users find the features offered by the prototype sufficient for practising the clinical skill for palpation.

As shown in the results, the force applied to the task trainer when the pain was identified increased. In a real-life scenario, the examiner would likely have gradually increased the force applied in painful regions, careful not to cause excessive pain to the patient. However, during testing of the task trainer, it recorded high pain levels in many cases even after the pain was located, as previously described. This behaviour is a sign of wrong technique and should be addressed by either a supervisor during training or by the task trainer itself. If used as a self-directed learning tool, corrective feedback should be given to trainees as avoidable pain to any actual patient should be mitigated. This phenomenon might be caused by a lack of immersion in the training scenario and lacking a more natural pain response. Test subjects were seen to be immersed going into the case. The physical appearance of the task trainer is realistic, as it is integrated into a full-size torso with accurate anatomical landmarks. Some participants were even concerned with warming up their hands before touching the 'patient' during an examination. However, the realism of the LED pain response significantly differs from the rest of the task trainer. This can have caused subjects to begin experiencing the case more as a game of 'finding the right spot' rather than a real scenario with a patient. The objective of the simulation can get lost and become overshadowed by the participants' will to succeed in what they experience as a game (Frank 2012; Rieber and Noah 2008). Human pain response includes all senses, including audible responses, visual responses like facial expression or body movement, and haptic responses from guarding, that is, muscle contraction in response to pain. Pain feedback from the task trainer is only addressing the visual aspect of the experience. Therefore, other sensory elements might need to be included in the future development to increase realism. This is in line with what (Stokes-Parish, Duvivier, and Jolly 2020) found, where a lack of immersion caused students to 'do tasks for the sake of doing it' during simulation, as opposed to performing as if the simulator was an actual patient.

Comments on fidelity

Simulation of clinical practices might be misrepresented as relatively simplistic (Motavalli and Nestel 2016), making it essential for designers of training equipment to acknowledge and consider the complexity and ambiguity of clinical simulation. Parts of this ambiguity and complexity can be attributed to finding the appropriate fidelity of a new medical simulator.

The fidelity of a simulator has been shown to affect trainee's engagement in training, where more realistic simulators tend to increase engagement and a more immersive experience (Stokes-Parish, Duvivier, and Jolly 2020; Hagiwara et al. 2016). Immersion is essential in simulation as it can improve knowledge transfer to real-world scenarios (Dede 2009). However, an increase in fidelity does not necessarily lead to a higher learning outcome (Issenberg et al. 2005; Hamstra et al. 2014; Scerbo and Dawson 2007; Norman, Dore, and Grierson 2012). Although different methods for benchmarking simulator fidelity exists (Wilson et al. 2018), in the early development stages, it is not evident what the required fidelity for a simulator is. It can, however, as shown in this paper, be elicited through early-stage user testing.

The aspect of fidelity is somewhat difficult to comprehend in the early design stages. It concerns both the development stage a prototype is in and the intended fidelity of a finished product. The fact that a prototype is being tested instead of a finished product will influence results. This introduces a significant limitation in this stage of product development, as building high-fidelity prototypes would be too costly. Nevertheless, results in this study show that functionality and fidelity are within what users believe is sufficient without relying on a high-fidelity prototype. This paper advocate that designers should focus on finding what is realistic enough to obtain valid test results during the early development of medical training equipment, as opposed to being as realistic as possible. In this way, designers can avoid unnecessary complexity and expensive, time-consuming challenges that could be addressed at a later stage of development.

In this development project, prototypes were tested throughout by expert users, thus pinpointing what had to be improved for each prototype iteration. In order to conclude early development, ensuring that functionality, appropriate fidelity, and fit within curricula are all sufficiently addressed in a concept, structured testing proved crucial, as it answered these questions.

Implications on the further development

The conceptual prototype presented in this paper has shown to be useful for tactile recognition and abdominal diagnosis exercise. It has been suggested that the increased fidelity in a task trainer does not necessarily result in better learning outcomes. Thus, future development will aim to create realistic training approaches and scenarios that can facilitate effective training. Moreover, the intent is to find ways to measure training outcomes to inform both trainees and educators.

To measure training outcome, a benchmark to compare participants against is needed. By comparing sensor data from participants with a gold standard examination recorded in expert user testing, trainees can be assessed. This can also allow for supplementing deficit-oriented corrective training approaches usually characterized with simulators, with an approach that focuses more on how good performance is produced, as described by (Dieckmann et al. 2017). Furthermore, sensor data could inform neural networks and algorithms, where pattern recognition can provide feedback and corrective suggestions. Location of touch, the areas covered by examination, peak forces, and timing is already possible to capture using the integrated sensors. This data could enable algorithms to locate not only simple pain points but also follow complete diagnostic scenarios.

Testing revealed that the simulation of enlarged organs was not satisfactorily captured in this prototype. Most participants did not even find the enlarged organs or realize that the simulator could emulate them. This exemplifies the importance of training before introducing a novel device, as discussed by Reid-McDermott et al. (2019), as trainees know what the simulator will test. It also means it might be necessary to exaggerate this functionality, that is, increase the size of organs to make them more distinguishable by trainees. This functionality could also be more dynamic where size and hardness could be altered to ensure a specific level of difficulty. Hence, training could be tailored to a specific scenario, curricula, or even individual skill levels.

Limitations

This study aimed to support further development of a conceptual prototype and not assert if trainees improve in abdominal examination using the prototype. Thus, it was considered redundant to correlate questionnaire data, sensor data, and clinical assessment of participants. This, for instance, implies that it is unknown how the sensor data of successful participants compare to those who are less successful in diagnosing the task trainer. However, it is of great interest for future tests as it speaks to the prototype's value as a learning tool. The ability to distinguish between skill levels is necessary to provide feedback to the user. Identifying areas of improvement can only be achieved by locating where a trainee's input varies from a standard for successful examinations.

Data presented here is based on small and unbalanced groups, thus limiting its general validity. However, test data was primarily intended to enlighten further development of a prototype, and it was not considered necessary to increase the number of participants. As previously shown, the data gathered has proved helpful in pinpointing shortcomings of the current prototype and eliciting ways of going forward, thus fulfilling its intention. Due to the uncertainty caused by small and unbalanced groups, the findings should not necessarily be generalized to all other projects of medical training equipment. However, as this way of developing and testing medical task trainer prototypes proved helpful in this case project, it might be worth investigating further in other cases.

Conclusion

This paper has presented the results of structured testing of an APT simulation prototype with expert users. Participants performed an abdominal examination of the prototype and provided a diagnosis based on their findings. Three data set were collected from each participant; a clinical assessment, a Likert scale rating of the prototype's features, and sensor data collected by the prototype. We found that participants correlated the simulated symptoms with high success rates to several medical diagnoses attempted recreated from the assessment. The continuity of these results further supports the concept as a valuable tool for the enabled practice of abdominal examination and to diagnose based on this procedure.

Further, subjective feedback from the questionnaire has highlighted areas of improvement for the conceptual prototype. The questionnaire results also support the concept as valuable for use in medical education. Obtained sensor data from the tests allow for technique and routine to be investigated in relation to the assessment results. It is discussed how this type of data could be leveraged for developing self-directed learning scenarios and objective feedback based on palpation data generated by experienced physicians. Using this data, insights concerning accumulated palpation force exerted (relating to pain), timing and coverage of the abdomen are observed. The prototypes' functionality and implications, and considerations for further development of the concept are discussed based on findings from the test results. By the mixed-method structured testing, insights and considerations for developing medical training equipment are also exemplified by this case project. Finally, we discuss why the findings in this project should not necessarily be generalized to all other projects of medical training equipment.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The Research Council of Norway supports this research financially through its industrial PhD funding scheme, project numbers 290404 and 321386.

Notes on contributors

Daniel Nygård Ege works in Nasjonalparken Næringshage as an industrial PhD candidate through the Department of Mechanical and Industrial Engineering at the Norwegian University of Science and Technology. He does research in early-stage product development, focussing primarily on prototyping, the role of makerspaces and open-source hardware. His research is part of TrollLabs, a Prototyping and product development lab/community at NTNU in Trondheim, Norway.

Marius Auflem currently works at Laerdal Medical as an industrial PhD candidate through the Department of Mechanical and Industrial Engineering at the Norwegian University of Science and Technology. Marius does research in early-stage Product Development in the context of medical training and simulation equipment. His research is part of TrollLabs, a Prototyping and product development lab/community at NTNU in Trondheim, Norway.

Oscar Lilleløkken is a mechanical engineering master student at the Department of Mechanical and Industrial Engineering at the Norwegian University of Science and Technology. He is doing his master thesis on prototyping medical training equipment concepts. The research is part of TrollLabs, a Prototyping and product development lab/community at NTNU in Trondheim, Norway.

Martin Steinert is a professor of engineering design and innovation in the Department of Mechanical and Industrial Engineering at the Norwegian University of Science and Technology (NTNU). He teaches fuzzy front-end engineering for radical new product/service/system concepts and graduate research seminars for PhD students engaged in topics related to new product design and development. His various research projects are usually multidisciplinary (ME/CS/EE/Neuro- and cognitive science) and are often connected with the industry. The aim is to uncover, understand, and leverage early-stage engineering design paradigms with a special focus on human-machine/object interactions.

References

ABSIM. 2020. “ACDET-ABSIM | Abdominal Simulator | USA.” Accessed 1 June 2020. https://www.acdet-absim.com/.

Arita, R., T. Numata, N. Saito, S. Takayama, T. Togashi, S. Kaneko, A. Kikuchi, et al. 2019. “Development of a Medical Education Program with Abdominal Palpation Simulators to

Support the Understanding of Traditional Japanese (Kampo) Medicine in Beginners." *Traditional & Kampo Medicine* 6 (3): 148–155. doi:10.1002/tkm2.1230.
Aubin, A., K. Gagnon, and C. Morin. 2014. "The Seven-Step Palpation Method: A Proposal to Improve Palpation Skills." *International Journal of Osteopathic Medicine* 17 (1): 66–72. doi:10.1016/j.ijosm.2013.02.001.
Auflem, M., J. F. Erichsen, and M. Steinert. 2019. "Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators." *Procedia CIRP* 84: 572–578. doi:10.1016/j.procir.2019.04.202.
Blandford, A., D. Furniss, and C. Vincent. 2014. "Patient Safety and Interactive Medical Devices: Realigning Work as Imagined and Work as Done." *Clinical Risk* 20 (5): 107–110. doi:10.1177/1356262214556550.
Cooper, J. B., and V. R. Taqueti. 2004. "A Brief History of the Development of Mannequin Simulators for Clinical Education and Training." *BMJ Quality & Safety* 13 (l1): 11–18. doi: 10.1136/qshc.2004.009886..
Dede, C. 2009. "Immersive Interfaces for Engagement and Learning." *Science* 323 (5910): 66–69. doi:10.1126/science.1167311.
Dieckmann, P., M. Patterson, S. Lahlou, J. Mesman, P. Nyström, and R. Krage. 2017. "Variation and Adaptation: Learning from Success in Patient Safety-Oriented Simulation Training." *Advances in Simulation* 2 (1): 21. doi:10.1186/s41077-017-0054-1.
Dinsmore, M., N. Langrana, G. Burdea, and J. Ladeji. 1997. "Virtual Reality Training Simulation for Palpation of Subsurface Tumors." Paper presented at the IEEE 1997 Annual International Symposium on Virtual Reality, Albuquerque, USA, March 1–5.
Ege, D., O. Lilleløkken, M. Auflem, and M. Steinert. 2020. "Dealing with Ecological Validity and User Needs When Developing Simulation Based Training Equipment – Case Study of a Medical Palpation Task Trainer." *Procedia CIRP* 91: 722–727. doi:10.1016/j.procir.2020.03.116.
Escobar-Castillejos, D., J. Noguez, L. Neri, A. Magana, and B. Benes. 2016. "A Review of Simulators with Haptic Devices for Medical Training." *Journal of Medical Systems* 40 (4): 104. doi:10.1007/s10916-016-0459-8.
Frank, A. 2012. "Gaming the Game: A Study of the Gamer Mode in Educational Wargaming." *Simulation & Gaming* 43 (1): 118–132. doi:10.1177/1046878111408796.
Gerstenberg, A., H. Sjöman, T. Reime, P. Abrahamsson, and M. Steinert. 2015. "A Simultaneous, Multidisciplinary Development and Design Journey – Reflections on Prototyping." In *Entertainment Computing - ICEC 2015*, edited by K. Chorianopoulos, M. Divitini, J. Baalsrud Hauge, L. Jaccheri, and R. Malaka, 409–416. Cham, Switzerland: Springer International Publishing.
Groeneveld, B., T. Dekkers, B. Boon, and P. D'Olivo. 2018. "Challenges for Design Researchers in Healthcare." *Design for Health* 2 (2): 305–326. doi:10.1080/24735132.2018.1541699.
Hagiwara, M. A., P. Backlund, H. M. Söderholm, L. Lundberg, M. Lebram, and H. Engström. 2016. "Measuring Participants' Immersion in Healthcare Simulation: The Development of An Instrument." *Advances in Simulation* 1 (1): 17. doi:10.1186/s41077-016-0018-x.
Hamstra, S. J., R. Brydges, R. Hatala, B. Zendejas, and D. A. Cook. 2014. "Reconsidering Fidelity in Simulation-Based Training." *Academic Medicine* 89 (3): 387–392. doi:10.1097/ACM.0000000000000130..
Houde, S., and C. Hill. 1997. "What Do Prototypes Prototype?." In *Handbook of Human-Computer Interaction (*Second Edition), edited by Marting G. Helander, Thomas K. Landauer, and Prasad V. Prabhu, 367–381. Amsterdam: North-Holland. doi:10.1016/B978-044481862-1.50082-0.

Issenberg, S. B., W. C. Mcgaghie, E. R. Petrusa, D. Lee Gordon, and R. J. Scalese. 2005. "Features and Uses of High-Fidelity Medical Simulations That Lead to Effective Learning: A BEME Systematic Review." *Medical Teacher* 27 (1): 10–28. doi:10.1080/01421590500046924.

Laerdal. 2021. "Abdominal Examination Trainer." Accessed 11 January 2021. https://laerdal.com/no/products/skills-proficiency/limbs-things/abdominal-examination-trainer/

Lawton, R., R. R. C. McEachan, S. J. Giles, R. Sirriyeh, I. S. Watt, and J. Wright. 2012. "Development of an Evidence-Based Framework of Factors Contributing to Patient Safety Incidents in Hospital Settings: A Systematic Review." *BMJ Quality & Safety* 21 (5): 369–380. doi:10.1136/bmjqs-2011-000443.

Leape, L. L., T. A. Brennan, N. Laird, A. G. Lawthers, A. R. Localio, B. A. Barnes, L. Hebert, J. P. Newhouse, P. C. Weiler, and H. Hiatt. 1991. "The Nature of Adverse Events in Hospitalized patients. Results of the Harvard Medical Practice Study II." *The New England Journal of Medicine* 324 (6): 377–384. doi:10.1056/NEJM199102073240605.

Lilleløkken, O., D. Ege, M. Auflem, and M. Steinert. 2020. "Proposing a Novel Approach for Testing Complex Medical Task Trainer Prototypes." Paper presented at NordDesign 2020, Lyngby, Denmark, August 12–14.

Money, A. G., J. Barnett, J. Kuljis, M. P. Craven, J. L. Martin, and T. Young. 2011. "The Role of the User within the Medical Device Design and Development Process: Medical Device Manufacturers' Perspectives." *BMC Medical Informatics and Decision Making* 11 (1): 15. doi:10.1186/1472-6947-11-15.

Motavalli, A., and D. Nestel. 2016. "Complexity in Simulation-Based Education: Exploring the Role of Hindsight Bias." *Advances in Simulation* 1: 3. doi:10.1186/s41077-015-0005-7.

Norman, G., K. Dore, and L. Grierson. 2012. "The Minimal Relationship between Simulation Fidelity and Transfer of Learning." *Medical Education* 46 (7): 636–647. doi:10.1111/j.1365-2923.2012.04243.x.

Reay, S. D., G. Collier, R. Douglas, N. Hayes, I. Nakarada-Kordic, A. Nair, and J. Kennedy-Good. 2017. "Prototyping Collaborative Relationships between Design and Healthcare Experts: Mapping the Patient Journey." *Design for Health* 1 (1): 65–79. doi:10.1080/24735132.2017.1294845.

Reid-McDermott, B., M. Browne, D. Byrne, P. O'Connor, E. O'Dowd, C. Walsh, C. Madden, and S. Lydon. 2019. "Using Simulation to Explore the Impact of Device Design on the Learning and Performance of Peripheral Intravenous Cannulation." *Advances in Simulation* 4 (1): 27. doi:10.1186/s41077-019-0118-5.

Ribeiro, M. L., H. M. Lederman, S. Elias, and F. L. S. Nunes. 2016. "Techniques and Devices Used in Palpation Simulation with Haptic Feedback." *ACM Computing Surveys* 49 (3): 1–48. doi:10.1145/2962723.

Rieber, L. P., and D. Noah. 2008. "Games, Simulations, and Visual Metaphors in Education: Antagonism between Enjoyment and Learning." *Educational Media International* 45 (2): 77–92. doi:10.1080/09523980802107096.

Rowe, A., M. Knox, and G. Harvey. 2020. "Re-Thinking Health through Design: Collaborations in Research, Education and Practice." *Design for Health* 4 (3): 327–344. doi:10.1080/24735132.2020.1841918.

Scerbo, M. W., and S. Dawson. 2007. "High Fidelity, High Performance?" *Simulation in Healthcare* 2 (4): 224–230. doi:10.1097/SIH.0b013e31815c25f1.

Steinert, M., and L. J. Leifer. 2012. "'Finding One's Way': Re-Discovering a Hunter-Gatherer Model Based on Wayfaring." *International Journal of Engineering Education* 28 (2): 251.

Stokes-Parish, J. B., R. Duvivier, and B. Jolly. 2020. “How Does Moulage Contribute to Medical Students’ Perceived Engagement in Simulation? A Mixed-Methods Pilot Study.” *Advances in Simulation* 5 (1): 23. doi:10.1186/s41077-020-00142-0.

Takeuchi, H., and I. Nonaka. 1986. “Stop Running the Relay Race and Take up Rugby.” *Harvard Business Review* 64 (1): 137–146.

Thomke, S., and D. Reinertsen. 1998. “Agile Product Development: Managing Development Flexibility in Uncertain Environments.” *California Management Review* 41 (1): 8–30. doi:10.2307/41165973.

Ullrich, S., and T. Kuhlen. 2012. “Haptic Palpation for Medical Simulation in Virtual Environments.” *IEEE Trans Vis Comput Graph* 18 (4): 617–625. doi:10.1109/TVCG.2012.46.

Wilson, E., D. G. Hewett, B. C. Jolly, S. Janssens, and M. M. Beckman. 2018. “Is That Realistic? The Development of a Realism Assessment Questionnaire and Its Application in Appraising Three Simulators for a Gynaecology Procedure.” *Advances in Simulation* 3 (1): 21. doi:10.1186/s41077-018-0080-7.

Yakubo, S., Y. Kinoshita, T. Aki, and H. Ota. 2008. “Improvement of a Simulator Production Project for Abdominal Palpation in Kampo Medical Training.” *Kampo Medicine* 59 (4): 595–600. doi:10.3937/kampomed.59.595.

Zenios, S., J. Makower, P. Yock, T. J. Brinton, U. N. Kumar, L. Denend, and T. M. Krummel. 2009. *Biodesign: The Process of Innovating Medical Technologies*. 1st ed. Cambridge, UK; New York, NY: Cambridge University Press.

Zhang, J., T. R. Johnson, V. L. Patel, D. L. Paige, and T. Kubose. 2003. “Using Usability Heuristics to Evaluate Patient Safety of Medical Devices.” *Journal of Biomedical Informatics* 36 (1–2): 23–30. doi:10.1016/S1532-0464(03)00060-1.

C7

Perception by Palpation: Development and Testing of a Haptic Ferrogranular Jamming Surface

Frontiers in Robotics and AI

Sigurd Bjarne Rørvik, Marius Auflem, Henrikke Dybvik, Martin Steinert (2021),

frontiers in Robotics and AI

ORIGINAL RESEARCH
published: 28 September 2021
doi: 10.3389/frobt.2021.745234

Perception by Palpation: Development and Testing of a Haptic Ferrogranular Jamming Surface

Sigurd Bjarne Rørvik, Marius Auflem*, Henrikke Dybvik and Martin Steinert

TrollLABS, Department of Mechanical and Industrial Engineering, Faculty of Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

Tactile hands-only training is particularly important for medical palpation. Generally, equipment for palpation training is expensive, static, or provides too few study cases to practice on. We have therefore developed a novel haptic surface concept for palpation training, using ferrogranular jamming. The concept's design consists of a tactile field spanning 260 x 160 mm, and uses ferromagnetic granules to alter shape, position, and hardness of palpable irregularities. Granules are enclosed in a compliant vacuum-sealed chamber connected to a pneumatic system. A variety of geometric shapes (output) can be obtained by manipulating and arranging granules with permanent magnets. The tactile hardness of the palpable output can be controlled by adjusting the chamber's vacuum level. A psychophysical experiment (N = 28) investigated how people interact with the palpable surface and evaluated the proposed concept. Untrained participants characterized irregularities with different position, form, and hardness through palpation, and their performance was evaluated. A baseline (no irregularity) was compared to three irregularity conditions: two circular shapes with different hardness (Hard Lump and Soft Lump), and an Annulus shape. 100% of participants correctly identified an irregularity in the three irregularity conditions, whereas 78.6% correctly identified baseline. Overall agreement between participants was high (κ= 0.723). The Intersection over Union (IoU) for participants sketched outline over the actual shape was IoU *Mdn* = 79.3% for Soft Lump, IoU *Mdn* = 68.8% for Annulus, and IoU *Mdn* = 76.7% for Hard Lump. The distance from actual to drawn center was *Mdn* = 6.4 mm (Soft Lump), *Mdn* = 5.3 mm (Annulus), and Mdn = 7.4 mm (Hard Lump), which are small distances compared to the size of the field. The participants subjectively evaluated Soft Lump to be significantly softer than Hard Lump and Annulus. Moreover, 71% of participants thought they improved their palpation skills throughout the experiment. Together, these results show that the concept can render irregularities with different position, form, and hardness, and that users are able to locate and characterize these through palpation. Participants experienced an improvement in palpation skills throughout the experiment, which indicates the concepts feasibility as a palpation training device.

Keywords: haptic interface, tactile surface, simulation, palpation, granular jamming, tactile perception, ferromagnetic granules

Edited by:
Carlos Rossa,
Carleton University, Canada

Reviewed by:
Nicolas Herzig,
University of Sussex, United Kingdom
Jay Carriere,
University of Calgary, Canada
Min Li,
Xi'an Jiaotong University, China

****Correspondence:***
Marius Auflem
mariuauf@stud.ntnu.no

Specialty section:
This article was submitted to Haptics, a section of the journal Frontiers in Robotics and AI

Received: *21 July 2021*
Accepted: *09 September 2021*
Published: *28 September 2021*

Citation:
Rørvik SB, Auflem M, Dybvik H and Steinert M (2021) Perception by Palpation: Development and Testing of a Haptic Ferrogranular Jamming Surface. Front. Robot. AI 8:745234. doi: 10.3389/frobt.2021.745234

1 INTRODUCTION

In simulated training environments (i.e., augmented, virtual, and mixed reality), realistic rendering of tactile interactions with the physical world is challenging, yet meaningful. This is because haptic interfaces enabling such tactile interactions must complement (and reflect) the vivid audiovisual feedback provided by the simulation (Woodrum et al., 2006). This combination could yield deeper immersion and thus facilitate the transfer of tactile experiences when transitioning to real-world scenarios. Furthermore, by realistically bridging the physical and digital world, users can develop, improve, and maintain critical psychomotor skills (Lathan et al., 2002; Zhou et al., 2012; Zhao et al., 2020). Hence, haptic interfaces in simulation can enable safe, repetitive, and available training alternatives for various professions that require dexterous hands-on experience (Carruth, 2017; Lelevé et al., 2020).

In a medical context, simulation can help narrow the gap of required clinical experience and mitigate the risk of harming or providing unsatisfactory patient treatment. However, various medical procedures require not only hands-on, but hands-only training. One of these procedures is palpation, which is used to examine a patient through touch. By palpation, diagnosis is based on tactile findings such as irregularities (lumps, fluids, tenderness) and locating pain-points. Unfortunately, common equipment such as wearable tactile devices and kinesthetic devices are less suited in this use-case given their current resolution, Degrees of Freedom (DOF), and tactile limitations (Licona et al., 2020). Consequently, simulated palpation exercises are mainly performed using static case-specific models (phantoms) or mannequins (patient simulators). While these can provide safe and repetitive training conditions, their fixed number of study cases, task-specific functionalities, and limited tactile realism are collectively obstacles for current healthcare training and education.

Haptic interfaces designed for palpation training should enable users to practice locating and describing tactile irregularities, as they would when palpating a real patient. Hence, multiple tactile displays are promising in this context by utilizing technology ranging from pin arrays (Wagner et al., 2002), to shape memory alloy actuators (Taylor et al., 1998) and airborne ultrasound (Iwamoto et al., 2008). However, such solutions are generally expensive, complex in operation and non-continuously available, thus limiting their use and widespread in research and education. Moreover, as these solutions rely on using a matrix of actuators or tactile outputs, it restricts the obtainable resolution, scalability and robustness of such interfaces. Furthermore, compliance and flexibility are often compromised by using rigid mechanisms to achieve haptic feedback. Therefore, attention has been brought to using soft robotics principles for haptic applications, as these can approximate soft body animations and organic behaviors suitable to medical training, among others (Manti et al., 2016).

An interesting area of soft robotics for medical training applications is the use of granular jamming mechanisms for haptic feedback. Granular jamming enables interfaces to alter stiffness and thus simulate compliant objects with variable hardness. This technology has been explored in medical training devices as embedded tactile modules (He et al., 2021), multi-fingered palpation interfaces (Li et al., 2014), and as actuation to enable objects and surfaces to alter shape and hardness for palpation (Stanley et al., 2016; Koehler et al., 2020). While this technology looks promising, current solutions often require complex pneumatic systems, since a matrix of actuated cells or objects is needed. Thus, this could limit the tactile resolution and geometrical freedom of rendered objects. Based on this existing work on granular jamming interfaces, we have developed a simple and low-cost technology utilizing ferromagnetic granulate. Our technology enables the granules to be remotely manipulated in an unjammed state and thus create customized tactile objects. Furthermore, when jammed, the hardness of these objects can be altered by the applied vacuum, i.e., how firmly the granules are packed together in a sealed chamber. In a haptic interface prototype described in **Figure 1**, the ferrogranular jamming principle is used to render palpable irregularities between two compliant layers. The prototype was developed to examine the feasibility and usability of this technology in a tactile display application. Moreover, this technology could be used to challenge the complexity, accessibility and cost of current haptic interfaces.

This work relates to the existing literature on tactile interactions, and more precisely, users' tactile perception of hardness and geometrical shapes. Hence, studies investigating the psychophysical perception of hardness and shapes have been of interest (Tan et al., 1992; Srinivasan and LaMotte, 1995; Bergmann Tiest and Kappers, 2009; Frisoli et al., 2011). However, the use-case of palpable interfaces that requires a perceptual exploration and manipulation is a less explored area with fewer examples (Lederman and Klatzky, 1993; Genecov et al., 2014). As this encourages more research on users' interaction and performance using haptic interfaces, our conceptual prototype has been piloted in a palpation experiment. This experiment investigates whether untrained users can locate and determine the form and hardness of rendered irregularities by palpation. Information of hardness, speed (time used to find irregularity) and accuracy of form and position has been collected, together with users' subjective experience throughout the experiment.

This paper examines using soft-robotics principles to alter the characteristics of a haptic interface for medical diagnostics training. This investigation has resulted in the concept shown in **Figure 1**, which uses granular jamming and ferromagnetic granulate manipulation to achieve various palpable outputs. The concept is used to assess untrained users' ability to locate and characterize the shape and hardness of different irregularities using palpation. Considering this concept for a novel haptic interface and the context of medical palpation training, we try to answer the following research questions in this paper:

i. Can the novel ferrogranular jamming concept be used as a haptic interface for palpation exercises?
ii. How well can untrained users determine the position, form and hardness of irregularities rendered by the haptic interface using palpation techniques?

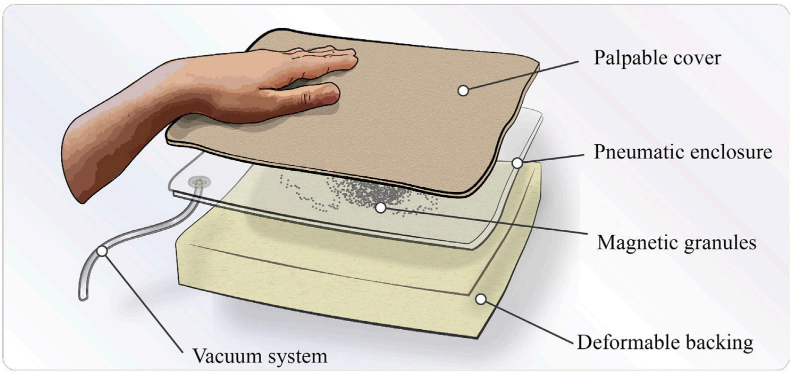


FIGURE 1 | Descriptive illustration of the haptic interface concept.

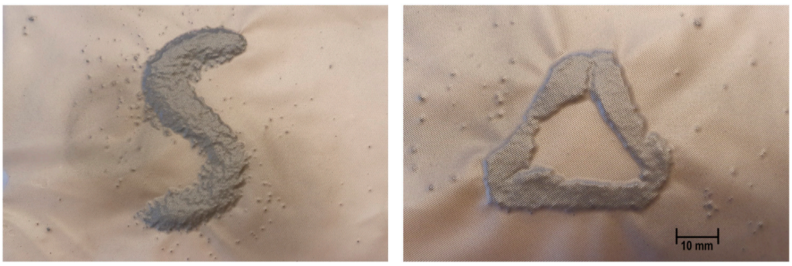


FIGURE 2 | Pictures of two arrangement possibilities.

iii. Did participants think their palpation skills improved during the experiment?

2 MATERIALS: DESIGN OF THE FERROGRANULAR JAMMING INTERFACE

This chapter starts with a short introduction to the ferrogranular jammer. Secondly, the theory of granular jamming and magnetic manipulation is presented. Lastly, the manufacturing of the magnetic granules and chamber is presented before the pneumatic setup.

The prototype was developed to examine the feasibility and usability of a ferrogranular jamming interface in a tactile display for palpation. The novelty of the proposed concept is the introduction of magnetic manipulation of granules in a jamming application. This innovation provides the opportunity to manipulate the granular media inside a compliant vacuum chamber, thus managing the position, form and hardness of the palpable outputs. Some examples are shown in **Figure 2**, where the jammed granulate shapes are visible within the translucent chamber. To act as a deformable and palpable structure the vacuum chamber is sandwiched between a deformable polyurethane (PU) foam backing (60 mm) and a flexible polyethylene (PE) fabric cover (4 mm) (as seen in **Figure 5B**).

2.1 Granular Jamming and Magnetic Manipulation

Granular jamming works by transitioning granular matter from a low-density compliant packing to a high-density rigid packing. This change is done by removing the fluid/medium surrounding the granulate, which produces an external hydrostatic pressure. From this, the granules can behave both like a fluid and a solid. When the granules are in a low-density packing, the intergranular friction is low, resulting in a fluid-like state. Vice versa, when the vacuum level increases, higher intergranular friction results in a jammed and solid-like state. In the jammed state, the granules distributes applied force through the grains so that the group of particles functions as a stiff and compliant material (Cates et al., 1998).

Particle jamming has been a big research topic for engineers and material scientists for the last few decades. The principle of reversibly transitioning the granular media from a fluid-like state to a more rigid state has been seen to be applicable to various domains, such as industrial grippers (Harada et al., 2016; D'Avella et al., 2020), minimally invasive surgery (Jiang et al., 2012) and robotic locomotion (Steltz et al., 2009). Granular jamming is a prevalent type of actuation within soft robotics applications because of two main reasons: 1) considerable stiffness variation with little volume change, and 2) possibility to adjust the stiffness variability area so it can be easily adapted to different soft robotics applications (Fitzgerald et al., 2020).

There has been research on optimizing granules for granular jamming with different aspects; size, shape and volume fraction (Jiang et al., 2012), chamber material (Jiang et al., 2012) and using soft granules (Putzu, Konstantinova, and Althoefer 2019). However, a common feature for these studies is the stasis of the granulate. To the best of our knowledge, there has been no research focusing on the movability of granules in a jamming context. For example, Follmer et al. (2012) reviewed jamming in a user-interface context, where none of the technologies utilized movement of the granules.

Using magnetic fields is an effective way to transport and position magnetic particles in a medium. The most prominent concept of ferromagnetic particles in a fluid is ferrofluid. This colloidal liquid consists of surfactant-coated magnetic particles with a size order of 10 nm suspended in a liquid medium. When the fluid is subjected to a magnetic field, it forms a shape like the magnetic field and acts more like a solid. Generally, ferromagnetic particles are induced by two types of interaction energy: the one between the particles and the magnetic field E^H, and between particles E^M (Cao et al., 2014). Using a magnetic field to manipulate magnetic particles has been used in microfluidic systems, such as magnetorheological fluid in user interfaces (Hook et al., 2009; Jansen et al., 2010) and biological analysis and catalysis (Gijs et al., 2010).

The advantages of using ferromagnetic granules include: 1) Controllability—Ferro-granulate can be arranged numerous ways by designing magnetic fields. 2) Noncontact—Magnetic particles can be remotely manipulated. 3) Precision—Ferromagnetic granules can be placed at a target region with high precision by precisely designing a magnetic field with local maximum field strength at preferred areas (Cao et al., 2014).

2.2 Manufacturing of Magnetic Granulate

Based on the previous research done on granular jamming, manufacturing of ferromagnetic granules to be used in a haptic interface were investigated. A central factor for the granulate in this research is how high interparticle friction yields higher viscosity in the un-jammed state but yields higher hardness when jammed and vice versa. Since moving the granules in the unjammed state is essential, we investigated the granule material and manufacturing methods that produce granules with lower interparticle friction in the unjammed state but still yielding sufficient hardness in the jammed state.

Ground coffee, which Putzu, Konstantinova, and Althoefer (2019) refer to as the gold standard within the field of granular jamming, was evaluated as the most viable option for our case. Ground coffee has been proven to be a successful granulate for jammers that need a large stiffness range (Brown et al., 2010; Cheng et al., 2012). The magnetic coffee ground was produced by mixing fine coffee ground and magnetic paint with a 1:1 volumetric ratio as seen in **Figure 3** (Magnetic undercoat, Lefranc and Bourgeois Déco). After the mixture dried, it was ground to a size of approximately 2 mm using a mortar. Using a crushing technique, instead of grinding, produced less size dispersion of the granulate. Granules with a 1–2.4 mm size were filtered out with a perforated filter with circular holes (see **Figures 3D,E**). It is advantageous to use homogeneous monodisperse granules to make the output more repeatable (Genecov et al., 2014).

The manipulation of the ferromagnetic granulate using a permanent magnet is presented in **Figure 4**. The same type of spikes can be observed in both ferrofluids and iron shavings when in the presence of a magnetic field.

2.3 Chamber Design

Since the concept of this technology is different from traditional granular jamming, the choice of chamber material was evaluated on having surface friction that enabled the granules to be remotely manipulated inside the sealed chamber. Further, the material needed to be flexible to jam the particles together when a vacuum was applied. Different heat-sealing plastic types were evaluated, and a corrugated polyvinylchloride (PVC) film (0.2 mm for vacuum sealing applications) was deemed the most viable due to its flexibility and least warping lines. With the corrugated pattern, we avoided self-sealing as this was a problem with other materials.

2.4 Pneumatic Setup

The pneumatic setup for the ferrogranular jamming concept is shown in **Figure 5**. The chamber is connected to the rest of the pneumatic system through a filter (**Figure 5D**). The 12 V vacuum pump (D2028B, SparkFun Electronics) delivers a vacuum level down to −0.54 bar. Next, a manometer is connected to measure the vacuum level. The vacuum pump is controlled using a speed controller. The chamber was made using an Impulse Heat Sealer (Audion Elektro Sealboy 235). A 3D-printed nozzle connects the chamber to the rest of the system, as seen in **Figure 5E**. Together with butyl vacuum sealant tape, it ensures minimal leakage at the inlet. A ball valve connects the system to atmospheric pressure when open.

3 METHOD: EXPERIMENT

A psychophysical experiment was designed to evaluate the functional abilities of the proposed concept by evaluating the user's performance in locating and characterizing rendered irregularities. The experiment encompassed a palpation task, where qualitative and quantitative data were gathered on both participant performance and prototype reliability.

3.1 Experimental Test Setup

The pneumatic system presented in 2.4 was integrated into the test cabinet shown in **Figure 6A**. A camera is fixed above the

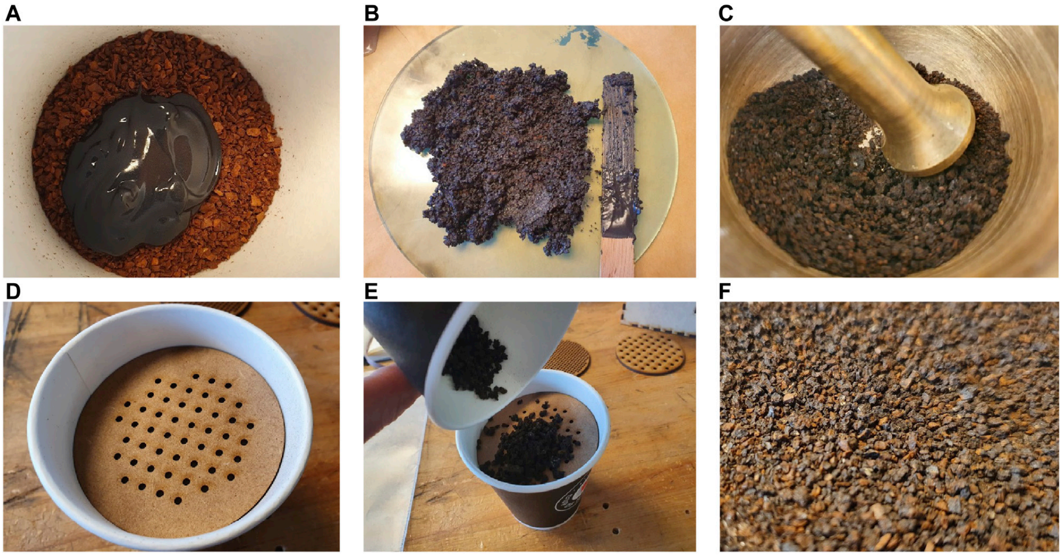


FIGURE 3 | Manufacturing of magnetic granulate **(A)** 1:1 mixing ratio of coffee ground and magnetic paint **(B)** Consistency of the mixture **(C)** Grounding using mortar **(D)** and **(E)** Filtering **(F)** Finished result.

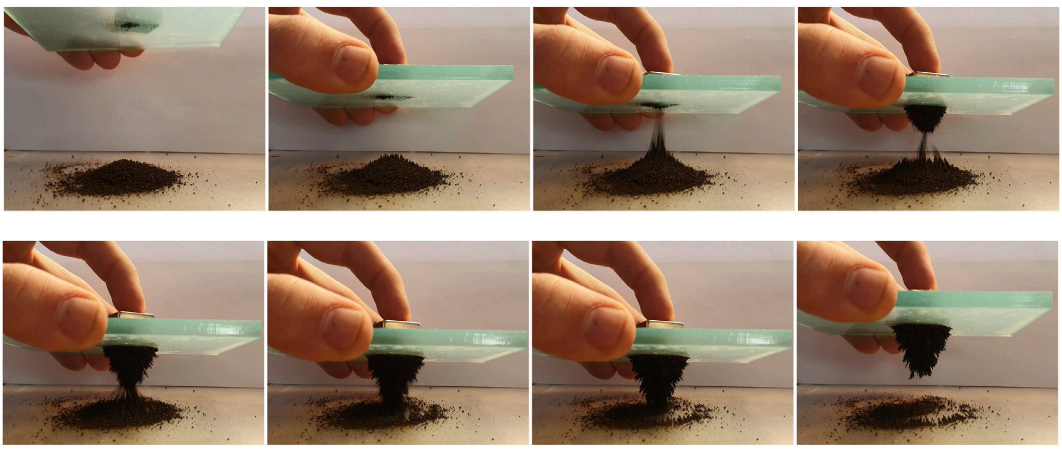


FIGURE 4 | Manipulation of the magnetic coffee ground using a permanent magnet.

haptic interface. The cabinet walls ensure no bias from visual perception during the transition between conditions and provides a consistent working environment. In addition, an overhead LED panel eases picture processing by ensuring consistent lighting. The two different geometrical shapes were created with two arrangements of permanent neodymium magnets, as seen in **Figure 6B**. These magnets were held above the vacuum chamber, arranging the granules in the desired shape, before applying the vacuum. When vacuum was applied, the magnets could be removed and the granulate remained jammed in place. To alter the shape, or remove it, the vacuum was released, before the granules were manually dispersed, rearranged, or moved out of the palpable field. The structural parts of the test rig are laser-cut MDF. The palpable field (260 × 160 mm) is seen as the pink

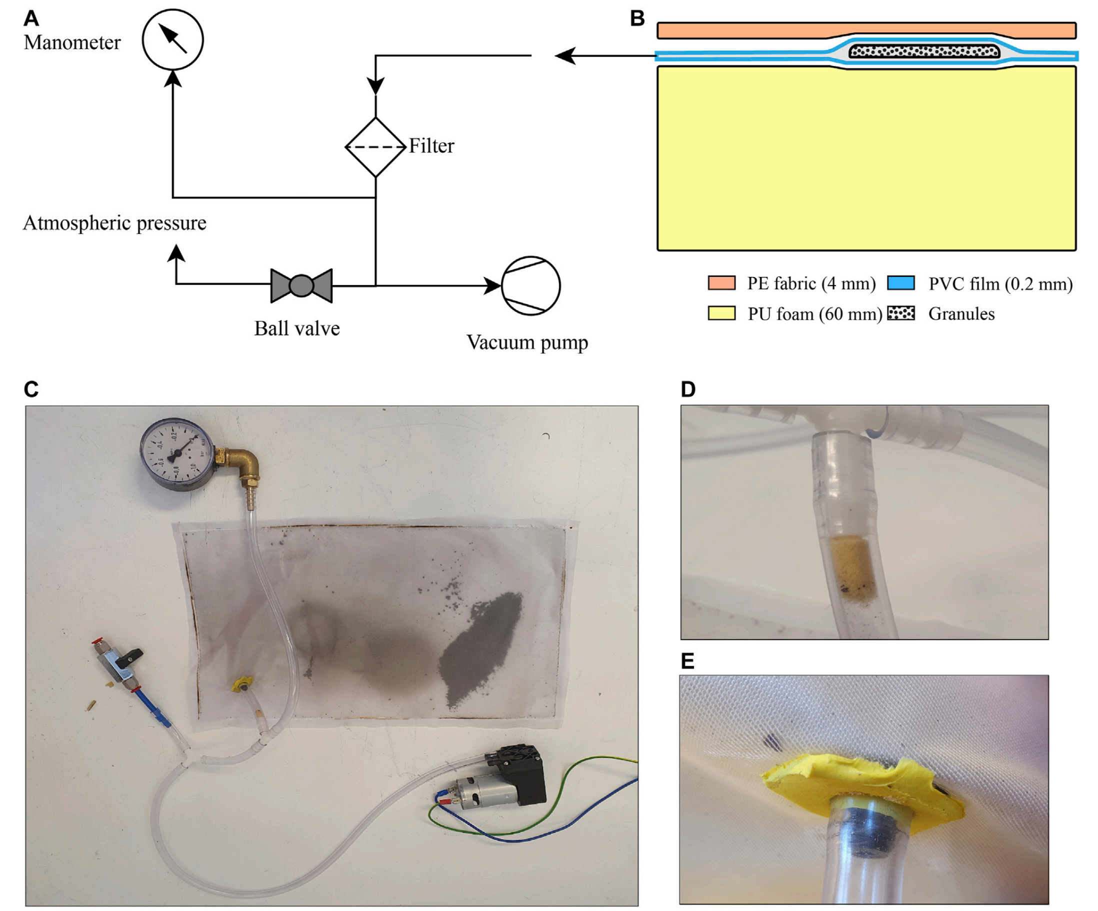


FIGURE 5 | (A) Schematic presentation of the pneumatic setup **(B)** Palpation interface with layer material and thickness **(C)** The pneumatic setup **(D)** Filter **(E)** Inlet seal for vacuum chamber.

area in **Figure 6A**. We used 12 g of filtered ferromagnetic granulate in the chamber.

3.2 Experiment Design

All participants repeated the palpation task four times, under four different conditions. The irregularity could differ in hardness, position and form. The four conditions were as follows:

- C1: Baseline. No irregularity in the palpation field.
- C2: Annulus. Annular-shaped irregularity rendered with the magnet configuration seen in **Figure 6B**. Vacuum level: −0.4 to −0.6 bar, whereas −1 bar is a complete vacuum. Located in the lower left part of the field. Approximately 82 mm outer diameter and 29 mm inner diameter with area M = 4,915 mm^2 SD = 371 mm^2 SE = 70 mm^2.
- C3: Hard Lump. A circular-shaped irregularity rendered with the magnet configuration seen in **Figure 5B**. Vacuum level: 0.4 to −0.6 bar. Located in the top right part of the field. Approximately 100 mm diameter with area M = 7,912 mm^2 SD = 474 mm^2 SE = 89 mm^2.
- C4: Soft Lump. A circular-shaped irregularity rendered with the magnet configuration seen in **Figure 5C**. Vacuum level: 0.1 bar. Located in the top right part of the field. Approximately 100 mm diameter with area M = 8,094 mm^2 SD = 641 mm^2 SE = 121 mm^2.

The sequence of the testing conditions was randomized to avoid potential learning or order effects. The order of conditions was also balanced, i.e., they appear the same number of times in each procedure step.

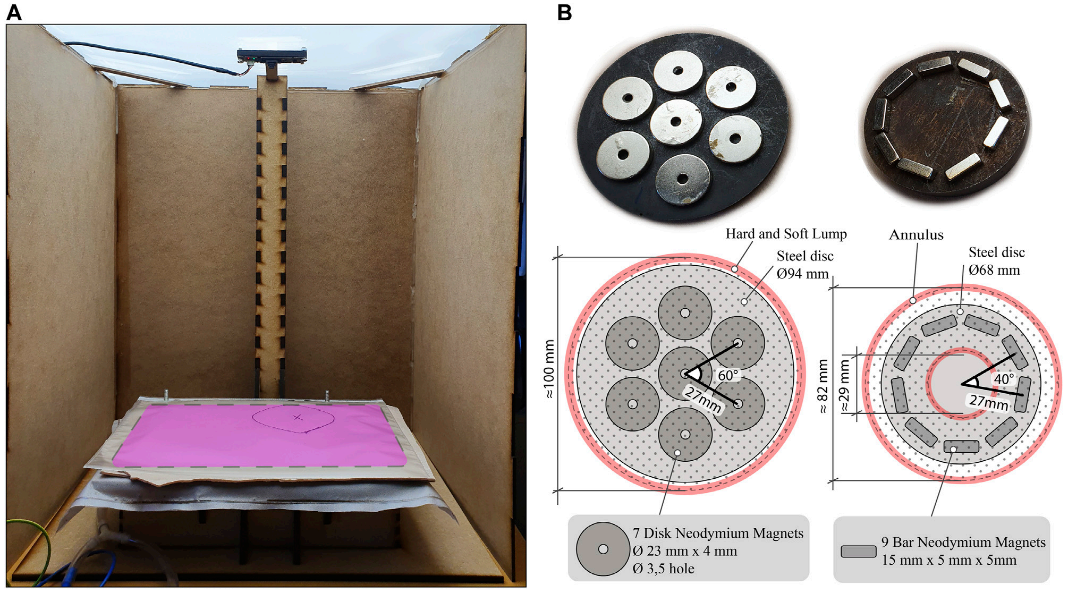


FIGURE 6 | (A) Test rig with camera and cabinet setup (pink area is the palpable field) **(B)** Magnet arrangement with angular and radial distance shown. Approximated outlines for the generated outputs are also illustrated with measurements. (Left) Circular shape for Hard and Soft Lump (Right) Annulus shape with the hollow center.

3.2.1 Participants

N = 28 healthy engineering students were recruited to participate (21 male (75%) and 7 female (25%)). Twenty-seven participants were in the 21–29 years range and one participant in the 18–20 years range. None of the participants were trained in the test or had any relevant knowledge about the technology before participation. Participation was voluntary, and all gave informed consent to be part of the study.

3.2.2 Experimental Procedure and Data Collection

The experimental procedure can be seen in **Figure 7**. After signing a consent form, the participants filled out a demographic questionnaire. Durometer and manometer readings and pictures of the granulate were sampled before the participant was seated in front of the test setup. The hardness of the irregularities was measured with a commercially available Shore durometer (Shenzhen Gairan Tech Co., X.F Type 00), following the requirements described in ISO 48–4:2018. A minimum of three measurements at different positions on the flat parts of the irregularity was performed. After objective data was collected, participants were instructed regarding the proceedings of the experiment. First, participants were told to palpate for a potential irregularity and say stop when they had control of the position and form. The participants did not get any instructions regarding technique to be used, other than using their hands to explore and feel for any irregularities in the field. We measured the time the participant used to find position and form of the irregularity. After completing each palpation, participants were asked to draw the contours of a possible irregularity on a sheet placed above the palpation field. More specifically, they were told to draw the outside and possible inside contours and put an x inside the area enclosing the proposedly identified irregularity (see **Figure 8B**). Pilot experiments showed that this instruction facilitated the participants who found an inside contour to also draw it, instead of drawing the outer contour only. Drawing data were captured using the camera.

To evaluate the hardness of the irregularity, a sampled selection of objects of varying hardness was used. These samples were numbered from 1 to 5 and had a Shore hardness of 00–20, 00–35, 00–55, 00–65, and 00–90, from soft to harder. The objects were presented similarly to the test setup using the same deformable backing and palpable cover as the palpation field. Thus, the participant could palpate the irregularity when doing the hardness test.

After each condition, participants reported their degree of agreement to a series of statements using a Likert Scale from 1 (Totally disagree) to 5 (Totally agree). The statements were: 1) It was hard to find the irregularity. 2) I am confident that I found the position and shape of the irregularity. 3) The irregularity had a constant/homogeneous hardness. To get a measure of a potential learning effect occurring during the experiments, participants also evaluated the statement: 4) I became better at finding the irregularity during the experiment., after completing the experiment.

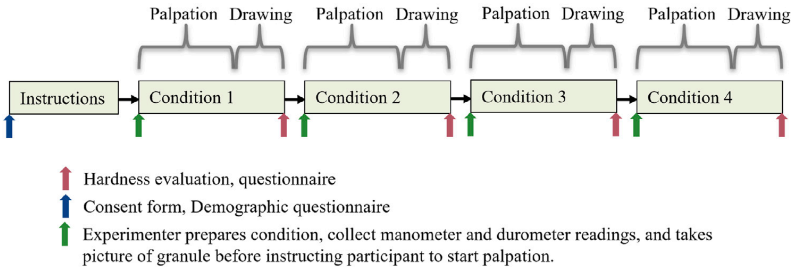


FIGURE 7 | Experimental procedure timeline. The order of the conditions was randomized.

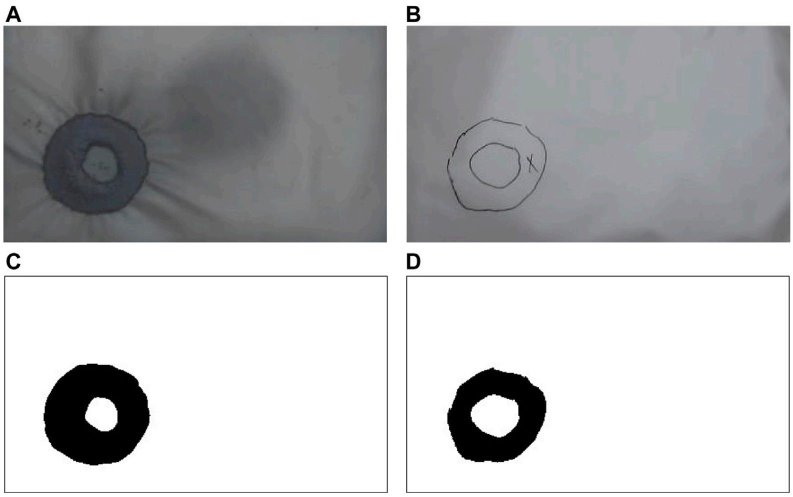


FIGURE 8 | Unprocessed images of **(A)** Granulate and **(B)** Drawing. Binarized pictures of **(C)** Granulate and **(D)** Drawing.

3.3 Data Analysis

The data was collected throughout the experiment to answer the study's research questions. Thus, experiment pictures were processed into binarized matrices that yielded objective data points describing irregularities' and drawings' respective positions and geometrical form. These data, together with the questionnaire and objective measurements, were statistically analyzed for reliability and differences between variables with SPSS Statistics (IBM SPSS Statistics 27, 2020).

3.3.1 Picture Processing

Data about position and form was collected through images. The images of the granulate and drawings were then processed and analyzed using package OpenCV 4.5.1 in Python 3.0. The capturing code also took photos of the manometer during each test. The images were blurred before grayscaling and binarizing to remove noise. An adaptive Gaussian threshold was used on the pictures of the drawings to improve accuracy. The binarized results are shown in **Figures 8C,D**.

3.3.1.1 Distance From Center to Center

The center point distance between granulate and drawings were calculated by finding the center of mass for both the granulate and the drawings using cv2.moments in Python. Then, the Euclidean distance (ΔD) was calculated between the two coordinates, using **Eq. 1**. x_1 and y_1 representing the coordinates for the granulate, while x_2 and y_2 representing the drawing.

$$\Delta D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{1}$$

3.3.1.2 Intersection Over Union

Intersection over Union (IoU) was used to evaluate the form. First, matrices of the intersection and union of the two binarized pictures were calculated using Python. Then, the number of black

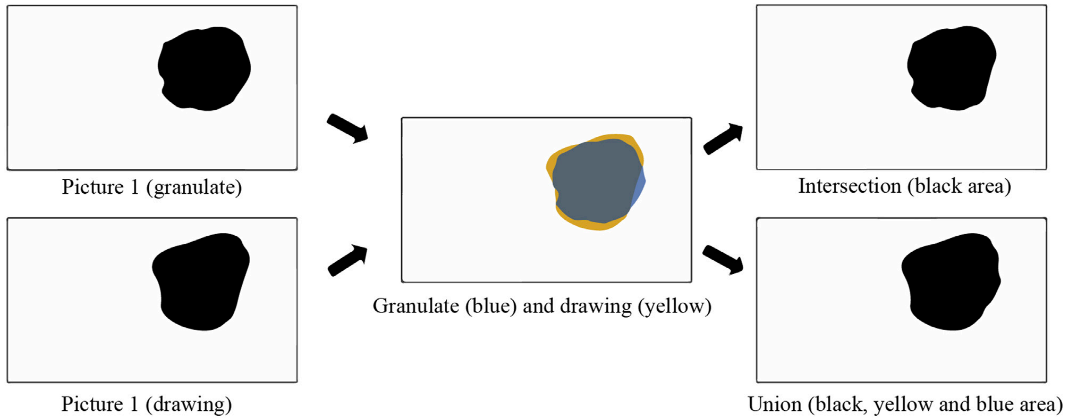


FIGURE 9 | Visual presentation of intersection and union.

pixels (pixels with value 0) in the intersection was divided by the number of black pixels in the union, using **Eq. 2**. A visual representation of intersection and union is shown in **Figure 9**.

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union} \quad (2)$$

3.4 Statistical Tests

To assess reliability, Fleiss' kappa was ran to determine if there was an agreement between participants' judgment of whether there was an irregularity or not (Lump or No Lump) in the four conditions. Fleiss' kappa does not assume that the raters are identical for each condition (which is the case here), but this is the only test we know of that assesses the case when there are multiple raters. Therefore, we report this test along with the frequency. One-way repeated measures ANOVAs were used to investigate differences between conditions for continuous variables. Those were: IoU, hardness (durometer reading) and vacuum level (manometer reading). Assumptions regarding no outliers, normality, and sphericity were inspected with boxplots, histograms and Normal Q-Q Plots, and Mauchly's Test of Sphericity. Violations of the outlier assumption were not removed since it only applied to Durometer and Manometer readings, which were used to corroborate that the conditions Hard Lump and Soft Lump differed in terms of hardness. In addition, a Friedman test was also conducted to ensure similar differences. A Greenhouse-Geisser correction was applied in the case of violating sphericity (Wickens and Keppel, 2004; Field, 2018). A Friedman test was used to investigate differences between conditions for discontinuous variables (the remaining variables), and in the case of more severe violations to ANOVA's assumptions. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons for both ANOVA and Friedman. Some variables produced a statistically significant Friedman test, but without any significant pairwise comparisons. One reason might be the conservative nature of the multiple comparisons correction. An additional approach, multiple Wilcoxon signed-rank tests, was therefore used to follow up the Friedman tests. We deemed it acceptable to be less conservative since it is the first investigation of an early-stage prototype, and it was important to gain an understanding of where potential differences were. The Wilcoxon signed-rank tests was also used to obtain a z-score, used to estimate effect size (r) (Rosenthal, 1986; Field, 2018). For the ANOVAs, the sample effect size partial eta squared (η^2), and population effect size partial omega squared (ω^2) (Rosenthal, 1986) are reported. The significance level $p < 0.05$ was chosen for highly significant differences. p–values ≤ 0.10 were considered as interesting effects, again due to the experiment involving human participants evaluating an early-stage prototype. We believe a 10% probability for Type 1 error is acceptable in this case.

4 RESULTS

Both objective and subjective data points were gathered throughout all four conditions described in 3.2.2. Each condition focused on localizing and characterizing a potential irregularity based on position, form and hardness. Additional descriptive statistics can be found in **Supplementary Material**.

4.1 Lump or No Lump: How Many Found an Irregularity?

In all three conditions with an irregularity (Annulus, Hard Lump and Soft Lump), all participants found an irregularity (100%

agreement). In Baseline condition six participants (21.4%) found an irregularity, despite there not being one. The remaining 22 participants (78.6%) failed to find an irregularity. Fleiss' kappa determined if there was an agreement between participants' judgment of whether there was an irregularity or not (Lump or No Lump) in the four conditions. The agreement between participants' judgements was statistically significant with κ= 0.723, 95% CI [0.722, 0.725], $p < 0.001$. The individual kappa's for Lump and No Lump categories were also κ= 0.723, 95% CI [0.722, 0.725], $p < 0.001$. This statistic is the proportion of agreement over and above chance agreement, with 0 being no agreement and 1 being perfect agreement. An agreement of 0.723 can be classified as a good agreement (Landis and Koch, 1977).

As stated, six out of 28 participants found an irregularity in the baseline condition. Of these six, three participants drew contours with areas of 22, 64 and 147 mm^2, which are small compared to the actual size of the irregularities. They are similar to granular remnants, which means they could be discarded as an error in the setup. Other participants commented on particle-sized irregularities in the Baseline condition but decided that they were not of sufficient size to be an actual irregularity. Removing these three participants results in three participants (12.0%) finding an irregularity in the Baseline condition, whereas 22 participants (88%) did not find an irregularity. Fleiss' kappa was ran again with these three participants removed to investigate the magnitude of the potential error from the setup. The agreement between the remaining 25 participants was statistically significant with κ= 0.840, 95% CI [0.783, 0.896], $p < 0.001$. The individual kappa's for Lump and No Lump categories were also κ= 0.840, 95% CI [0.783, 0.896], $p < 0.001$.

22 of 28 (78.57%) of the participants found the inner circle. In the two irregularity conditions 50 of 56 (89.29%) drawings were filled circles without any inner contour.

In summary, all participants agreed that there was an irregularity present in all irregularity conditions. Despite a few participants finding an irregularity where there was none, the overall agreement between participants was high.

4.2 Time

The users were not instructed to be as fast as possible but rather spend enough time to be sure of position and form of the irregularity. Therefore, the time represents the procedure time needed to find position and form of the irregularity to the best of the participant's ability.

Time was statistically significantly different in the four conditions, $\chi^2(3) = 29.460$, $p < 0.001$ as shown in **Figure 10**. Post hoc analysis revealed significant differences in Time from Baseline (Mdn = 37.0 s), 95% CI [25.0, 61.0] to Annulus (Mdn = 13.50 s), 95% CI [7.0, 21.0] ($p < 0.001$) and Soft Lump (Mdn = 16.00), 95% CI [11.0, 40.0] ($p = 0.001$) condition. It took longer to determine that there was no irregularity in Baseline condition, compared to finding it in Annulus and Soft Lump condition. The contrast comparing Annulus to Hard Lump (Mdn = 17.50), achieved a significance level $p = 0.050$ and effect size r = −0.50, and Hard Lump to Baseline had a significance level of $p = 0.067$ and effect size r = 0.48. We interpret this to be a notable difference. There was no significant difference between Annulus and Soft Lump (p = 0.80, r = −0.28) and Soft Lump and Hard Lump (p = 1.00, r = 0.22).

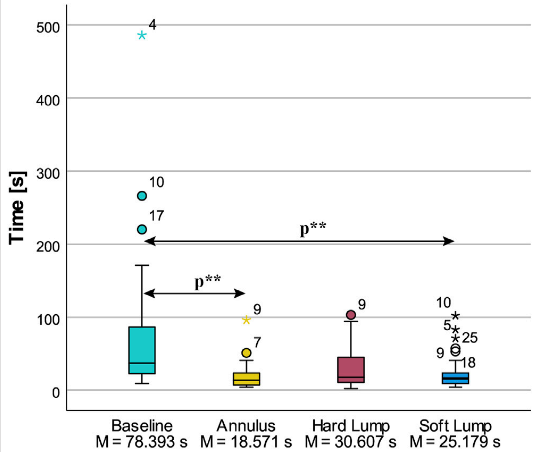


FIGURE 10 | Descriptive statistics of procedure time to find the position and form of the irregularities. Statistically significant differences at $p < 0.05$ are indicated by p^{**}.

4.3 Position: Distance From Center to Center

The distance between the center of the irregularity to the center of the participants' drawing was statistically significantly different in the three irregularity conditions, χ^2 (2) = 16.357, $p < 0.001$. Post hoc analysis revealed statistically significant differences in center distance from Annulus (Mdn = 5.2920 mm), 95% CI [2.978, 7.097], to Hard Lump (Mdn = 7.4366 mm), 95% CI [6.331, 12.217] ($p < 0.001$), and from Soft Lump (Mdn = 6.3908 mm), 95% CI [3.836, 9.654], to Hard Lump condition (*Mdn* = 7.4366 mm) ($p = 0.01$). There was no significant difference in center distance between Annulus and Soft Lump ($p = 1$). We also observe that there was a greater spread in the Hard Lump condition. These results are plotted in **Figure 11A**.

4.4 Form: IoU

IoU was statistically significantly different in the three conditions, χ^2 (2) = 12.071, $p = 0.002$. Post hoc pairwise comparisons with a Bonferroni correction for multiple comparisons yielded one significant difference between Soft Lump (Mdn = 0.793), 95% CI [0.728, 0.825], and Annulus (Mdn = 0.688), 95% CI [0.579, 0.735], $p = 0.002$, and a corrected $p = 0.247$ for both the Soft Lump vs Hard Lump (Mdn = 0.767), 95% CI [0.568, 0.754] comparison, and Hard Lump vs Annulus comparison (uncorrected p-value was p = 0.082). Post hoc Wilcoxon tests revealed a statistically significant difference between Annulus (Mdn = 0.688) and Soft Lump (Mdn = 0.793), $T = 316.00$, $p = 0.010$, r = 0.49, and a significant difference between Hard Lump (Mdn = 0.767), 95%

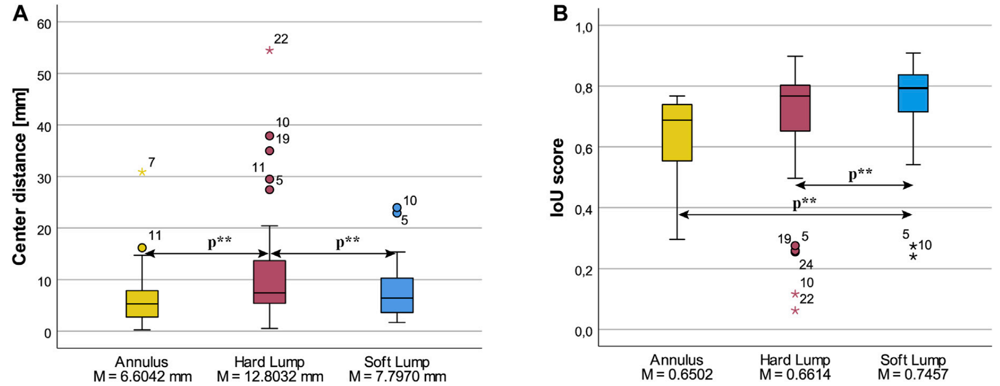


FIGURE 11 | Box plots of **(A)** Center distance of granulate and drawing, and **(B)** Intersection over Union, for the three lump conditions. Statistically significant differences at $p < 0.05$ are indicated by p^{**}.

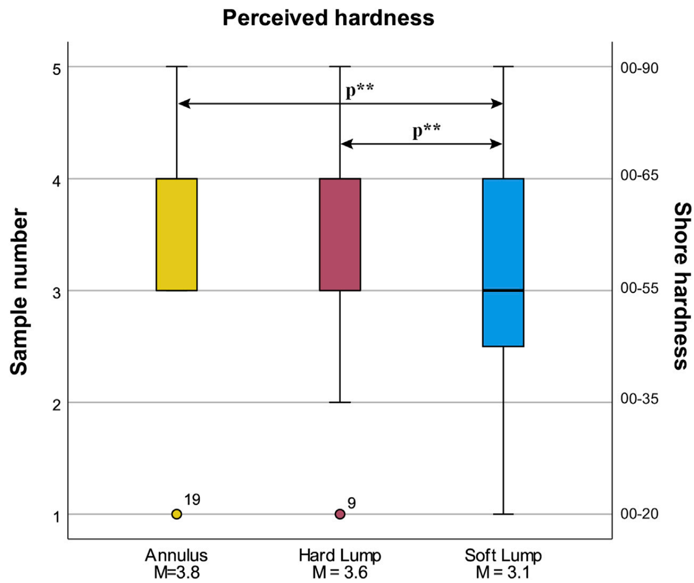


FIGURE 12 | Box plot of perceived hardness (subjective) for the three lump conditions. Statistically significant differences at $p < 0.05$ are indicated by p^{**}.

CI [0.703, 0.0794], and Soft Lump (Mdn = 0.793), $T = 304.00$, $p = 0.021$, $r = 0.43$. There was no difference between Annulus and Hard Lump, $T = 263.00$, $p = 0.021$, $r = 0.26$ as plotted in **Figure 11B**.

4.5 Hardness

We compared perceived hardness, objective hardness measurements, and vacuum levels of the irregularity conditions.

4.5.1 Perceived Hardness

Perceived hardness was statistically significantly different in the three conditions, χ^2 (2) = 9.129, $p = 0.010$. Post hoc pairwise comparisons with a Bonferroni correction for multiple comparisons yielded one significant difference between Soft Lump (Mdn = 3) and Annulus (Mdn = 4), $p = 0.033$, and a corrected $p = 0.184$ for the Soft Lump and Hard Lump comparison (uncorrected p-value was $p = 0.061$). Post hoc Wilcoxon tests revealed a statistically significant difference

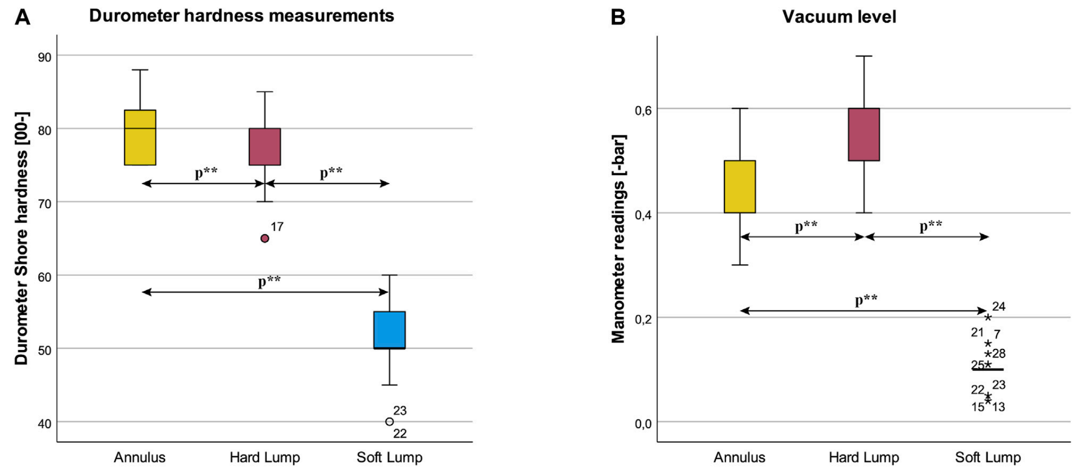


FIGURE 13 | Box plots of **(A)** Durometer hardness measurements and **(B)** vacuum level for the three lump conditions. Statistically significant differences at $p < 0.05$ are indicated by p^{**}.

between Soft Lump (Mdn = 3) and Hard Lump (Mdn = 4), T = 132.00, p = 0. 032, r = –0.41, and a significant difference between Annulus (Mdn = 4) and Soft Lump (Mdn = 3), T = 55.00, p = 0.015, r = –0.46. There was no difference between Annulus and Hard Lump, T = 87.00, p = 0. 474, r = –0.14. These results are as expected. Participants perceived the hardness of Soft Lump to be less than that of both Hard Lump and Annulus which is shown in **Figure 12**.

4.5.2 Durometric Measurements

There were 3 outliers as assessed by boxplot in **Figure 13A**. By visual inspection, the data was approximately normally distributed. The assumption of sphericity was violated, as assessed by Mauchly's test of sphericity, $\chi^2(2)$ = 6.825, p = 0.033. Therefore, a Greenhouse-Geisser correction was applied (ε = 0. 812). Results was statistically significant different in the three conditions F (1.625, 43.872) = 278.699, p < 0.001, η^2 = 0.912, ω^2 = 0.869. Durometric readings were: Annulus (M = 79.96), Hard Lump (M = 76.79), Soft Lump (M = 51.25). Post hoc analysis with a Bonferroni correction yielded statistically significantly difference between Annulus and Hard Lump (M = 3.179, 95% CI [0.63, 5.72], p = 0.011), between Annulus and Soft Lump (M = 0.28.714, 95% CI [24.71, 32.72], p < 0.001), and between Hard Lump and Soft Lump (M = 25.536, 95% CI [22.04, 0.29.03], p < 0.001).

4.5.3 Manometer

There were several outliers as assessed by boxplot in **Figure 13B**. The data was approximately normally distributed by visual inspection. The assumption of sphericity was violated, as assessed by Mauchly's test of sphericity, $\chi^2(2)$ = 13.713, p = 0.001. Therefore, a Greenhouse-Geisser correction was applied (ε = 0. 709). Manometer was statistically significant different in the three conditions F (1.419, 38.301) = 228.636, p <0.001, η^2 = 0.894, ω^2 = 0.844. Manometer readings were: Annulus (M = 0.431), Hard Lump (M = 0.536), Soft Lump (M = 0.101). Post hoc analysis with a Bonferroni adjustment was statistically significantly different between Annulus and Hard Lump (M = –0.105, 95% CI [–0.17, –0.04], p = 0.002), between Annulus and Soft Lump (M = 0.330, 95% CI [0.29, 0.368], p < 0.001), and between Hard Lump and Soft Lump (M = 0.435, 95% CI [0.38, 0.49], p <0.001).

4.6 Questionnaire

Participants completed the questionnaire in the three irregularity conditions.

4.6.1 How Hard Was It to Find the Position?

Participants' evaluation of how hard it was to find the irregularity was statistically significantly different in the three conditions, $\chi^2(2)$ = 7.423, p = 0.024. Post hoc pairwise comparisons with a Bonferroni correction for multiple comparisons yielded no significant differences. Post hoc Wilcoxon tests revealed a statistically significant difference between Soft Lump (Mdn = 1) and Hard Lump (Mdn = 1), T = 63.00, p = 0.006, r = –0.52. There were no significant differences between Annulus (Mdn = 1) and Soft Lump (Mdn = 1), T = 27.00, p = 0.957, r = –0.01, or between Annulus and Hard Lump, T = 89.00, p = 0.096, r = –0.32. Participants found it hardest to locate the irregularity in the Hard Lump (see **Figure 14A**).

4.6.2 Confidence in Finding Position and Shape of the Irregularity

Participants' confidence in finding position and shape of the irregularity was statistically significantly different in the three irregularity conditions, $\chi^2(2)$ = 8.926, p = 0.012. Post hoc pairwise comparisons with a Bonferroni correction for multiple

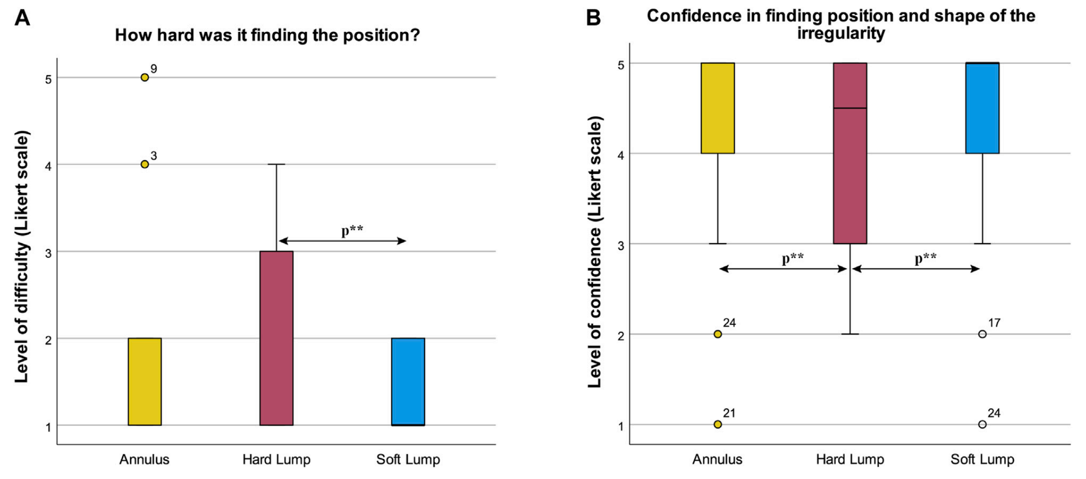


FIGURE 14 | Box plot of results from the questionnaire **(A)** How hard was it finding the position, and **(B)** Self-assessment of confidence in the accuracy of their drawing.

comparisons yielded no significant differences. Post hoc Wilcoxon tests revealed a statistically significant decrease in confidence from Annulus (Mdn = 5) to Hard Lump (Mdn = 4.5), $T = 12.00$, $p = 0.032$, r = –0.40, and a significant decrease in confidence from Soft Lump (Mdn = 5) to Hard Lump (Mdn = 4.5), $T = 13.50$, $p = 0.038$, $r = -0.39$. There was no significant difference between Annulus and Soft Lump, $T = 31.00$, $p = 0.276$, $r = -0.21$. Participants were most confident in finding the position and shape of the irregularity in the Annulus and Soft Lump condition and less confident in finding the irregularity's position and shape in the Hard Lump condition (see **Figure 14B**).

4.6.3 Homogeneous Hardness

Participants' evaluation of whether the irregularity had a constant/homogeneous hardness was not significantly different in the three irregularity conditions, χ^2 (2) = 3.410, $p = 0.182$.

4.6.4 Self-Assessed Improvement in Palpation

After completing the experiment, participants evaluated whether they thought they improved their palpation skills throughout the experiment. 20 participants (71.4%) thought they improved, 3 disagreed (10.7%), and 5 were neutral (17.9%).

5 DISCUSSION

Introducing ferromagnetic granules in a jamming haptic interface has the quality to be a promising solution to produce larger tactile displays cheaply with high accuracy. Palpation trainers need to be robust, safe and have a high level of repeatability. Using adaptable palpation trainers increases the number of study cases and task-specific functionalities the trainer can accomplish. Thus, we think our concept can be taken further for use in a medical training equipment environment. Before that, however, there is a need for further development of the technology and contextual testing.

When comparing our data with relevant research (as mentioned in the Introduction), we have good results for people's perception of both hardness and position. Bergmann Tiest and Kappers (2009) states that users are pretty good at determining hardness. Frisoli et al. (2011) states that cutaneous sensor modality is not affected by size, but kinesthetic performance is reduced with smaller-sized objects. To our knowledge, there is a lack of research on people's perceptual exploration and characterization. Thus, this study could add to the body of knowledge concerning this aspect of both machine interaction and human tactile perception. The following section discusses the objective and subjective results gathered and how they answer our research questions. Further, an evaluation of the participant sampling is presented before we discuss the limitations and outlook of the study.

5.1 Interpreting Results

We defined that our haptic interface should be able to change the position, form and hardness of an irregularity recognizable by palpation. We chose a circle and an annulus as our two shapes to evaluate if people could locate and characterize them. A major part of the participants could differentiate between the circular lumps and the Annulus (78.57% recognized the inner circle of the Annulus, and 89.29% drew the circular lumps with no inner circle). Furthermore, all the participants found an irregularity in all conditions that had an irregularity. For the Baseline condition, six participants found a false positive. When determining the participants' ability to describe the form, we used Intersection over Union (IoU). The median was promising for all three irregularity conditions, with the highest score for Soft Lump (0.793). From these three observations, we could conclude that the overall agreement between participants was great for form. Thus, our concept can manipulate the granules into different shapes that laypersons can distinguish by palpation.

However, an interesting result is that the IoU was significantly lower for Annulus than for Hard Lump and Soft Lump, while Annulus scored best at center point distance. A more logical assumption would be that IoU and center point distance is inversely proportional. There could be at least two reasons why we get a lower IoU score for Annulus. Initially, we observed from the participants contouring the Annulus that they struggled to get the size and position of the inner circle right. Due to how we calculated IoU, a wrong positioned hole yielded a more considerable difference in IoU than a similar error in outer contour. Also, the same error in the center point difference for Lump and Annulus gives a more significant change in IoU for Annulus because of the inner circle.

The results show a statistical difference for both objectively measured and perceived hardness between Soft lump and Hard Lump. Furthermore, when performing Wilcoxon tests on perceived hardness, there was a significantly lower value of the Soft Lump than the Hard Lump and Annulus condition. Thus, we have shown that participants can distinguish between hard and soft objects that the prototype produces, which is essential for palpation tasks, whereas characterizing the physical attributes such as form and hardness of identified irregularities is essential.

Considering how participants conducted the palpation tasks, time spent is of interest. The five most extended procedure times were on baseline condition, and three of them had baseline as the first condition. For example, participant No. 4 stated that there was no irregularity after 90 s before spending six more minutes to palpate before finding a false positive. No. 21 expressed hesitancy after 1 minute, and then spent two more minutes palpating before concluding with a true negative. No. 17 expressed insecurity before spending 2–3 more minutes searching, ending with a true negative. From the respective participant's confidence data, the participant with the false positive (no. 4) answer a four on the Likert scale, i.e., partially agreeing that they were sure they found the correct position and form of the irregularity. Also, all three participants had good results in all irregularity conditions. This connection could mean that some people struggle to trust their sense of touch. The baseline condition presents ambiguity as there is nothing to palpate, and we believe having it as the first condition increased ambiguity and thus uncertainty in participants who probably expected an irregularity.

Another aspect is the repeatability of our testing equipment. We tried to develop a haptic interface that can alter and maintain hardness, position and form of palpable outputs with high repeatability. However, while prototyping the granular jammer, it became apparent that repeatedly creating geometries with identical shape and hardness was challenging. While the outline for the shapes varied for each sample as a result of the manual setup and granule dispersion, the gathered images showed only a small deviation of rendered area for each irregularity condition. The durometric measurements did, however, show a wide hardness range within a prepared condition. This inhomogeneous hardness from granular jamming is similar to the findings of Genecov, Stanley, and Okamura (2014). We thought of two reasons for this, firstly, how the hardness is highly dependent on how the granules interlock or position themselves across the irregularity. Secondly, because the arrangement of granules was made manually, there was an unavoidable variation in the produced output geometries and thus granulate concentration across the area.

Considering the pressure readings for the setup, Soft Lump had more outliers as a result of the vacuum level being manually set (and adjusted). Compared to Hard Lump and Annulus which had a hard stop, governed by the maximum vacuum the setup could provide. Given these being different geometries thus yielding different volumes to drain the air from, this could cause the difference in obtained vacuum level. However, our results from the perceived hardness showed no significant difference between Annulus and Hard Lump. As this being a first prototype, challenges concerning repeatability is expected and the overall results show great promise for this concept to be improved further to address these limitations.

When looking at the questionnaire data, confidence was highest in finding the position and shape of the Annulus and Soft Lump condition. We expected that the Hard Lump would be the easiest to find due to a sharper edge and thus greater difference between the Hard Lump and the palpable area. However, this was not the case. In retrospect we suspect this was due to the increased vacuum level instantly jamming the granules not allowing them to evenly distribute and conform to a smooth shape. This could cause the edges of the Hard Lump to be more jagged/rivet than the edge of the Soft Lump, which had a more circular shape in comparison. We therefore believe this might have made participants more uncertain of where the edge was. Moreover, since the edge of the Hard Lump varied more compared to a circle, this may have contributed to that it was more difficult to find the center of it.

Participants reported they got better at finding the irregularity throughout the experiment, meaning it could be used as a training device for palpation exercises. However, a reported high level of confidence and low level of difficulty for each condition could mean the task being too easy to perform, not leaving much room for progression and learning. Given the ability to find and characterize the irregularities sufficiently, and having a high confidence in doing so, further steps should be made to tailor level of difficulty to specific scenarios and investigating the use of the device in a medical context. The concept has been experimentally, shown to facilitate users' palpation skills by speed, location, shape, and hardness differentiation of palpable findings. Other learning objectives could involve motoric technique and following procedural algorithms, which should be explored in further development of the concept.

5.2 Participant Sampling

In this experiment, the 28 participants were all engineering students who did not have any previous experience with palpation as a medical examination technique. The participants did not get any technical or strategic instructions, meaning their palpation approach would be different to a medical professional. Therefore, we have shown that our concept works for presenting generic geometric shapes for laypersons, which is a promising result considering this a training device. Moreover, having participants with prior medical experience, would thus require a higher level of difficulty. A sample size of 28 was adequate compared to similar studies (Gerling et al., 2003; Asgar-Deen et al., 2020). Also, as we got statistically significant differences between relevant data points, such as the hardness of Soft Lump and Hard Lump, more participants would most likely not produce other results. However, a higher sample size would reduce the possibility of an accepted hypothesis being incorrect.

5.3 Limitations and Further Work

In this research, sensing, automation or participant feedback has not been addressed nor implemented in the palpation concept. The prototype and subsequently the experiments with the prototype are not tied to a medical context. Instead, it explores some of the capabilities and extreme conditions the haptic device can output. Hence, it has not been within the scope of this research to model, synthesize, or simulate physiological attributes for palpation. Nevertheless, we lack to prove that our haptic interface is helpful in medical training because of a simplified experiment focusing on planar perceptual exploration. Therefore, in further development, more levels of difficulty, complex geometries, hardness profiles, locations, and dynamic abilities should be explored. As palpation tasks seldom concerns irregularities in one plane, investigating multiple jamming layers, or simulating depth of palpation by dynamic stiffness control should be investigated. A positive backpressure in combination with ferrogranular jamming could yield 3D-shapes with high tactile resolution and geometrical freedom (Koehler et al., 2020; He et al., 2021). In further work, we seek to test the concept with users who can provide feedback and evaluation on a medical basis. This could reveal hitherto unexplored concept potentials and critical functions to pursue.

6 CONCLUSION

This work has described the development and testing of a novel haptic interface concept that uses ferrogranular jamming. This concept was developed as a compliant simulation interface for medical palpation training, with the ability to simulate geometrical objects of various shapes and tactile properties. The concept was tested by having 28 untrained participants perform a set of structured palpation tasks in an experiment. The experiment consisted of four conditions, one baseline and three containing a palpable irregularity. These conditions were chosen to evaluate if the interface could produce various shapes and hardness levels, while also investigating participants' palpation skills. Given the results of the experiment, we conclude that the concept can create palpable objects with variable hardness by adjusting the jamming vacuum. Laypersons can distinguish these objects by palpation, both by the hardness, location, and shape of objects with good accuracy. Thus, this study also provides insights on peoples' perceptual abilities in explorative palpation. It shows the ability to locate and characterize palpable objects of varying shape and hardness in a satisfactory manner. Further, the results show that, the task was not considered very challenging. This combined with participants reported high level of confidence in performance, indicates that increased difficulty might be required to ensure room for improvement and learning. However, as participants also reported improvements in their palpation skills during the experiment, the technology looks promising to be further developed for medical training applications.

Considering this being an early conceptual prototype, this study revealed opportunities and challenges yet to be addressed. In further work, we want to explore whether the interface can be used as a palpation tool in medical simulation by qualitative testing with expert users. This will require palpable objects where both hardness, shape, and difficulty are tailored to the medical scenarios we want to simulate. Other technical aspects of the ferrogranular jamming concept we want to explore are sensing and feedback, automation, and dynamic and responsive tactile abilities. Collectively, this could improve the experience of using this technology in simulation-based medical training.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article is available in the following repository. SBR; MA; HD; and MS, 2021, Replication Data for: Perception by Palpation: Development and Testing of a Haptic Ferrogranular Jamming Surface, https://doi.org/10.18710/OCMXVP, DataverseNO, V1.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors contributed ideas and concepts throughout the development of the device. SR did the technical work and execution of the experiment. MS provided laboratory and testing facility. All authors contributed and agreed to the final wording of the submission. HD conducted the statistical analysis. MA contributed to the development of the concept and design of the experiment.

FUNDING

This research is supported by the Research Council of Norway through its industrial Ph.D. funding scheme, project number 290404.

ACKNOWLEDGMENTS

Thanks to Sindre Wold Eikevåg for the experiment procedure timeline figure and thanks to Håvard Vestad for creative ideas concerning the ferrogranular jamming concept.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frobt.2021.745234/full#supplementary-material.

REFERENCES

Asgar-Deen, D., Carriere, J., Wiebe, E., Peiris, L., Duha, A., and Tavakoli, M. (2020). Augmented Reality Guided Needle Biopsy of Soft Tissue: A Pilot Study. *Front. Robot. AI* 7, 72. doi:10.3389/frobt.2020.00072

Bergmann Tiest, W. M., and Kappers, A. (2009). Cues for Haptic Perception of Compliance. *IEEE Trans. Haptics* 2, 189–199. doi:10.1109/TOH.2009.16

Brown, E., Rodenberg, N., Amend, J., Mozeika, A., Steltz, E., Zakin, M. R., et al. (2010). Universal Robotic Gripper Based on the Jamming of Granular Material. *Proc. Natl. Acad. Sci.* 107, 18809–18814. doi:10.1073/pnas.1003250107

Cao, Q., Han, X., and Li, L. (2014). Configurations and Control of Magnetic fields for Manipulating Magnetic Particles in Microfluidic Applications: Magnet Systems and Manipulation Mechanisms. *Lab. Chip* 14, 2762–2777. doi:10.1039/c4lc00367e

Carruth, D. W. (2017). "Virtual Reality for Education and Workforce Training," in 2017 15th International Conference on Emerging eLearning Technologies and Applications (ICETA) (IEEE), 1–6. doi:10.1109/ICETA.2017.8102472

Cates, M. E., Wittmer, J. P., Bouchaud, J.-P., and Claudin, P. (1998). Jamming, Force Chains, and Fragile Matter. *Phys. Rev. Lett.* 81, 1841–1844. doi:10.1103/PhysRevLett.81.1841

Cheng, N. G., Lobovsky, M. B., Keating, S. J., Setapen, A. M., Gero, K. I., Hosoi, A. E., et al. (2012). "Design and Analysis of a Robust, Low-Cost, Highly Articulated Manipulator Enabled by Jamming of Granular media," in 2012 IEEE International Conference On Robotics And Automation (IEEE), 4328–4333. doi:10.1109/ICRA.2012.6225373

D'Avella, S., Tripicchio, P., and Avizzano, C. A. (2020). A Study on Picking Objects in Cluttered Environments: Exploiting Depth Features for a Custom Low-Cost Universal Jamming Gripper. *Robotics and Computer-Integrated Manufacturing* 63, 101888. doi:10.1016/j.rcim.2019.101888

Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics*. Los Angeles: SAGE.

Fitzgerald, S. G., Delaney, G. W., and Howard, D. (2020). A Review of Jamming Actuation in Soft Robotics. *Actuators* 9, 104. doi:10.3390/act9040104

Follmer, S., Leithinger, D., Olwal, A., Cheng, N., and Ishii, H. (2012). "Jamming User Interfaces," in Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology UIST '12 (New York, NY, USA: Association for Computing Machinery), 519–528. doi:10.1145/2380116.2380181

Frisoli, A., Solazzi, M., Reiner, M., and Bergamasco, M. (2011). The Contribution of Cutaneous and Kinesthetic Sensory Modalities in Haptic Perception of Orientation. *Brain Res. Bull.* 85, 260–266. doi:10.1016/j.brainresbull.2010.11.011

Genecov, A. M., Stanley, A. A., and Okamura, A. M. (2014). "Perception of a Haptic Jamming Display: Just Noticeable Differences in Stiffness and Geometry," in 2014 IEEE Haptics Symposium (HAPTICS) (IEEE), 333–338. doi:10.1109/HAPTICS.2014.6775477

Gerling, G. J., Weissman, A. M., Thomas, G. W., and Dove, E. L. (2003). Effectiveness of a Dynamic Breast Examination Training Model to Improve Clinical Breast Examination (CBE) Skills. *Cancer Detect. Prev.* 27, 451–456. doi:10.1016/j.cdp.2003.09.008

Gijs, M. A. M., Lacharme, F., and Lehmann, U. (2010). Microfluidic Applications of Magnetic Particles for Biological Analysis and Catalysis. *Chem. Rev.* 110, 1518–1563. doi:10.1021/cr9001929

Harada, K., Nagata, K., Rojas, J., Ramirez-Alpizar, I. G., Wan, W., Onda, H., et al. (2016). Proposal of a Shape Adaptive Gripper for Robotic Assembly Tasks. *Adv. Robotics* 30, 1186–1198. doi:10.1080/01691864.2016.1209431

He, L., Herzig, N., Lusignan, S. d., Scimeca, L., Maiolino, P., Iida, F., et al. (2021). An Abdominal Phantom with Tunable Stiffness Nodules and Force Sensing Capability for Palpation Training. *IEEE Trans. Robot.* 37, 1051–1064. doi:10.1109/TRO.2020.3043717

Hook, J., Taylor, S., Butler, A., Villar, N., and Izadi, S. (2009). "A Reconfigurable Ferromagnetic Input Device," in Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology UIST '09 (New York, NY, USA: Association for Computing Machinery), 51–54. doi:10.1145/1622176.1622186

IBM SPSS Statistics 27 (2020). *Downloading IBM SPSS Statistics 27*. NY: IBM Corp..

Iwamoto, T., Tatezono, M., and Shinoda, H. (2008). "Non-contact Method for Producing Tactile Sensation Using Airborne Ultrasound," in *Haptics: Perception, Devices and Scenarios Lecture Notes in Computer Science*. Editor M. Ferre (Berlin, Heidelberg: Springer), 504–513. doi:10.1007/978-3-540-69057-3_64

Jansen, Y., Karrer, T., and Borchers, J. (2010). "MudPad," in ACM International Conference on Interactive Tabletops and Surfaces ITS '10 (New York, NY, USA: Association for Computing Machinery), 11–14. doi:10.1145/1936652.1936655

Jiang, A., Xynogalas, G., Dasgupta, P., Althoefer, K., and Nanayakkara, T. (2012). "Design of a Variable Stiffness Flexible Manipulator with Composite Granular Jamming and Membrane Coupling," in 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IEEE), 2922–2927. doi:10.1109/iros.2012.6385696

Koehler, M., Usevitch, N. S., and Okamura, A. M. (2020). Model-Based Design of a Soft 3-D Haptic Shape Display. *IEEE Trans. Robot.* 36, 613–628. doi:10.1109/TRO.2020.2980114

Landis, J. R., and Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 159–174. doi:10.2307/2529310

Lathan, C. E., Tracey, M. R., Sebrechts, M. M., Clawson, D. M., and Higgins, G. A. (2002). "Using Virtual Environments as Training Simulators: Measuring Transfer," in *Handbook of Virtual Environments* (Boca Raton, FL: CRC Press).

Lederman, S. J., and Klatzky, R. L. (1993). Extracting Object Properties through Haptic Exploration. *Acta Psychologica* 84, 29–40. doi:10.1016/0001-6918(93)90070-8

Lelevé, A., McDaniel, T., and Rossa, C. (2020). Haptic Training Simulation. *Front. Virtual Real.* 1, 2. doi:10.3389/frvir.2020.00003

Li, M., Ranzani, T., Sareh, S., Seneviratne, L. D., Dasgupta, P., Wurdemann, H. A., et al. (2014). Multi-fingered Haptic Palpation Utilizing Granular Jamming Stiffness Feedback Actuators. *Smart Mater. Struct.* 23, 095007. doi:10.1088/0964-1726/23/9/095007

Licona, A. R., Liu, F., Pinzon, D., Torabi, A., Boulanger, P., Lelevé, A., et al. (2020). "Applications of Haptics in Medicine," in *Haptic Interfaces for Accessibility, Health, and Enhanced Quality of Life*. Editors T. McDaniel and S. Panchanathan (Cham: Springer International Publishing), 183–214. doi:10.1007/978-3-030-34230-2_7

Manti, M., Cacucciolo, V., and Cianchetti, M. (2016). Stiffening in Soft Robotics: A Review of the State of the Art. *IEEE Robot. Automat. Mag.* 23, 93–106. doi:10.1109/MRA.2016.2582718

Putzu, F., Konstantinova, J., and Althoefer, K. (2019). "Soft Particles for Granular Jamming," in Annual Conference towards Autonomous Robotic Systems (Springer), 65–74. doi:10.1007/978-3-030-25332-5_6

Rosenthal, R. (1986). "Meta-Analytic Procedures for Social Science Research," in *Educational Researcher* (Beverly Hills: Sage Publications), 14818–14820. doi:10.3102/0013189x015008018

Srinivasan, M. A., and LaMotte, R. H. (1995). Tactual Discrimination of Softness. *J. Neurophysiol.* 73, 88–101. doi:10.1152/jn.1995.73.1.88

Stanley, A. A., Hata, K., and Okamura, A. M. (2016). "Closed-loop Shape Control of a Haptic Jamming Deformable Surface," in 2016 IEEE International Conference on Robotics and Automation (ICRA) (IEEE), 2718–2724. doi:10.1109/icra.2016.7487433

Steltz, E., Mozeika, A., Rodenberg, N., Brown, E., and Jaeger, H. M. (2009). "JSEL: Jamming Skin Enabled Locomotion," in 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IEEE), 5672–5677. doi:10.1109/IROS.2009.5354790

Tan, H. Z., Pang, X. D., and Durlach, N. I. (1992). Manual Resolution of Length, Force, and Compliance. *Adv. Robotics* 42, 13–18. doi:10.5802/aif.1307

Taylor, P. M., Moser, A., and Creed, A. (1998). A Sixty-Four Element Tactile Display Using Shape Memory alloy Wires. *Displays* 18, 163–168. doi:10.1016/S0141-9382(98)00017-1

Wagner, C. R., Lederman, S. J., and Howe, R. D. (2002). "A Tactile Shape Display Using RC Servomotors," in Proceedings 10th Symposium on Haptic Interfaces for Virtual Environment and Tele Operator Systems. HAPTICS 2002 (IEEE), 354–355.

Wickens, T. D., and Keppel, G. (2004). *Design and Analysis: A Researcher's Handbook*. Upper Saddle River, NJ: Pearson Prentice-Hall.

Woodrum, D. T., Andreatta, P. B., Yellamanchilli, R. K., Feryus, L., Gauger, P. G., and Minter, R. M. (2006). Construct Validity of the LapSim Laparoscopic Surgical Simulator. *Am. J. Surg.* 191, 28–32. doi:10.1016/j.amjsurg.2005.10.018

Zhao, X., Zhu, Z., Cong, Y., Zhao, Y., Zhang, Y., and Wang, D. (2020). Haptic Rendering of Diverse Tool-Tissue Contact Constraints during Dental Implantation Procedures. *Front. Robot. AI* 7, 1. doi:10.3389/frobt.2020.00035

Zhou, M., Tse, S., Derevianko, A., Jones, D. B., Schwaitzberg, S. D., and Cao, C. G. L. (2012). Effect of Haptic Feedback in Laparoscopic Surgery Skill Acquisition. *Surg. Endosc.* 26, 1128–1134. doi:10.1007/s00464-011-2011-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

C8

Embedded Soft Inductive Sensors to Measure Arterial Expansion of Tubular Diameters in Vascular Phantoms

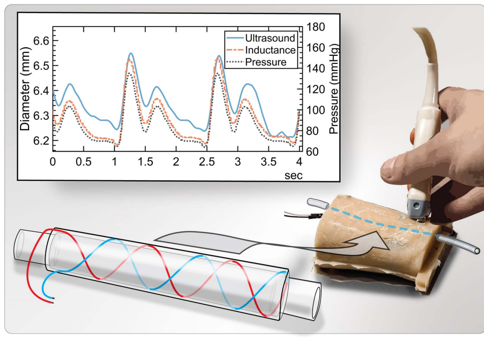
IEEE Sensors Journal

Torjus Steffensen, Marius Auflem, Håvard Vestad, Martin Steinert (2022),

Embedded Soft Inductive Sensors to Measure Arterial Expansion of Tubular Diameters in Vascular Phantoms

Torjus L. Steffensen, Marius Auflem, Håvard N. Vestad, *Member, IEEE*, and Martin Steinert

***Abstract*—Measuring diameter change in flexible tubular structures embedded in opaque material is challenging. In this article, we present a soft braided coil embedded in an elastomer tube as a method to continuously measure such a change in diameter. By measuring the inductance change in the braided coil, we estimate the instantaneous diameter with a simple inductance model. In applying this method, we demonstrate that diameter waves in a vascular phantom, a model of a radial artery embedded in a viscoelastic wrist structure, can be recorded continuously. Four sensors were made, and their ability to measure physiologically relevant simulated pulse waves was assessed. Several pressure pulse profiles were generated using a precision digital pump. Inductance of the coil was measured simultaneously as the change in diameter was recorded using an optical laser/mirror deflection measurement. One sensor was then embedded in a vascular phantom model of the human wrist. The diameter of the simulated radial artery was recorded via ultrasound and estimated from coil inductance measurements. The diameter estimates from the inductance model corresponded well with the comparator in both experimental setups. We demonstrate that our method is a viable alternative to ultrasound in recording diameter waves in artery models. This opens opportunities in empirical investigations of physiologically interesting fluid-structure interaction. This method can provide new ability to measure diameter changes in tubular systems where access is obstructed.**



***Index Terms*—In vitro experimentation, inductance, soft electronics, measurement methods.**

I. Introduction

MEASURING pressure propagation in soft tubular structures is of interest in the field of arterial mechanics, where pulse wave propagation can be studied to make conclusions about arterial function [1]–[3]. Many tissues including arteries are viscoelastic [4], making accurate mathematical modeling challenging. Therefore, empirical in vitro experiments are frequently employed [5], [6]. The relationship between fluid forces acting on the wall structure and the resulting change in geometry from the induced strains is one topic of interest. The propagation of the strain energy through the vessel wall can be visualized as waves of expanding and contracting vessel diameter. In vitro models are often considered as freely suspended tubes, for sake of simplicity and observation. However, in cases where the interaction between the model vessel and the surrounding tissue is of interest, this is not feasible. In the case of radial tonometry, a technology that is seeing interest for applications in wearable blood pressure monitoring [7]–[9], the interaction between the vessel wall, the underlying bone, and the tissue between the vessel and skin surface are all important. This can be achieved by placing the vessel model inside tissue-mimicking material, obstructing view of the tube [10].

There are few methods to continuously measure a varying diameter in a soft tube inside an opaque material. Here, we propose a soft braided coil cast inside an elastomer tube wall as such a method. To allow for radial expansion from pressurizing the tube, the sensor consists of a high pitch helical coil. The coil is routed back and forth over the length of the tube section, resulting in a weaved braid pattern, as opposed to a sequentially stacked coil (Fig. 1).

A change in the tube's diameter also changes the diameter of the coil, which in turn alters its inductance. This principle has previously been applied in sensing actuation lengths in

Manuscript received January 22, 2022; revised February 23, 2022; accepted February 23, 2022. Date of publication February 25, 2022; date of current version March 31, 2022. The work of Marius Auflem was supported by the Research Council of Norway under Grant 290404. The associate editor coordinating the review of this article and approving it for publication was Dr. Edward Sazonov. *(Corresponding author: Torjus L. Steffensen.)*

Torjus L. Steffensen is with the Department of Circulation and Medical Imaging, Norwegian University of Science and Technology, 7025 Trondheim, Norway (e-mail: torjus.l.steffensen@ ntnu.no).

Marius Auflem, Håvard N. Vestad, and Martin Steinert are with the Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, 7025 Trondheim, Norway.

Digital Object Identifier 10.1109/JSEN.2022.3155071

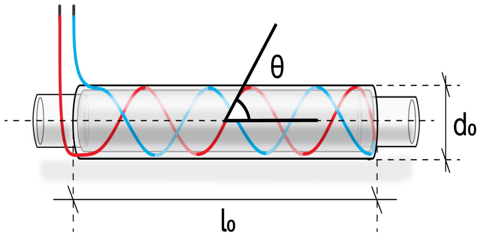


Fig. 1. Geometry of the sensor tube, showing the braided coil structure and connecting pressure tubes.

pneumatic actuators (McKibben muscles) [11], [12]. Similar flexible coils have also been used to determine angular change from the induced current resulting from the change in magnetic field caused by bending the coil [13]. Here, we focus on small, sub-millimeter changes in diameter caused by differences in internal pressure.

A. Detection Principle

With a fixed wire length, a braid pattern causes axial contraction of the helical structure from altered braid angles as the diameter increases. As the wire is routed in the same rotational direction around the tube the electromagnetic properties of the solenoid and current direction are preserved. We can then estimate the relationship between the cross-sectional area of the coil and the inductance L as in a long solenoid [11]:

$$L = \frac{\mu N^2 A}{l} \tag{1}$$

where N is the number of effective coil windings, μ is the magnetic permeability of the core (for air, we set μ approximately equal to vacuum permeability μ_0), and A is the cross-sectional area of the "solenoid". For a winding angle of 20° the model is roughly linear to the full extent of expansion and compression as limited by the braid [11]. In this work we operate well within this range.

The accuracy of the long solenoid model in braided soft actuators has been compared to more sophisticated inductance models [11], and while accuracy decreases with deformation size, for small deformations the error is small.

For a braided tube of length l, relaxed diameter D_0 and wire winding angle θ with respect to the long axis of the braid, assuming constant coil length and a circular cross section the length of the coil helix b is given by

$$b = \frac{l_e}{\cos(\theta)} \tag{2}$$

where l_e is the length of the fully relaxed tube. The number of turns each helix of the braid makes around the axis of the cylinder n is given by:

$$n = \frac{b \sin(\theta_0)}{\pi D_0} \tag{3}$$

TABLE I
TUBE DIMENSIONS

Tube no.	D_0	l_e	θ_0
1	6.2 mm	90 mm	25°
2	6.8 mm	58 mm	23°
3	5.3 mm	100 mm	18°
4	7.2 mm	75 mm	25°

Equation (3) along with the number of helices in the sensor gives us a value for the effective number of coil windings N [11], [14]. To determine a change in tube radius r from an associated change in inductance, we expand (1):

$$\Delta L = (L - L_0) = \frac{\mu N^2 \pi (r^2 - r_0^2)}{l} \tag{4}$$

Which is equivalent to:

$$r = \sqrt{\frac{(L - L_0) l}{\mu N^2 \pi} + r_0^2} \tag{5}$$

In a material with a positive Poisson's ratio the change in diameter necessarily also results in a contraction along the length of the tube. In the case of small deformations in a long thin tube, where $l \gg D$, we assume that the change in cylinder length is negligible in relation to the change in diameter and treat l as constant. The resulting relationship is nonlinear, but locally behaves approximately linearly for small deformations [12].

B. Sensor Fabrication

We made sensors by winding flexible silicone shielded wire onto a dowel. The resulting braids were then molded in Ecoflex 00-10 (Smooth-On, USA) by pouring and continuous rotation. Ecoflex 00-10 is a very soft elastomer, with a reported static Young's modulus of 0.05 MPa at 10% strain [15]. Removing the dowel, the resulting structure is a braided flexible wire embedded in a silicone tube of roughly uniform thickness.

We made four sensors to compare their behavior and to account for differences in coil characteristics due to the artisanal nature of the manufacturing process (Table I).

II. EXPERIMENTS & METHODS

A. Pressure Expansion

We conducted an experiment to assess the change in sensor induction due to expansion under pressure (Fig. 2). The sensor tube was fixed horizontally to an inclined plane, and the ends of the tube were fixed to prevent angular deformation during expansion and contraction.

The ends of the wire braids in the tube sensor were connected to a parallel capacitor, forming an LC oscillator circuit. This circuit was connected to an induction-to-digital converter (LDC1612, Texas Instruments, Texas, USA). The LDC1612 operates by applying a drive current to an LC circuit

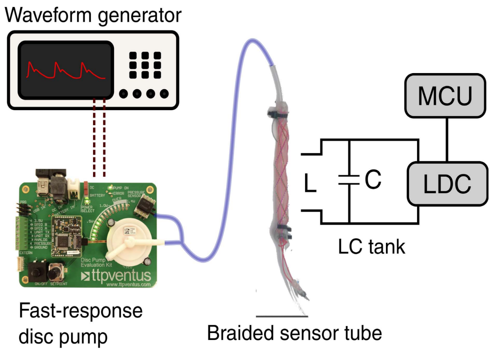


Fig. 2. Schematic illustration of the experimental setup.

and measuring the resulting primary oscillation frequency f_0. In our configuration the LDC was connected to an external 40 MHz reference oscillator, which is used to determine f_0. Given that the parallel capacitance C is known, this allows for the determination of the inductance L through the relationship

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \tag{6}$$

Each of the four tubes was in turn connected to a precision digital control disc pump (XP-S2-028, TTP Ventus, UK). The other end of the tube was clamped shut, resulting in a controllable internal pressure when the pump was active. A differential pressure sensor (TSC 015PD, Honeywell International, USA) gauged the air pressure inside the tube against ambient atmosphere.

As the diameter of the tube changes in response to the pressure variation, as does the diameter of the coil, resulting in an increase in L. Thus f_0 varies over the course of the load cycle, allowing a calculation of L from (6).

The disc pump was controlled with an arbitrary waveform generator (UTG2025A, UNI-T, China). Inductance and pressure were recorded at three different load states (Table II). First, after setting up the experiment, one minute of inductance readings was recorded to establish L_0 for each tube in the unloaded state. Second, a 0.5 Hz square wave pressure cycle was applied to the tube for a period of 4 minutes. Pressure was then allowed to equalize to ambient pressure. Lastly an arterial pulse pressure waveform was applied for 1 minute. The wave profile was recorded from the proper palmar digital artery at the middle phalanx of the left middle finger using a volume-clamp apparatus (NIBP Nano INL382, ADInstruments, Dunedin, New Zealand).

Data was collected on a Windows PC and processed in MATLAB r2021a (The Mathworks, Massachusetts, USA).

B. Response Validation

To record the change in tube diameter, the tube was fixed to an angled plate and restricted in a single point with cyanoacrylate glue, allowing it to expand and retract with the center of the tube moving normally to the angled plate.

TABLE II
EXPERIMENTAL PROTOCOL

Load type	Cycle Frequency	Pressure range	Duration
Zero load	n/a	Ambient	1 minute
Square wave	0.5 Hz	30 - 90 mmHg	4 minutes
Arterial wave	1 Hz	40 – 100 mmHg	1 minute

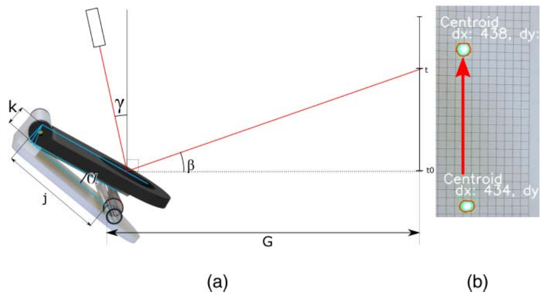


Fig. 3. (a) Schematic of laser measurement geometry used to measure expansion and contraction of the tubes. (b) Illustration of the laser dot tracking. Not to scale.

A mirror, rotating freely on a hinge, was placed resting on top of the tube. The expansion of the tube then resulted in a small angular change of the mirror, proportional to the increase in diameter of the tube (Fig. 3). A laser was focused on the mirror at an angle, and the resulting laser dot was recorded moving against a reference scale placed at a distance to amplify the movement [16].

The diameter of the tube could then be determined from the motion of the laser dot through a trigonometric relationship:

$$r \cong \tan\left(\frac{\tan^{-1}\left(\frac{t_1}{G}\right) + \gamma + 2\alpha_0}{4}\right) \times \left(j + k\tan\left(\frac{\tan^{-1}\left(\frac{t_1}{G}\right) + \gamma + 2\alpha_0}{4}\right)\right) \tag{7}$$

where r is the radius, t is the height of the laser dot on the target relative to the height of the mirror, k is the vertical distance from the hinge point to the baseplate where the tube is fixed (8.5mm), j is the lateral distance from hinge point to fixation point of tube (40mm), G is the horizontal distance between mirror and target (9980mm), and α_0 is the angle of the mirror hinge where the resulting angle between mirror and the horizontal plane is 45 degrees (7.5 degrees). γ is the angular deviation of the laser from the vertical axis. This is tuned with an adjustment screw on the laser to hit the target regardless of the different diameters of the tubes. γ is calculated using (7) and a caliper measurement of the diameter of the tube.

We recorded the movement of the laser dot with a digital camera at a resolution of 3840×2160 pixels. Frame to frame position was determined using open-source video analysis

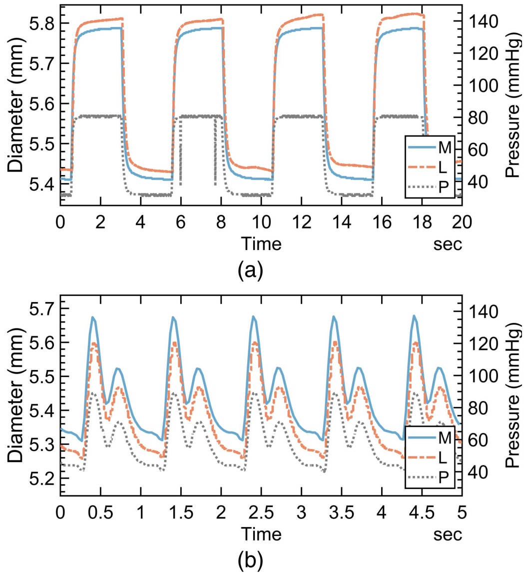


Fig. 4. Change in tube diameter from applied internal pressure, measured via mirror deflection, and estimated from inductance. M: diameter measured via mirror deflection, L: diameter estimate from inductance measurement, P: internal tube pressure. (a) 0.5 Hz square wave. (b) 1 Hz radial pulse wave.

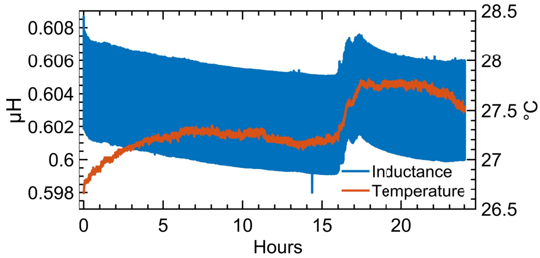


Fig. 5. Inductance recorded over a 24-hour 0.5 Hz load cycle in ambient conditions. Drift appears to behave linearly in the first 15 hours.

tools (OpenCV) and referenced against the background scale. The result is a time series of the tube diameter with a time step resolution of 30 samples per second and a geometric resolution of 3.56 pixels per mm of laser movement (Fig. 4).

Equation (7) is simplified and does not account for vertical movement of the mirror. These effects result in a vertical displacement of the laser dot less than the total diameter change of the tube, which is negligible in comparison to the displacement of the dot due to angular change.

C. Drift Measurement

We observed significant sensor drift during long recordings. To assess the scale of drift in ambient conditions and the influence of temperature we placed a tube sensor on a vibration isolated table exposed to ambient atmosphere over 24 hours and applied a 0.5 Hz square-wave pressure load cycle. Inductance was measured from the coil alongside ambient temperature with a BMP388 atmospheric pressure and temperature sensor (Bosch Sensortec, Germany) in the immediate vicinity of the tube (Fig. 5). Drift was confirmed and appeared to behave approximately linearly in the first regime (zero to sixteen hours), before a temperature impulse caused a dramatic upwards spike. We believe this impulse to have been caused by the sun shining on our lab, a hypothesis supported by in situ meteorological observations.

For the first 7 hours, the temperature sensor reported rising temperatures with a seemingly logistic growth. A likely explanation is internal temperature increase in the temperature sensor itself once powered. In time the temperature sensor reaches equilibrium and disperses the same amount of heat as it produces to the surroundings, after which the temperature data seems to correspond well with the drift of the inductance sensor. Afterwards, the response of the inductance measurement to the temperature spike suggests the drift is dominated by thermal effects. This sensitivity to small changes in temperature must be considered in applications.

We performed a linear regression on the component of the drift occurring before the temperature impulse took place, resulting in a slope of -1.48×10^{-10} H/hour.

D. Application in a Radial Artery Phantom

To assess the usability of these sensors in a vascular phantom we embedded one of the previously tested braided tube sensors in a wrist model. The wrist was cast in silicone elastomer (Ecoflex, Smooth-On, USA) in a mold produced via fused filament fabrication (FFF). The mold was modeled from a high-resolution scan of the human wrist. We also placed an FFF-printed radius and ulna bone from the same anatomical model into the cast to provide internal structure. The embedded tube had an inner diameter of 3 mm, matching a realistic range for the human radial artery[17] (Fig. 6).

We pressurized the tube system as before, but this time flushed it with water instead of air to allow ultrasound imaging. The pump compressed a small amount of air in the upstream tubing, acting as a piston on the internal water reservoir. The water was mainly static, although the pressure difference in the air piston resulted in minor oscillatory back-and-forth flow. We did not expect exchanging water for air to have a significant effect on the inductance in the coil, as the magnetic permeability of water is similar to that of air [18].

We investigated the flushed and pressurized phantom under ultrasound (Vivid E95, GE Healthcare) to compare the change in diameter of the embedded tube to that estimated from the inductance. We applied the same arterial pulse pressure profile. Due to the surrounding tissue now resisting compression, we had to apply a higher pressure to see the same range of diameter motion. The load varied between 70 – 135 mmHg, as measured in the air piston. This corresponds well to a realistic blood pressure range [4].

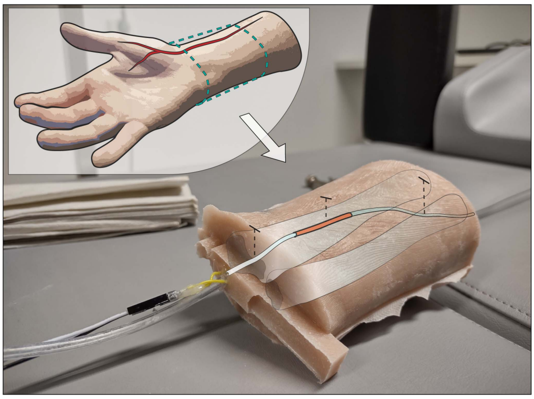


Fig. 6. Illustration of the wrist phantom with the placement of the radius, ulna, and sensor tube superimposed. The phantom simulates a section of the right wrist, with bone, blood vessel, and mediating tissue.

It is important to note that the material in the phantom has significantly different acoustic properties from those of human tissue. This difference must be accounted for when using equipment calibrated for real tissue. Most importantly for this application, the speed of sound differs, which will affect the calculated depth scale of the ultrasound machine. A rough compensation can be done by correcting for the difference in local speed of sound, also referred to as the propagation velocity:

$$z_{corr} = z\frac{C_{EF}}{C_T} \tag{8}$$

where z_{corr} is the corrected depth z, and C_t and C_{EF} are the propagation velocities in soft tissue and Ecoflex polymer respectively. As C varies between tissues, C_t is commonly taken to be the average value of several common types of soft tissue, approximately 1.54 m/s [19]. Together with the propagation velocity in the silicone material, 0.97 m/s [20], a correction factor of 0.63 is obtained. The propagation velocity in the phantom tissue is slower than the expected value for human tissue. The uncorrected depth therefore assumes a greater distance has been traveled by the reflected sound wave. If not corrected the estimated depth values will be erroneously large.

We measured the relaxed diameter a priori and confirmed this measurement under ultrasound B-mode to confirm good correspondence. In the M-mode configuration, a sequence of single ultrasound scan lines is collected and presented to illustrate the spatial movement of structures in the scan line over time. By recording the motion of the wall of the artery, a "diameter wave" can be obtained (Fig. 7). By combining the motion of a point in the tube wall with the initial, or smallest, diameter, the diameter of the vessel over time can be reconstructed. (Fig. 8) shows the diameter of the embedded soft tube determined in this way, alongside the diameter of the tube estimated using the same nominal starting diameter and the inductance measurement.

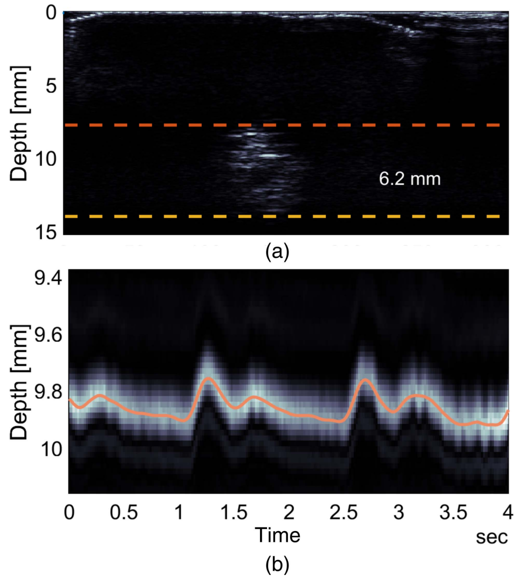


Fig. 7. (a) B-mode ultrasound of the wrist phantom, approximately 5 cm distal to the styloid process. Note that no effort was made to induce echogenicity in the phantom, resulting in a poorly defined image. (b) M-mode recording illustrating the pulsatile movement of the "lumen". A centerline trace of the tube wall has been superimposed.

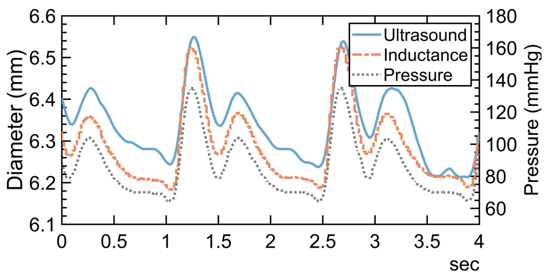


Fig. 8. Diameter calculated from phantom lumen wall movement from ultrasound plotted with the estimate from inductance. The profile of the pressure trace is consistent over the three cycles shown. At around 3 seconds a peak is shifted in the ultrasound trace. This is not reflected in the pressure or inductance curves, implying that it is an artifact introduced in the ultrasound recording.

III. Results and Discussion

Over the course of contraction and expansion, the sensor tube system experiences some hysteresis, as can be expected from the viscoelastic properties of the structural material (Fig. 9). Within the isolated loading and unloading regimes, a linear fit can still be made to estimate sensor sensitivity. Linear regression models were fitted on monotonically increasing tube expansion series using a robust fit method implemented in MATLAB's fitlm function. The results presented strong linear fits to the data (R^2 between 0.994 and 0.997). Apparent sensitivities corresponding to the regression

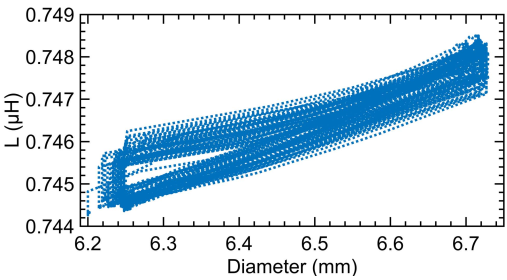


Fig. 9. Superposition of inductance and diameter traces measured over 4 minutes of cyclic loading, showing hysteresis as well as drift.

TABLE III
REGRESSION ESTIMATES OF BEST LINEAR FIT

Tube no.	Closest linear fit	Corresponding apparent sensitivity, 10^{-7} H/mm
1	$r = 6.58 \times 10^{-7} L - 46.08$	7.12
2	$r = 3.97 \times 10^{-7} L - 18.60$	4.89
3	$r = 4.96 \times 10^{-7} L - 20.15$	4.27
4	$r = 7.02 \times 10^{-7} L - 42.77$	6.20

model are presented in Table III. The sensitivity range agrees broadly with previous work [11]. The mean absolute error of the inductance estimate compared to the laser measurement ranged from 0.02 to 0.06 mm. This result is associated with some uncertainty, which is discussed later.

Sensor resolution in our setup is determined by the output of the LDC1612 converter, which has a total resolution of 28 bits. In practice, effective resolution is limited by the electromagnetic properties of the sensor element in the LC circuit and by the sample rate. As the LDC can essentially be viewed as a frequency-to-digital converter, the LSB is measured in Hz, and is determined primarily by the reference count of the converter and the variation of the signal frequency. We observed a typical frequency variation in our target signal of around 0.5%, which with a reference clock timed at 40 MHz and a sample rate of 300 Hz corresponds to 11 bits of effective resolution over the range of interest [21]. For the same configuration, reducing the sample rate to 100 Hz results in an increase in effective resolution to 12 bits as the reference count increases.

A. Error Sources

There are benefits and drawbacks to using LDCs compared to a benchtop LCR meter. Perhaps the clearest benefits are sampling rate, which can reach several thousand Hz under some circumstances, high portability, and low cost of implementation. But because the measured value of the LDC is the oscillation frequency of an LC circuit, it is necessary for the designer to have precise control of their coil characteristics. As tables I and III show, the sensitivity of the coil varies significantly with coil dimensions. As loosely wound, high-pitch coils such as the ones we have used here typically have small self-inductance values, the oscillation frequency is typically high without a very small parallel capacitance, as can be seen from (6). In our experiments, the sensor frequencies were between 6-8 MHz in this configuration. Lowering the parallel capacitance or adding inductors in series can bring the frequency range down, but this increases noise outside of the manufacturer's recommended capacitance range as the effects of parasitic capacitance in the coil become more noticeable [22] and reduces effective signal resolution. Skin effects also become significant at these frequency ranges, which may contribute some error to the estimates we derive from the inductance values. Skin effect is not accounted for in our model.

The parallel capacitor in our LC circuit was placed at the end of the connecting wires from the coil, rather than immediately outside the coil, which is the recommended configuration [21]. This may have contributed to a greater parasitic capacitance, resulting in an unstable source of error in the calculation of inductance from the oscillation frequency. In future applications the influence of the coil placement in relation to the measurement electronics should be considered.

Other sources of error relate to the uncertainty in our coil parameters. Much of this uncertainty is due to the hand-made nature of the coils. The winding angle, for example, may differ slightly over the length of the coil, as may the thickness of the polymer layer under the conductive material. The ends of the coil are also nonuniform because of the fabrication process. Improving this process so that coil geometry is more uniform would reduce this uncertainty in future applications.

Determining the diameter expansion of the tube using the laser and mirror setup is dependent on a geometric relationship which is subject to several simplifying assumptions and precision measurements. There is some uncertainty associated with the absolute values determined in this way. The pixel resolution of the laser measurement puts a lower bound on the absolute certainty of the measurement. The resolution of the video, at 0.278 mm per pixel, with an average of 255 pixels of laser travel corresponding to approximately 0.5 mm of expansion range results in a theoretical measurement resolution around 2 μm using the laser mirror setup. Errors due to simplifying assumptions in calculating the geometric relationship and imprecisions in the experimental setup are certainly larger than this resolution limit.

Significant drift was observed over the course of several-minute long recordings. The drift appeared to behave in a broadly linear manner, suggesting it could be compensated for. The sensitivity of the system to small changes in temperature poses some challenges to practical use. Temperature-sensitive elements in our experimental setup include the coil itself, the parallel capacitor, and the reference clock. The coil was not separated thermally from the measurement circuitry in our experiment, but as this more closely reflects the likely realities of a real-world application, the observed behavior may be more informative for the application designer. Perhaps most importantly, handling of the sensor might adversely affect the measurement. Regular calibration of the estimate value should be done against a known reference value.

B. Application Validation

Ultrasound recording over time is challenging. Even when recording structures that are relatively motionless, the probe must be kept perfectly still. Robot-assisted systems do exist, but are rare, and in most real cases ultrasound recordings will be done by hand [23]. The quality of the recording is therefore dependent on the skill of the operator. The practical consequence of this is that without skilled operators, many otherwise high-quality recordings may not be stable over time. In our data, the shape of the inductance measurement corresponds more closely to that of the internal pressure than the trace from the ultrasound, especially in the third cycle of the M-mode trace (Fig. 8). A likely explanation for this is instability of the ultrasound probe in relation to the "artery" over the time of the recording. In some situations, measurement of structural diameter from coil inductance could be a more reliable measure of internal geometry than operator-guided ultrasound in continuous recordings.

IV. Conclusion

We have presented the manufacture and application of soft sensor tubes consisting of braided wire coils wound inside a soft silicone rubber sleeve. As part of an LC tank together with an inductance-to-digital converter, these sensors can detect small changes in their diameter caused by differences in internal pressure with high fidelity. The sensor was sampled at 300 Hz and corresponded well with comparator measurements of the tube diameter collected optically and with ultrasound. The sample rate of the inductance measurement was higher than those easily achievable with the comparator methods, implying that the sensor might be useful in applications with high requirements for temporal resolution, such as pressure wave analysis.

While the sensor exhibits drift and hysteresis that could be prohibitive to very high precision measurements, presumably because of thermal and viscoelastic effects, the resolution was satisfactory for deformations in the tenths of a millimeter range. Linear drift seemed to be in correspondence to temperature changes in the ambient atmosphere, indicating that temperature compensation could be necessary in a practical application.

Soft braided coils embedded in flexible polymer can be used to estimate small deformations of embedded tubular structures using self-inductance, presenting an alternative to ultrasound in vascular phantoms. Possible applications include transient flow-through pressure impulse monitoring, pipe inspection, or mold channel cleaning.

In future work we want to recreate the results presented here using compliant conductors to avoid the composite effects of the coil structure on the mechanical properties of the tube. We intend to apply this well-performing measurement method to investigate numerical models of pressure coupling in the wrist for noninvasive wearable blood pressure sensors.

References

[1] D. Agnoletti *et al.*, "Pulse pressure amplification, pressure waveform calibration and clinical applications," *Atherosclerosis*, vol. 224, no. 1, pp. 108–112, Sep. 2012, doi: 10.1016/j.atherosclerosis.2012.06.055.

[2] M.-M. Laurila *et al.*, "Evaluation of printed P(VDF-TrFE) pressure sensor signal quality in arterial pulse wave measurement," *IEEE Sensors J.*, vol. 19, no. 23, pp. 11072–11080, Dec. 2019, doi: 10.1109/JSEN.2019.2934943.

[3] W. W. Nichols, M. F. O'Rourke, and C. Vlachopoulos, "Pressure pulse waveform analysis," in *McDonald's Blood Flow in Arteries. Theoretical, Experimental and Clinical Principles*, 6th ed. London, U.K.: Hodder Arnold, 2011.

[4] W. W. Nichols, M. F. O'Rourke, and C. Vlachopoulos, *McDonald's Blood Flow in Arteries. Theoretical, Experimental and Clinical Principles*, 6th ed. London, U.K.: Hodder Arnold, 2011.

[5] Y. Ma, J. Choi, A. Hourlier-Fargette, Y. Xue, H. U. Chung, and J. Y. Lee, "Relation between blood pressure and pulse wave velocity for human arteries," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 44, pp. 11144–11149, 2018, doi: 10.1073/pnas.1814392115.

[6] S. G. Yazdi, P. H. Geoghegan, P. D. Docherty, M. Jermy, and A. Khanafer, "A review of arterial phantom fabrication methods for flow measurement using PIV techniques," *Ann. Biomed. Eng.*, vol. 46, no. 11, pp. 1697–1721, Nov. 2018, doi: 10.1007/s10439-018-2085-8.

[7] F. S. Solberg, S. Kohtala, H. Vestad, and M. Steinert, "A combined photoplethysmography and force sensor prototype for improved pulse waveform analysis," in *Proc. IEEE SENSORS*, Oct. 2019, pp. 1–4, doi: 10.1109/SENSORS43011.2019.8956487.

[8] M. Kaisti *et al.*, "Clinical assessment of a non-invasive wearable MEMS pressure sensor array for monitoring of arterial pulse waveform, heart rate and detection of atrial fibrillation," *NPJ Digit. Med.*, vol. 2, no. 1, May 2019, doi: 10.1038/s41746-019-0117-x.

[9] S. Chen, J. Qi, S. Fan, Z. Qiao, J. C. Yeo, and C. T. Lim, "Flexible wearable sensors for cardiovascular health monitoring," *Adv. Healthcare Mater.*, vol. 10, no. 17, Sep. 2021, Art. no. 2100116, doi: 10.1002/adhm.202100116.

[10] G. Jo, T.-H. Yang, J.-H. Koo, M.-H. Jun, and Y.-M. Kim, "A transfer function model development for reconstructing radial pulse pressure waveforms using non-invasively measured pulses by a robotic tonometry system," *Sensors*, vol. 21, no. 20, p. 6837, Oct. 2021, doi: 10.3390/s21206837.

[11] W. Felt, K. Y. Chin, and C. D. Remy, "Contraction sensing with smart braid McKibben muscles," *IEEE/ASME Trans. Mechatronics*, vol. 21, no. 3, pp. 1201–1209, Jun. 2016, doi: 10.1109/TMECH.2015.2493782.

[12] W. Felt and C. D. Remy, "Smart braid: Air muscles that measure force and displacement," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2014, pp. 2821–2826, doi: 10.1109/IROS.2014.6942949.

[13] A. V. Prituja, H. Banerjee, and H. Ren, "Electromagnetically enhanced soft and flexible bend sensor: A quantitative analysis with different cores," *IEEE Sensors J.*, vol. 18, no. 9, pp. 3580–3589, May 2018, doi: 10.1109/JSEN.2018.2817211.

[14] W. Felt and C. David Remy, "A closed-form kinematic model for fiber-reinforced elastomeric enclosures," *J. Mech. Robot.*, vol. 10, no. 1, Nov. 2017, doi: 10.1115/1.4038220.

[15] J. Vaicekauskaite, P. Mazurek, S. Vudayagiri, and A. L. Skov, "Mapping the mechanical and electrical properties of commercial silicone elastomer formulations for stretchable transducers," *J. Mater. Chem. C*, vol. 8, no. 4, pp. 1273–1279, Jan. 2020, doi: 10.1039/C9TC05072H.

[16] T. Ha, J. Ma, J. Blindheim, T. Welo, G. Ringen, and J. Wang, "In-line springback measurement for tube bending using a laser system," *Proc. Manuf.*, vol. 47, pp. 766–773, 2020, doi: 10.1016/j.promfg.2020.04.233.

[17] H. V. Riekkinen, K. O. Karkola, and A. Kankainen, "The radial artery is larger than the ulnar," *Ann. Thoracic Surgery*, vol. 75, no. 3, pp. 882–884, Mar. 2003, doi: 10.1016/S0003-4975(02)04557-5.

[18] B. D. Cullity and C. D. Graham, *Introduction to Magnetic Materials*, 2nd ed. New York, NY, USA: Wiley, 2009.

[19] M. K. Feldman, S. Katyal, and M. S. Blackwood, "U.S. artifacts," *Radio Graph.*, vol. 29, no. 4, pp. 1179–1189, Jul. 2009, doi: 10.1148/rg.294085199.

[20] A. Cafarelli, P. Miloro, A. Verbeni, M. Carbone, and A. Menciassi, "Speed of sound in rubber-based materials for ultrasonic phantoms," *J. Ultrasound*, vol. 19, no. 4, pp. 251–256, Apr. 2016, doi: 10.1007/s40477-016-0204-7.

[21] C. Oberhauser, *Optimizing L Measurement Resolution for the LDC161x and LDC1101*. Dallas, TX, USA: Texas Instruments, Feb. 2016.

[22] *Sensor Design for Inductive Sensing Applications Using LDC. Rev C.* Texas Instruments, Dallas, TX, USA, May 2021.

[23] R. Monfaredi, "Robot-assisted ultrasound imaging: Overview and development of a parallel telerobotic system," *Minimally Invasive Therapy Allied Technol.*, vol. 24, no. 1, pp. 54–62, 2015, doi: 10.3109/13645706.2014.992908.

Torjus L. Steffensen was born in Bergen, Norway, in 1993. He received the M.Sc. degree in mechanical engineering from the Norwegian University of Science and Technology (NTNU) in 2019, where he is currently pursuing the Ph.D. degree in medical technology with the TrollLABS.

His research interests include soft sensors, wearable longitudinal health monitoring, and data processing methods for physiology time series data.

Marius Auflem was born in Stavanger, Norway, in 1992. He received the M.Sc. degree in mechanical engineering from the Norwegian University of Science of Technology (NTNU) in 2018.

He currently works as an Industrial Ph.D. candidate at the TrollLABS, NTNU, together with Laerdal Medical. His research interests include medical simulation technology, soft robotics, and prototyping in the early stages of product development.

Håvard N. Vestad (Member, IEEE) was born in Bergen, Norway, in 1993. He received the M.Sc. degree in mechanical engineering from the Norwegian University of Science and Technology in 2018, where he is currently pursuing the Ph.D. degree with the Prototyping Laboratory TrollLABS.

His research interests include soft robotics and sensors, prototyping methodology, and early prototyping concept development for complex sensor problems.

Martin Steinert was born in Dresden, Germany. He received the B.A., M.A., and Ph.D. (Dr.Rer.Pol) degrees from the University of Fribourg, Switzerland, majoring in technology management.

He has been an Assistant Professor at the University of Fribourg, Switzerland, and a Visiting Scholar at the MIT and Stanford University before changing full-time to Stanford University as the Deputy Director of the Center for Design Research (CDR), and an Assistant Professor (Acting) of Mechanical Engineering with Stanford University. Since 2013, he has been a Full Professor of Engineering Design at the Department of Mechanical and Industrial Engineering (MTP), Norwegian University of Science and Technology (NTNU). He is a Professor of the NTNU. As of August 2021, he has more than 200 publications. His research interests focus on the fuzzy front end of new product development and design: optimizing the intersection of engineering design thinking and new product development, mechatronics/sensors, and computer sciences (especially machine learning). A special focus is on conceptual development and alpha prototype generation of high-performance requirements, as well as on experimental tools and setups.

Dr. Steinert has been a member of the Norwegian Academy of Technological Sciences (NTVA) since 2015. He has several prizes in both teaching and research.

C9

Lost in Transit: Implications and Insights for Making

Medical Task Trainer Prototypes with an Open-Source Hardware Paradigm

Proceedings of the Design Society: DESIGN2022 Conference

Daniel Nygård Ege, Marius Auflem, Martin Steinert (2022),

INTERNATIONAL DESIGN CONFERENCE – DESIGN 2022
https://doi.org/10.1017/pds.2022.127

Lost in Transit: Implications and Insights for Making Medical Task Trainer Prototypes with an Open Source Hardware Paradigm

D. N. Ege, M. Auflem and M. Steinert
Norwegian University of Science and Technology, Norway
danieneg@stud.ntnu.no

Abstract
This paper presents an open-source novel intravenous cannulation task trainer developed during the Covid-19 pandemic for unsupervised clinical skill practice. Multiple user errors were uncovered when observing 13 registered nurses using the task trainer during a two-hour unsupervised skill training session. These insights raise the question of how OSH needs to share more than just device descriptions and assembly instructions- as designs are being shared only in its current state of an ongoing project, sharing insights, user errors and test results should be encouraged and prioritized.

Keywords: open source design, prototyping, medical task trainer, healthcare design, prototype testing

1. Introduction

With the Covid-19 pandemic requiring remote working, virtual training, and education, we are witnessing increased initiatives that transfer product development from professionals to the public through the emergence of Open-Source Hardware (OSH). These initiatives are sparked by the ease of sharing information and digital blueprints online and the growing accessibility of distributed manufacturing equipment such as 3D printers (Bonvoisin et al., 2017; Boujut, 2019). OSH is based on the open-source principles that originated for software (Open Source Initiative, 2007) and transfer these to tangible artifacts, machines, and devices whose design has been released so that anyone can make, modify, distribute, and use them (Balka and Herstatt, 2011). This transformation is not straightforward, as while software can be displayed and shared online without specific tools, physical objects need to be described through more complex constructs (Bonvoisin and Boujut, 2015), such as CAD models and assembly guides. Still, recent efforts of openly sharing hardware designs and projects for free allow for more collaboration with others with similar needs and increase the speed of innovation (Daniel K. and Peter J., 2012). OSH enables designs and product solutions to be distributed instantly by relying on people's ability to manufacture and assemble equipment/products/hardware themselves. It has given researchers access to expensive test equipment for a fraction of the cost (Pearce, 2014). The medical technology sector also calls for more OSH tools to reduce costs and democratize access to new technology (Linde and Kunkler, 2015). Examples exist of OSH solutions for this sector, including Open Source CT scanners, infusion pumps, and prostheses (Niezen, Eslambolchilar, and Thimbleby, 2016).

Given this context, we consider a case of a novel intravenous cannulation task trainer developed during the Covid- 19 pandemic. Throughout this development project, prototypes were used to probe an ambiguous solution space (Gerstenberg et al., 2015) to learn as quickly as possible (Lauff, Kotys-

Schwartz, and Rentschler 2018) and validate assumptions (Lim et al., 2008; Menold et al., 2017). Users were consulted from the initial phases to uncover current task trainers' needs and pain points (Sutcliffe and Sawyer 2013). Sufficient fidelity and realism were captured by continuously testing prototypes with expert users, as suggested by Ege et al. (2021), while mitigating the risk of severe obstacles later in the project (Thomke and Reinertsen 1998; Takeuchi and Nonaka 1986). Intentional and reflective prototypes were built and tested with users to explore specific dimensions of the concept (Houde and Hill, 1997) and to elicit experts' tacit knowledge at an early stage of development, as previously demonstrated by Auflem et al. (2019) and Ege et al. (2020). Prototypes were developed through design-build-test cycles, as described by Gerstenberg et al. (2015), driving the development as new insights and knowledge got revealed.

The case presented in this paper showcase both the development journey of an IV cannulation trainer, the deployment of this trainer in structured user testing, and the obtained insights and implications for utilizing this device in unsupervised training scenarios as an OSH design. The case exemplifies and highlights the importance of structured user-testing and user-centered design mindset when providing OSH healthcare training solutions. By showcasing the prototyping journey of this case project, we aim to exemplify user-centered and prototype-driven development of medical training equipment. Moreso, we show how to effectively deploy prototypes with users to gain valuable insights for further development, and for sharing insights in the context of OSH solutions. By piloting the concept, the authors observed several procedural errors performed by users. Without testing, these errors and other interactions would not have been identified. By the insights presented we raise the question of how OSH needs to share more than just device descriptions and assembly instructions- as designs are being shared only in its current state of an ongoing project, sharing insights, user errors, and test results should be encouraged and prioritized.

2. Background

Nurses are expected to apply cognitive, psychomotor, and procedural skills when performing essential medical procedures in clinical environments. Intravenous (IV) cannulation is such an essential skill, in which a cannula is placed inside a vein to provide venous access to treat various types of patients, sample blood, and administer fluids in acutely or critically ill patients (Jones et al. 2014). Intravenous cannulation is among the most common procedures nurses face, and most of all hospitalized patients undergo it (Devenny et al., 2018). Generally, educators use medical task trainers or fellow students for teaching IV cannulation skills. Several task trainers for practicing IV cannulation exist on the market, ranging from full-size human mannequins to small, wearable pads to simulate cannulation, injections, and blood tests. These, to different degrees, embody both physical and visual attributes of the human arm to facilitate realistic motor skill training without the risk of infections and injuries associated with practicing on fellow students.

Jacobson and Winslow (2005) found that almost a quarter of all IV insertion attempts were unsuccessful. Nurses needed two and even three shots before correctly placing the cannula in a vein. This low success rate can be critical in acute situations where patients rely on quick access to medicine and blood (Witting 2012). Experienced nurses show significantly higher success rates than those less skilled (Jacobson and Winslow 2005), indicating a need for easily accessible equipment for skill training. Available commercial training equipment can provide equivalent procedural skill training to training using fellow students, without the percutaneous and infectious risk involved in practicing on peers (Jones et al. 2014). However, these task trainers are both often expensive and tedious to use and maintain. Even though they are available to educational institutions, they are not widely available or used by practicing health care professionals. Furthermore, with the Covid- 19 pandemic requiring remote working, virtual training, and education, the availability of skill training equipment has decreased for medical students alike.

As a response to the COVID-19 pandemic, designs for medical personal protective equipment PPE, ventilators and oxygenation were shared in various communities (Mora, Duarte, and Ratti 2020). Furthermore, medical training equipment is also being shared and published free online, such as intubation task trainers made using low-cost materials and 3D printers (Park et al. 2019; Tasnim, Parikh, and Naik 2020) that simplify access to task training equipment. Because of postponed and

https://doi.org/10.1017/pds.2022.127 Published online by Cambridge University Press

canceled clinical skill exercises, the demand for self-directed learning in healthcare education has increased. Open-source solutions for practicing procedures have therefore been made available, and several tasks are now being taught through self-directed learning. For instance, Gieswein et al. (2021) developed low-fidelity simulation models from household items for medical students to make themselves to practice essential procedural skills without access to the high-fidelity training equipment in a simulation center traditionally used. These skills included central venous catheterization and emergency airway management.

3. Development of an open-source IV cannulation task trainer

In response to the highlighted challenges with IV cannulation training, the authors have developed a novel task trainer concept for use in medical education and the repetition of skill training (Figure 3). In the preliminary design phases, expert users, in this case, registered nurses, were consulted to uncover needs and functional requirements. Initial requirements included a realistic feel and look, as well as easy use and maintenance. The task trainer must also simulate blood flashback, in which blood fills a chamber in the back of the cannula when correctly placed in a vein. Issenberg et al. (1999) argue that medical training equipment needs to provide users with feedback, allow repetitive practice, integrate with the curriculum, and simulate clinical variation. These factors also influenced the design of the final concept.
Initial prototypes, illustrated in Figure 1, were made to elicit expert users' tacit knowledge quickly as a primary bounding object between designer and practitioner. A latex glove (Prototype 1) with thin-walled tubes glued to the inside was presented to expert users to observe their interaction with a wearable task trainer. Prototype 2 informed material selection and tactile and visual realism of the eventual final design. It showed that a silicone skin and thin silicone tubing could recreate parts of the tactile sensation of inserting a cannula in a patient's vein but lacked the possibility of veins to move more freely under the skin. Prototype 3 addressed this by stretching silicone over a sponge with silicone tubing laying free underneath. Iterating on this concept, Prototype 4 implemented a 3D printed case. Prototype 5 improved this design before it was optimized for 3D printing and simpler manufacturing in the final design.

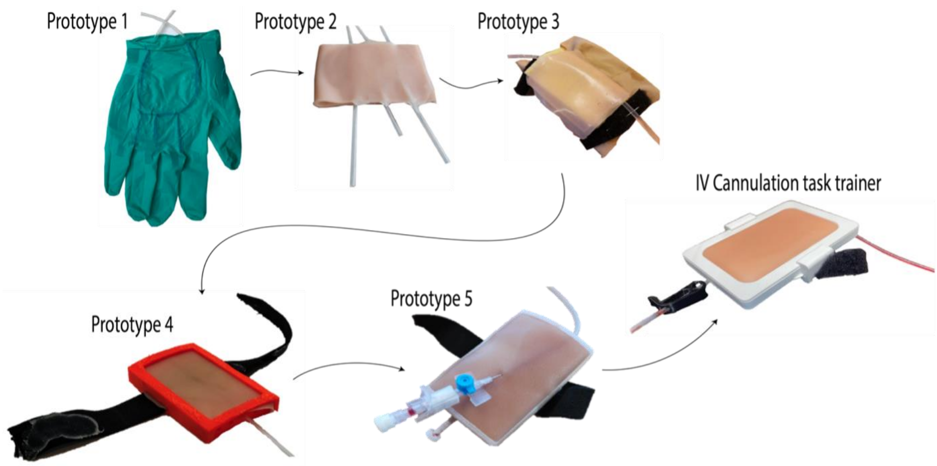


Figure 1. Iterative prototyping journey

Figure 2 show and describe all the components of the final IV task trainer concept. The 3D- printed case consists of a top and bottom part that snaps together to keep all components in place. Depending on the training scenario, a clamp and flexible strap make it possible to attach the task trainer to someone's hand, arm, or leg with Velcro. The case is 120mm by 80mm, with an opening of 100mm by 60mm for the skin and vein to be palpated and a cannula inserted. Tabs in the corners of the top part stretch the skin, while openings at the front and back allow a silicone tube to move freely back and forth. The silicone skin is made from Smooth-On Ecoflex 30 platinum cure silicone with skin color. Cotton fabric that was laser cut to shape and embedded during casting gives the perimeter of the skin

https://doi.org/10.1017/pds.2022.127 Published online by Cambridge University Press

reinforcement. A casting mold made from acrylic sheets has a faux leather bottom that provides the silicone with skin texture and more realistic looks. Soft silicone tubing represents veins. The tube has an outer diameter of 4 mm and a wall thickness of 0,5 mm. The veins are filled with artificial blood and pressurized by an external reservoir (Figure 3) hung at a variable height (40 cm to 70 cm) to simulate different blood pressures. A 3D printed hose clamp is attached to the opposite end of the tube to make the system pressurized. When pierced by a cannula, blood fills the chamber in the back, indicating blood flashback. When it starts to leak after repeated use, the vein can be pulled thru the task trainer, and the clamp moved up. Used tubing can simply be cut off and thrown away, eliminating the need for flushing and washing after use as with traditional equipment. Filling the remaining of the task trainer is an 8 mm thick high-density melamine foam. A channel is cut in the foam for the vein. Changing the shape and size of this channel influences the palpability of the vein and how much it is allowed to roll, allowing for clinical variation and difficulty during exercise.

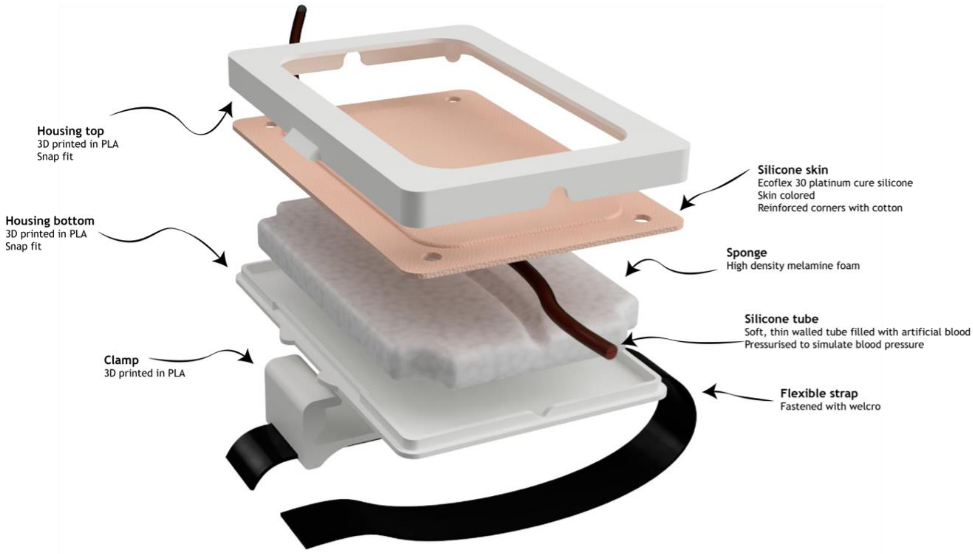


Figure 2. Exploded view of the conceptual prototype

Figure 3. Variable stand with artificial blood

https://doi.org/10.1017/pds.2022.127 Published online by Cambridge University Press

To maximize the availability of OSH, it should use readily available materials and components or custom components that anyone can easily manufacture. Therefore, the conceptual prototype presented in this paper can be manufactured using a 3D printer, laser cutter, and readily available materials. The case can be printed without support structures, minimizing sources of error during printing. To make the skin, silicone moulding was simplified by a laser-cut mould. The rest of the components can be supplied from well-equipped hardware stores or online. The final design combines the functional requirements elicited in the initial stage of development while being in line with the Open Source Hardware Associations best practices ("Best Practices for Open Source Hardware 1.0," n.d.)

4. Method

We want to observe how the task trainer prototype is used in a realistic scenario, in which nurses practice IV cannulation skills in an unsupervised group setting. To capture realistic interactions and potential errors during use, thirteen nurses and two nursing students participated in a practice session for IV cannulation in their workplace. The session lasted 2 hours. Participants were placed in pairs and practiced IV cannulation on each other with the task trainer. During the session, participants were observed and filmed to capture their interactions with the task trainer. The number of attempts each participant needed to insert the cannula correctly was counted.

Figure 4. Task trainer in use during the training session.

5. Results

Thirteen nurses and two students used on average 3,2 attempts (SD= 2,57) to correctly insert the IV cannula in the task trainer's vein. However, half of the participants used two or fewer attempts. The histogram in Figure 5 shows how many participants used the same number of attempts.

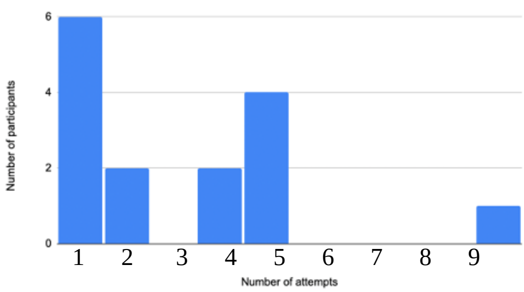


Figure 5. Histogram showing how many attempts participants used

https://doi.org/10.1017/pds.2022.127 Published online by Cambridge University Press

After the training session, an experienced nurse assisted in a video review to determine user errors and procedural mistakes. Figure 6 illustrates some of the user errors that were observed. In A, the needle was inserted far into the task trainer without receiving blood flashback. Instead of pulling the needle out and trying again, the participant further inserted the needle, which would cause excessive pain on an actual patient. B depicts a participant touching the insertion site while the needle is inserted, which significantly increases the risk of infection. In C, a participant that missed the vein moves the needle around under the skin, causing excessive pain to the patient. D shows a participant stopping blood flow while removing the needle by pressing directly on top of the catheter. The nurse should instead put pressure further up to decrease infection risk and pain for the patient. E shows how a participant left the entire needle inside the vein instead of pulling it out. This would likely perforate the vein and lead to a hematoma on an actual patient. Whereas the first 5 cases show procedural mistakes, F shows how one participant managed to simplify the procedure using the prototype's case as a guide. The vein lies perpendicular to the case, so this participant could perfectly insert the needle in the vein by using the case as a guide and sliding the catheter against it. Most of the participants were also observed touching the vein after cleaning it with alcohol, increasing the risk of infection. Several participants also used the same needle multiple times when missing the vein, which can lead to parts of the cannula breaking inside the patient's vein.

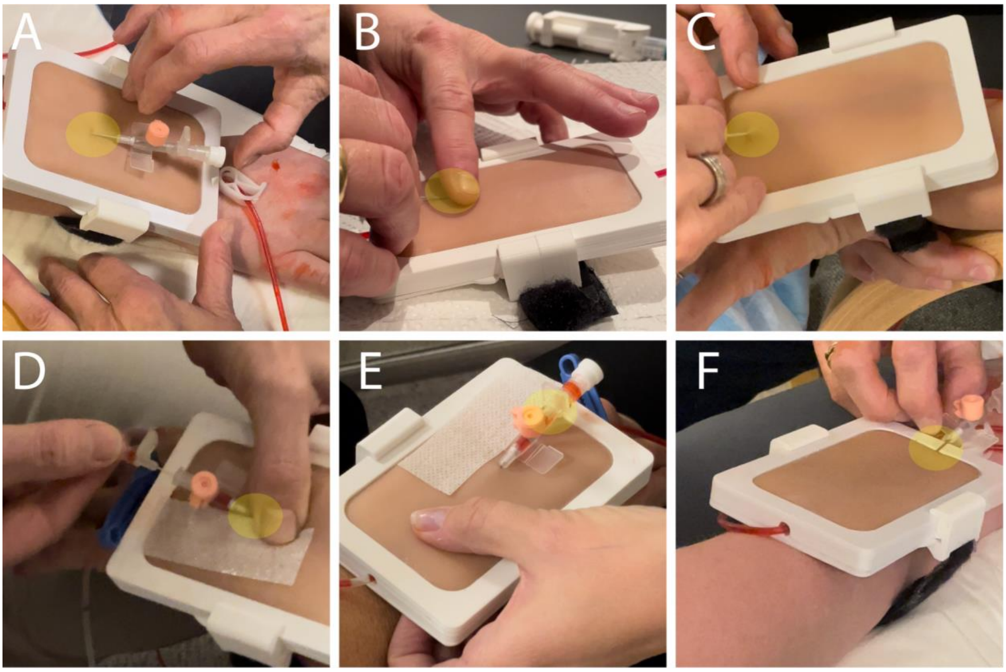


Figure 6. User errors with a yellow circle indicating the mistake. A: needle inserted too far, B: touching insertion site, C: needle moved around under the skin, D: pressure applied on top of the catheter, E: needle not pulled out, F: prototype case used a guide to align the needle with the vein.

6. Discussion

This paper describes how prototypes can be leveraged to uncover user errors through user interaction in realistic, ecologically valid scenarios. In this case, new needs and requirements arose because of the observed user errors. All participants were able to insert the cannula in a vein, indicating that the task trainer can, in fact, be used to practice IV cannulation. However, the observed user errors, both regarding the clinical procedure and use can affect the learning outcome of using the task trainer unsupervised. By learning a procedure wrong from the start, it is harder to correct it later. We,

https://doi.org/10.1017/pds.2022.127 Published online by Cambridge University Press

therefore, argue that emphasis should be put on mitigating such user errors during the design process. As shown in this paper, testing a conceptual task trainer can help discover potential user errors. This should arguably be done as early as possible so that it is possible to address them, either by design changes or by also providing user instructions. When designing training equipment emulating the human body, an inherent trade-off the designers need to address is realism and foolproof design. If the aim is to make a task trainer as realistic as possible, it needs to allow for the same mistakes that can be done on actual patients, such as the procedural errors witnessed in this study, but can at the same time allow for learning errors. The solution might be to focus more on user feedback to reduce improper use. Other OSH devices, such as open-source 3D printers, use sensors and self-calibration to mitigate user errors. Using smart sensor systems or machine vision could mitigate specific errors of the task trainer we are presenting but would increase the complexity significantly.

The development project presented here exemplifies how prototyping can be used for learning and knowledge generation. Prototypes are used actively to elicit critical functionalities to make the product function as intended and sufficiently. In this case, it concerns both building, testing, and iterating on concepts and learning about their potentials and pitfalls. A critical notion to make is how to convey this information, the tangible insights obtained through development, to mitigate sources of user error, both in the context of assembly and use. We argue that the development of OSH is more than sharing of technical blueprints. There is a need for detailed instructions and testing of design in context to reveal sources of error extending beyond the technical capabilities of the hardware itself. We also need to consider the intended end user and how we address the user when developing OSH solutions. Providing sufficient assembly instructions and making sure people can assemble and use the device is quite simple. Understanding how the product will be used/altered/interacted with is more ambiguous. Compared to traditional prototype testing and evaluation, designers lose control of this situation. As exemplified in this paper, a solution might be to observe how users interact with a device in an unsupervised setting. Then, insight and observations can be used to address this uncertainty.

While current OSH research has focused on the availability of manufacturing tools, sharing practices, and licensing, there is a lack of research and best practices regarding the development process and sharing of knowledge generated throughout development. In this paper, we consider an IV-cannulation trainer prototype, which was developed using an iterative development framework. Prototypes were developed through design-build-test cycles, driving the development as new insights and knowledge got revealed. A prototype to evaluate the concept potential was tested in the context of unsupervised cannulation skill training, thus revealing areas for improvement and sources of wrong learning for end-users. Hence, we exemplify how the design and build stages of hardware development should be shared and results from testing and evaluation of the shared designs.

As OSH facilitates new product designs to be shared, built, tested, re-designed and re-distributed, we emphasize the importance of sharing more than merely technical blueprints and building instructions. We argue that this is especially important when distributing open hardware designs, as opposed to in a conventional new product development process, because its unknown who and in what state the design is going to be used in. Considering designs being shared is only the current state of an ongoing project, sharing insights, user errors, and test results should be encouraged. This way, new designs could be shared with the assurance that they function as intended and thus mitigate sources of error both considering assembly and of end-use. Alternatively, novel designs could be shared before being tested, re-designed, and thus improved by other researchers and developers to mitigate the risk of not meeting the targeted requirements. Further research is necessary to determine effective knowledge transfer of OSH designs. As the transfer of tacit knowledge trough physical artefacts is challenging to obtain given physical distance and virtual sharing of ideas, sharing development stories, experiences of manufacturing and deploying the designs could be valuable. Moreover, this could be critical given use-cases (such as the one presented in this paper) where critical learning outcome could be affected by premature design choices and designs not sufficiently tested.

7. Conclusion

With the Covid-19 pandemic requiring remote working, virtual training, and education, we are witnessing an increase in initiatives that transfer product development from professionals to the public

https://doi.org/10.1017/pds.2022.127 Published online by Cambridge University Press

through the emergence of Open-Source Hardware. Open-Source designs grant access to tangible artifacts by relying on people's ability to manufacture and assemble these themselves. Because of postponed and canceled clinical skill exercises, open-source solutions for practicing procedures have been made available, and several tasks are now being taught through self-directed learning. We, however, question how this affects the learning outcome of task training. We consider a case of a novel intravenous cannulation task trainer developed during the Covid- 19 pandemic. The task trainer is designed as an Open-source device for unsupervised clinical skill practice. Multiple user errors were uncovered by observing 13 registered nurses and two students using this task trainer during a two-hour unsupervised skill training session. Without user testing, these errors and other interactions might not have been identified. These insights raise the question of how OSH needs to share more than just device descriptions and assembly instructions- as designs are being shared only in its current state of an ongoing project, sharing insights, user errors, and test results should be encouraged. Therefore, we argue that the insights and learning from an iterative design process should be shared on the same basis as final hardware designs to take full advantage of OSH's potential.

Acknowledgment

This research is funded by the Research Council of Norway, project numbers 321386 and 290404.
We also thank the Home Health Nursing employees in Oppdal, Norway, for their contribution to this study.

References

Auflem, M., Erichsen, J.F., Steinert, M., 2019. Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators. Procedia CIRP 84, 572–578. https://doi.org/10.1016/j.procir.2019.04.202

Balka, K., Herstatt, C., 2011. Open source product development: the meaning and relevance of openness, 1. ed. ed, Gabler Research Forschungs-/Entwicklungs-/Innovationsmanagement. Gabler, Wiesbaden.

Best Practices for Open Source Hardware 1.0. (n.d.). Open Source Hardware Association. Retrieved February 15, 2022, from https://www.oshwa.org/sharing-best-practices/

Bonvoisin, J., Boujut, J.-F., 2015. OPEN DESIGN PLATFORMS FOR OPEN SOURCE PRODUCT DEVELOPMENT: CURRENT STATE AND REQUIREMENTS. 80-8 Proc. 20th Int. Conf. Eng. Des. ICED 15 Vol 8 Innov. Creat. Milan Italy 27-300715.

Bonvoisin, J., Mies, R., Boujut, J.-F., Stark, R., 2017. What is the "Source" of Open Source Hardware? J. Open Hardw. 1, 5. https://doi.org/10.5334/joh.7

Boujut, J.-F., 2019. The Emerging Phenomenon of Open Source Hardware Design (No. hal-02086837), Post-Print, Post-Print. HAL.

Daniel K., F., Peter J., G., 2012. Open-Source Hardware Is a Low-Cost Alternative for Scientific Instrumentation and Research. Mod. Instrum. 2012. https://doi.org/10.4236/mi.2012.12002

Devenny, A., Lord, D., Matthews, J., Tuhacek, J., Vitlip, J., Zhang, M., Bucha, A.C., Wang, L., Buckley, J.M., Cowperthwait, A., 2018. Avstick: An Intravenous Catheter Insertion Simulator for Use with Standardized Patients. Nurs. Educ. Perspect. 1. https://doi.org/10.1097/01.NEP.0000000000000225

Ege, D.N., Auflem, M., Lilleløkken, O., Steinert, M., 2021. User involvement in early-stage design of medical training devices – case of a palpation task trainer prototype. Design for Health 5, 214–232. https://doi.org/10.1080/24735132.2021.1980270

Ege, D., Lilleløkken, O., Auflem, M., Steinert, M., 2020. Dealing with Ecological Validity and User Needs when Developing Simulation Based Training Equipment – Case Study of a Medical Palpation Task Trainer. Procedia CIRP 91, 722–727. https://doi.org/10.1016/j.procir.2020.03.116

Gerstenberg, A., Sjöman, H., Reime, T., Abrahamsson, P., Steinert, M., 2015. A Simultaneous, Multidisciplinary Development and Design Journey – Reflections on Prototyping, in: Chorianopoulos, K., Divitini, M., Baalsrud Hauge, J., Jaccheri, L., Malaka, R. (Eds.), Entertainment Computing - ICEC 2015. Springer International Publishing, Cham, pp. 409–416. https://doi.org/10.1007/978-3-319-24589-8_33

Gieswein, J., Tennill, R., Austin, R., Delfino, K., Dynda, D., Kim, S., 2021. Developing a Novel, At-Home Procedure Curriculum for Fourth-Year Medical Students in Response to the COVID-19 Pandemic. Cureus 13. https://doi.org/10.7759/cureus.15215

Houde, S., Hill, C., 1997. What do Prototypes Prototype?, in: Helander, M.G., Landauer, T.K., Prabhu, P.V. (Eds.), Handbook of Human-Computer Interaction (Second Edition). North-Holland, Amsterdam, pp. 367–381. https://doi.org/10.1016/B978-044481862-1.50082-0

https://doi.org/10.1017/pds.2022.127 Published online by Cambridge University Press

Issenberg, S.B., McGaghie, W.C., Hart, I.R., Mayer, J.W., Felner, J.M., Petrusa, E.R., Waugh, R.A., Brown, D.D., Safford, R.R., Gessner, I.H., Gordon, D.L., Ewy, G.A., 1999. Simulation Technology for Health Care Professional Skills Training and Assessment. JAMA 282, 861–866. https://doi.org/10.1001/jama.282.9.861

Jacobson, A.F., Winslow, E.H., 2005. Variables influencing intravenous catheter insertion difficulty and failure: An analysis of 339 intravenous catheter insertions. Heart Lung 34, 345–359. https://doi.org/10.1016/j.hrtlng.2005.04.002

Jones, R.S., Simmons, A., Boykin, G.L., Stamper, D., Thompson, J.C., 2014. Measuring Intravenous Cannulation Skills of Practical Nursing Students Using Rubber Mannequin Intravenous Training Arms. Mil. Med. 179, 1361–7. https://doi.org/10.7205/MILMED-D-13-00576

Lauff, C., Kotys-Schwartz, D., Rentschler, M., 2018. What is a Prototype? What are the Roles of Prototypes in Companies? J. Mech. Des. 140. https://doi.org/10.1115/1.4039340

Lim, Y.-K., Stolterman, E., Tenenberg, J., 2008. The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. ACM Trans. Comput.-Hum. Interact. TOCHI 15, 7:1-7:27. https://doi.org/10.1145/1375761.1375762

Linde, A., Kunkler, K., 2015. The Evolution of Medical Training Simulation in the U.S. Military. https://doi.org/10.13140/RG.2.1.3643.4647

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

Mora, S., Duarte, F., Ratti, C., 2020. Can Open Source Hardware Mechanical Ventilator (OSH-MVs) initiatives help cope with the COVID-19 health crisis? Taxonomy and state of the art. HardwareX 8, e00150. https://doi.org/10.1016/j.ohx.2020.e00150

Niezen, G., Eslambolchilar, P., Thimbleby, H., 2016. Open-source hardware for medical devices. BMJ Innov. 2. https://doi.org/10.1136/bmjinnov-2015-000080

Park, L., Price-Williams, S., Jalali, A., Pirzada, K., 2019. Increasing Access to Medical Training With Three-Dimensional Printing: Creation of an Endotracheal Intubation Model. JMIR Med. Educ. 5, e12626. https://doi.org/10.2196/12626

Pearce, J., 2014. Open-source lab: how to build your own hardware and reduce research costs. Elsevier, Amsterdam ; Boston.

Sutcliffe, A., Sawyer, P., 2013. Requirements elicitation: Towards the unknown unknowns. Presented at the 2013 21st IEEE International Requirements Engineering Conference, RE 2013 - Proceedings, pp. 92–104. https://doi.org/10.1109/RE.2013.6636709

Takeuchi, H., and I. Nonaka. 1986. “Stop Running the Relay Race and Take up Rugby.” Harvard Business Review 64 (1): 137–146.

Tasnim, H., Parikh, N., Naik, N., 2020. Designing/Building a Low-Fidelity SimulationIn Collaboration with Weill Cornell’s Emergency Department, in: ISAM 2019. Presented at the ISAM 2019.

Open Source Initiative, “The Open Source Definition”, 2007. URL https://opensource.org/osd (accessed 11.8.21).

Thomke, S., Reinertsen, D., 1998. Agile Product Development: Managing Development Flexibility in Uncertain Environments. Calif. Manage. Rev. 41, 8–30. https://doi.org/10.2307/41165973

Witting, M.D., 2012. IV Access Difficulty: Incidence and Delays in an Urban Emergency Department. J. Emerg. Med. 42, 483–487. https://doi.org/10.1016/j.jemermed.2011.07.030

C10

Ultrasound Phantom

US Patent App. 17/641,651

Torjus Lines Steffensen, Carlo Kriesi, Martin Steinert, Thomas Lafrenz,
Jostein Rødseth Brede, Marius Auflem (2022),

C11

A Pain in the Neck: Prototyping and Testing of a Patient Simulator Neck for Spinal Immobilization Training

Design for Health Journal

Emil Matias Henriksen, Marius Auflem, Martin Steinert (2022),

DESIGN FOR HEALTH
https://doi.org/10.1080/24735132.2023.2184995

Routledge
Taylor & Francis Group

ARTICLE

OPEN ACCESS Check for updates

A pain in the neck: prototyping and testing of a patient simulator neck for spinal immobilization training

Emil Matias Henriksen, Marius Auflem and Martin Steinert

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway

ABSTRACT

This paper presents the development and user-testing of a novel concept for patient simulators, aiming to enhance spinal immobilization training. Iterative prototype interactions with critical stakeholders revealed a need for evidence-based guidelines and training for suspected neck injuries. A realistic and compliant neck prototype with sensor feedback was developed to address the need for objective performance metrics. The conceptual prototype was used in an experimental study ($n = 12$) to obtain subjective and objective feedback on its characteristics and use in medical training. In the experiment, users were asked to perform spinal immobilization techniques on a simulator while sensor data recorded head and neck movements. Furthermore, a Likert-scale questionnaire and subjective feedback were gathered. Results are used to discuss proposed performance metrics and whether they can be used as quality performance indicators for formative and summative training feedback. The results also suggest the neck prototype to realistically simulate an unconscious patient regarding the obtained range of motion and spinal compliance.

ARTICLE HISTORY
Received 16 November 2022
Accepted 15 February 2023

KEYWORDS
Patient simulator; medical training; spinal immobilization; sensor; prototyping; healthcare design

Introduction

In medical cases where a patient has been subjected to traumatic accidents and needs care and transport from an out-of-hospital location, the neck of the patient is particularly important. Since the neck is less protected than the rest of the spine, avoiding eventual further damage to the spinal cord is crucial, as this could lead to paralysis or death. Cervical spine injuries represent 29% of all injuries to the spinal cord (Domeier et al. 1997). Of these, Theodore et al. (2013) estimate that up to 25% of spinal injuries occur after

CONTACT Marius Auflem mariuauf@stud.ntnu.no Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway

the initial impact, during medical treatment of the patient, or transportation to the hospital. High-quality medical treatment at the trauma site is essential as the mortality rate drops between 48.3 and 79% to 4.4 and 16.7% after the patient is admitted to the hospital (Sekhon et al. 2001). Hence, medical personnel need to be well-versed in handling a suspected spinal injury, and there is a need for adequate equipment to facilitate such training.

Guidelines and training of suspected neck injury treatment and care

If a spinal injury is suspected, the standard procedure has been to apply full immobilization of the patient (White, Domeier, and Millin 2014; ACS-COT 2018). However, several studies have shown the use of supporting equipment such as cervical collars and a backboard can cause problems like increased pain (Papadopoulos et al. 1999), increased intracranial pressure (Lemyze et al. 2011; Davies, Deakin, and Wilson 1996), and ultimately increased mortality rate in trauma cases involving penetrating spinal injuries (Vanderlan, Tew, and McSwain 2009; Haut et al. 2010). Thus, spinal motion restriction, referred to as SMR, has risen as an alternative to traditional spinal immobilization (Swartz et al. 2018). Using SMR, the patient is immobilized either manually or by applying a cervical collar. In this case, a medical professional is responsible for keeping the neck as still as possible during the assessment, treatment, and transportation.

When treating and transporting trauma patients with suspected spinal injuries, it is not evident what movements are considered optimal (or acceptable). Furthermore, which approach (full immobilization or SMR) that accounts for the best results in different scenarios is ambiguous, considering unique patient characteristics, different injuries, and the challenge of obtaining measurable and comparable data in out-of-hospital events.

In training and simulation of trauma scenarios, using standardized patients, and real human markers is often valuable. However, for real people to simulate an unconscious patient, not restricting neck movements, is challenging. Furthermore, performance evaluation and feedback are subjective, as it is based on external observations or experience from the human actor. Therefore, human patient simulators are widely used in training of trauma scenarios, as an alternative to human actors.

Human patient simulators

The use of human patient simulators (or mannequins) in medical training, is an expanding field aiming to replicate clinical scenarios in a safe and repeatable environment (Nehring, Ellis, and Lashley 2001). Several studies have suggested that the transfer of knowledge using human patient simulators, is superior to traditional teaching methods such as interactive case studies (Cant and Cooper 2010; Howard et al. 2010).

For patient simulators to be effective in training, the design needs to enable dextrous hands-on procedures and patient handling that can facilitate transfer of learning back into clinical scenarios (Issenberg et al. 2005). Current patient simulators, however, are considered rigid, and have limited humanlike mobility, range of motion, and tactility. Furthermore, training of pre-hospital treatment is also a field that lacks key performance feedback to trainees, especially regarding suspected neck injuries and neck immobilization.

Objective feedback in medical simulation training

Objective feedback in simulation-based medical training affects the learning outcome and retention of skills for various procedures and medical interventions (McGaghie et al. 2010). One example is the low-dose high-frequency skills training of cardiopulmonary resuscitation (CPR) using mannequins equipped with sensors providing objective feedback and quality improvement metrics (Spooner et al. 2007; Sutton et al. 2011). Feedback should be given during the learning experience, and is an important for retention of skills, and knowledge (Issenberg et al. 2005). Feedback given to trainees could be both formative and summative, but should be supported by objective performance indicators (McGaghie et al. 2010).

Aim and scope

This paper aims to showcase the development and testing of a new neck concept facilitating realistic interactions, as means of training for real patient encounters. The new neck, its functionality, and obtained sensor data will be presented and used to drive a discussion on the applicability of it to be used in medical training. Data obtained from testing will be reviewed and discussed considering trainee performance feedback and quality metrics for patient handling and neck immobilization. The development of a new neck was initiated to explore opportunities for future simulator neck topologies and functionality. As the project is considered a pre-requirement engineering design task, the scope has been to elicit requirements and understand user-needs. Furthermore, the project aim has been to prototype a solution and explore the potential for this to improve (and inform) medical training for suspected spinal injuries.

Prototyping a new neck by Iterative designing, building, and testing cycles

Development process

In the project of creating a new neck, the team utilized a highly iterative and prototype-driven development approach. This was to rapidly generate prototypes that could be tested with users, to gain access to experience-based tacit knowledge and making the user-needs tangible (Auflem, Falch Erichsen, and Steinert 2019). Prototypes were continuously designed, built, and tested to bring answers to open design questions (Ege et al. 2021). When interacting with users and stakeholders, prototypes served as a common ground for both gaining tangible feedback and inform further development (Houde and Hill 1997; Santos et al. 2021).

The following stakeholders were used for consultation through interviews, prototype interactions, and concept testing:

- Paramedics
- 5th year medical students working part time as ambulance personnel
- PhD candidate in neurosurgery

Iterative development of prototypes and testing

The development of the new neck concept can be described as iterative consultation-design-build-test cycles. A timeline, shown in Figure 1, visualizes this process in retrospect, highlighting the milestones being key prototyping activities and iterations.

Initially, the team conducted interviews with potential users regarding the routines and procedures during trauma scenarios and the shortcomings in simulation-based training. This need-finding uncovered the lack of flexibility in current mannequins to be a big limitation. It was stated that current simulators are generally not suited to the required interactions during trauma scenarios, due to limited range of motion and compliance. This was show-cased by movements and handling of a widely used trauma simulator. It was noted that care and handling of suspected neck injuries are an important part of training, and that a flexible neck would enable spinal immobilization to be simulated more realistically.

Based on the insights uncovered through need-finding, the first prototype, *iteration 1,* was created using a biomimetic approach. By stacking rigid discs and silicone bushings in an alternating pattern, the movement could more closely imitate that of the human neck anatomy. Suspending the discs by spring-loaded wires, the stiffness of the neck could also be altered from selection and pre-tensioning of the springs. This allowed for a similar range of motion to a human neck. This was uncovered by empirical testing, using an existing mannequin head to observe and compare prototype performance with real human head movement and anatomical constraints.

In *Iteration 2,* replacing the flat discs with 3D printed elements, the bending radius, and degrees of freedom could be restricted by the topology of each disc element. During user-testing of this prototype, it was noted the

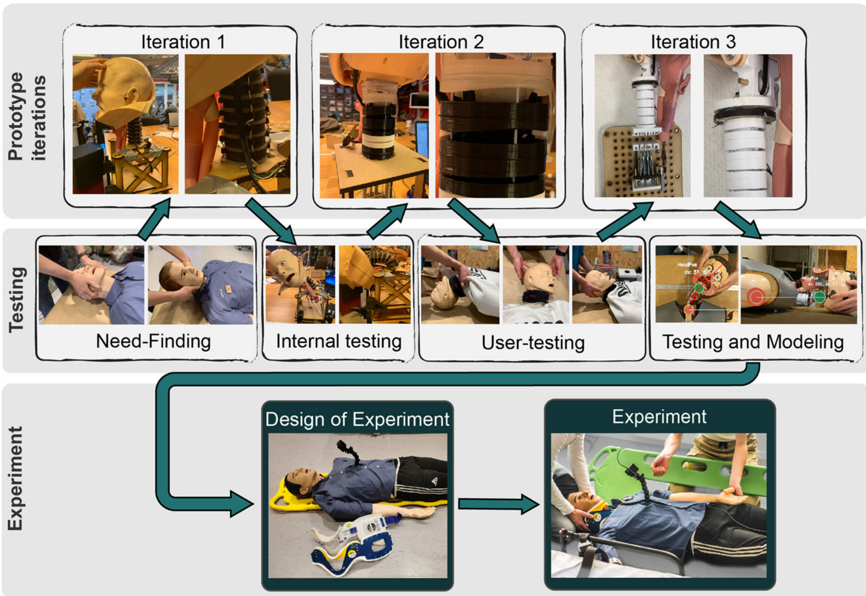


Figure 1. Prototyping activities in the project showing the iterations, the tests performed, design of experiment, and performing the final user-test. The setup, design and execution of experiment (dark boxes) will be further described in following sections.

looseness of the neck could resemble an unconscious patient, and that the lateral flexion felt realistic. However, the flexion and extension were missing, as the joint between the top vertebra and skull was not considered in the prototype. It was also noted the rotation of the head is mainly observed in this joint. Furthermore, an important finding during this user-testing was the lack of objective measurements and training performance indicators when it comes to suspected neck injuries. While it is evident that mitigating movements and strain to the neck is important, what type and magnitude of movement that could harm or cause further injury is not well understood nor researched.

Iteration 3 was designed to substitute the existing neck assembly in (but not restricted to) the SimMan3G from Laerdal Medical, a commonly used medical simulator (Shinnick and Woo 2013). Figure 2 shows the conceptual prototype and the different components making up the neck assembly. Four wires are led through the discs, evenly spaced in the four main directions of motion. Each wire is coupled in series with a spring, and all springs are pre-tensioned to the same level. By moving the head, the springs will either have a positive or negative relative displacement.

To address the lack of objective measurements and performance indicators, each spring was connected to a sliding potentiometer. The joint between the neck and head was equipped with rotary potentiometers at the

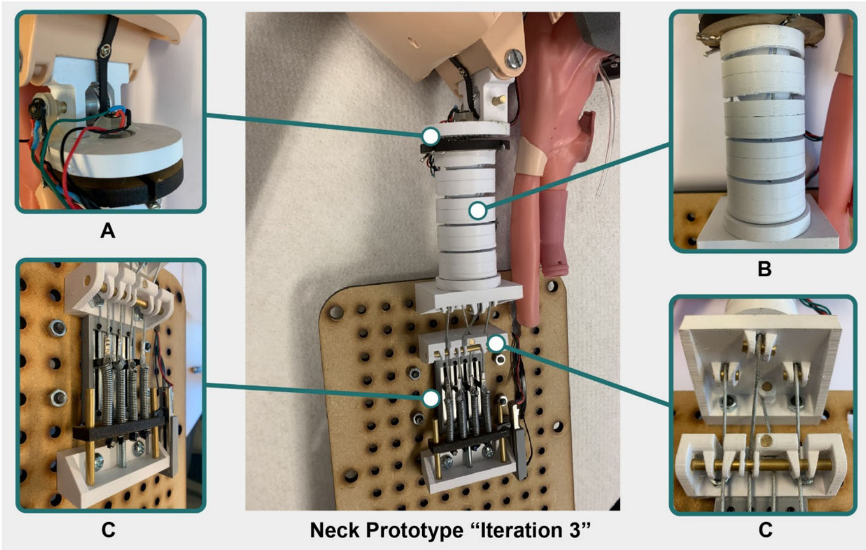


Figure 2. The design of the conceptual prototype (iteration 3). (A) Rotation and tilt recorded by rotary potentiometers. (B) 3D printed discs emulating the vertebrae. (C) Wires coupled in series with springs, with sliding potentiometers measuring the spring displacement. (D) Pulley system for wire routing.

rotational and tilting axis. This way, the spine's relative movement and the head's relative rotation and tilt can be recorded.

Modelling neck behaviour

Continuum robotics principles allows for modelling of suspended element structures, such as the neck in the developed prototype. This has been shown in a variety of applications such as artificial muscles (Pritts and Rahn 2004), surgery equipment (Chen, Pham, and Redarce 2009; Kato et al. 2015), and artificial human fingers (Suzumori, Iikura, and Tanaka 1992). Despite robots having different coordinate frames and analytical formalisms, Webster and Jones (2010) describe theoretical modelling that apply to all continuum robotics that can be assumed to express piecewise constant curvature (Rao et al. 2020).

However, the design of the neck prototype is not ideal and hence, relying solely on the ideal model is not feasible. Thus, empirical measurements using video tracking was performed, and the results were compared to the ideal model. The main objective was to correlate the spring displacements to the spine's angle, to track the neck and head's motion during handling.

Using OpenCV, as seen in Figure 3(A), a computer vision library and toolkit, a local coordinate system (green marker) was created relative to the red marker on the torso, with the x-axis (y-axis for lateral tilt) aligned with the green marker. The relative displacement of the green marker was then

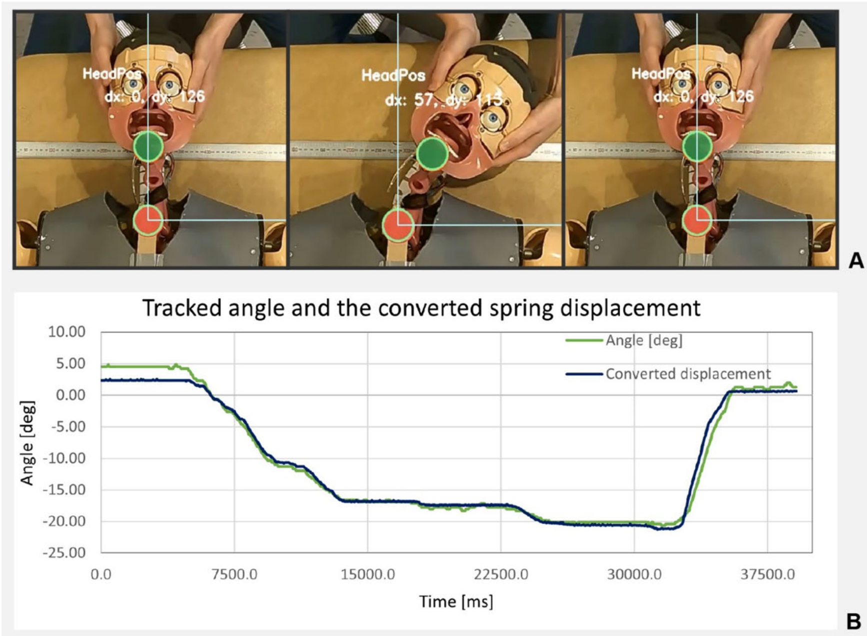


Figure 3. Tracking frames for the lateral flexion (A) and the corresponding graph (B) showing both the tracked angle and the converted displacement from the sensor readings. The same approach was repeated for the other directions (and spring displacements).

measured accurately based on the number of pixels it moved while moving the head to its extremes. Synchronizing the sensor readings with the derived angles from the video tracking can be seen in Figure 3. By dividing the tracked angles by the measured spring displacements, correlation factors could be derived for the different degrees of freedom. The angles from the spring displacements represent the frontal and sagittal components of the angular motion of the neck.

A linear correlation factor was also found for the relation between the spring displacement and the top of the spine's deviation from the neutral axis. The correlation factor for each spring, for both angular and distance conversion, is shown in Table 1.

Design of experiment

To maximize learning potential from the developed prototype, a structured experiment was designed. Based on mixed-method research, both objective and subjective data was captured and analyzed. The aim of the experiment was to, firstly, gain subjective evaluations of the performance and characteristics of the prototype, and secondly, to gather and analyze objective data on simulated patient handling and spinal immobilization techniques.

Table 1. Correlation factors for the angular and translational displacement for the four directions of movement.

Motion and affected spring	Correlation factor [degree/mm]	Correlation factor [mm/mm]
Backwards tilt (Front spring)	1.4402	4.53070
Forward tilt (Back spring)	1.0793	5.40365
Lateral deflection (Right spring)	1.7438	3.40950
Lateral deflection (Left spring)	1.7438	3.40950

Participants

Twelve users with relevant background and experience participated in the experiment. Each participant's gender, general age, occupation, and years of clinical experience is stated in Table 2. All participants gave informed consent to the results being used in analysis and publication.

Table 2. Overview of participant demographics.

Participant	Gender	Age [years]	Occupation	Clinical experience [years]
Participant 1	Male	30–39	Ambulance personnel	10
Participant 2	Female	30–39	Ambulance personnel	10
Participant 3	Male	30–39	Medical doctor	6
Participant 4	Male	21–29	Ambulance personnel	2
Participant 5	Female	21–29	Ambulance personnel	1
Participant 6	Female	21–29	Ambulance personnel	1
Participant 7	Male	21–29	Ambulance personnel	1
Participant 8	Male	21–29	Student/ambulance	1
Participant 9	Male	21–29	EMT	7
Participant 10	Female	21–29	Student/ambulance	5
Participant 11	Male	21–29	Student/ambulance	1
Participant 12	Male	21–29	Student/ambulance	2

Testing protocol

Using the neck prototype (*iteration 3)* previously described, the participants performed three tests to evaluate the prototype during trauma scenarios, its validity, and realism. The testing protocol consisted of:

- Simulation of trauma scenario requiring spinal immobilization.
- Free interaction with the prototype for feedback.
- Post-test questionnaire and Likert scale.

Trauma scenario simulation

The participants were introduced to a simulated scenario where an unconscious patient was lying on the ground wearing a ski helmet, as shown in Figure 4. It was emphasized the objective was to immobilize the neck to prevent further spinal injuries. In pairs, they were asked to:

1. Remove the ski-helmet.
2. Immobilize the patient.
3. Perform a lift to transfer the patient to a bed placed a couple of metres away.

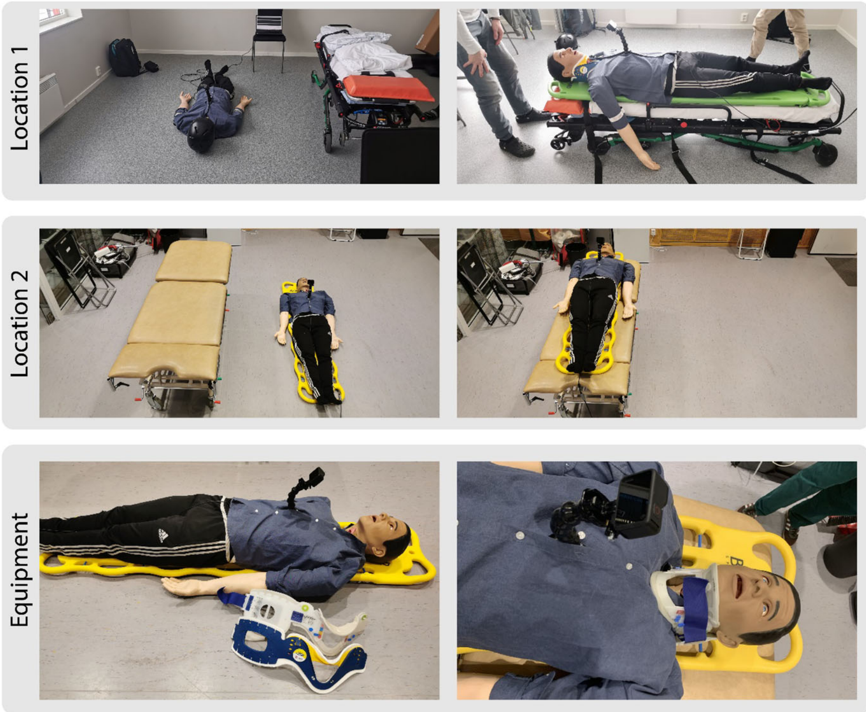


Figure 4. The locations used for the user experiments and the utilized equipment for performing spinal immobilization and patient transfer.

This procedure was performed twice. Firstly, it was done without any aiding equipment, with full manual immobilization during the lift. Secondly, the participants were required to use a backboard and a cervical collar. The aim was to gather quantitative data from the sensors to evaluate the quality and resolution obtained given the different scenarios. A camera mounted on the mannequin torso, also recorded the handling of the head to be able to compare it to the movement measured by the sensors. User tests 1–5 were performed at location 2, and user test 6 was performed at location 1, shown in Figure 4.

Free interaction with the prototype

The participants were asked to freely interact with the neck and compare the movement to an unconscious patient in terms of the range of motion and tactile experience. Feedback was gathered through free dialogue while participants interacted with the prototype. The aim was to obtain insights on the viability of the prototype in trauma simulations, and the realism of the prototype's tactility and range of motion.

Questionnaire

All participants filled out a questionnaire, collecting demographical information and years of clinical experience. A Likert scale was part of the form, consisting of five statements the participants rated from 1 to 5, where numbers one to five meant 'Strongly disagree', 'Disagree', 'Neutral', 'Agree', and 'Strongly Agree', respectively. The questionnaire gathered feedback concerning specific aspects of the prototype by the statements listed below:

1. The neck feels like an unconscious patient's neck.
2. The patient lift from the ground to the stretcher felt realistic.
3. The weight (of the head) feels realistic.
4. The looseness of the neck feels realistic.
5. The range of motion of the neck feels realistic.

MotionScore

There is a lack of benchmarks and/or established formulations of how much (and what) movement is considered critical during handling of patients with spinal injuries (Swartz et al. 2018). Therefore, we utilize an assumption that large deviations from a stable and neutral position, are worse than keeping the head still (which is the formal instruction for handling patients with traumatic neck injuries). A score for evaluating relative movement during testing was derived based on this assumption, called the MotionScore (MS). Scoring of head movement has also been utilized in other studies such as (Nolte et al. 2021), which proposed a similar metric to measure the relative angle between the torso and the head. The equation for calculating the MotionScore is given in Equation 1. The angular motion, denoted as ω and given in degrees, is the sum of the angular motion in the frontal direction, sagittal direction, rotation, and tilting of the head, at a given moment during the simulation. The angular motion can be used to evaluate the neck's relative movement over time by summing up the angular motion at each time-point, obtaining ω_{tot}, dividing it by the length of the interval in seconds, and multiplying it by 0.01 [sec/degree]. The result is the MotionScore, a dimensionless value that represents the area underneath the displacement-time graphs obtained during simulated events.

$$MS = \frac{\omega_{tot}}{t} \times 0.01 \qquad (1)$$
$$\text{MotionScore (MS)}$$

Momentaneous spine deviation (MSD)

To detect and analyze sudden and large movement of the neck during simulations, a performance metric was developed, hereafter referred to as the momentaneous spinal deviation, or MSD. This metric considers the distance between the top of the spine and the natural axis of the neck. We utilize the distance conversion factor of the spring displacements, ending up with the spinal deviation throughout the simulation. The time-derivative of the resulting distance-time graph provides the momentaneous spinal deviation at each point in time during simulation events.

Results

Sensor raw data

To visualize the objective data recorded from the simulations we look at the tracked angles along the timeline of the event. As an example, we have plotted the results from user test 6 showing both testing conditions of the scenario, being patient handling and transport with and without equipment. Furthermore, in addition to the data from both conditions, we have extracted descriptive still frames from the camera pointing at the mannequin face. The frames are correlated to events of interest throughout the scenario, and the timing is indicated where the event appear on the graph. The events for the no use of equipment condition are the start position for the simulation, the removal of the ski helmet, a 'fork-grip' supporting the head and lifting the back, the patient lift, releasing the patient from the 'fork-grip', and the end position of the simulation. The same procedure was done for the condition requiring the participants to utilize a cervical collar and a backboard. The events mapped from the still frames are the start position of the simulation, the removal of the ski helmet, applying the cervical collar, log-rolling the patient onto the backboard, the patient lift, and the end position of the simulation. The recorded data, event frames, and mapping of events onto the graphs is shown in Figure 5.

Likert Scale results and feedback from free interaction discussion

From the questionnaire, the results of the Likert Scale are presented in Table 3. The table shows the averages and standard deviations of the twelve participants' ratings on the Likert scale.

From freely handling and examining the functionality of the prototypes the participants provided feedback and comments. These are presented below in the form of quoted comments, and summarized findings obtained from general discussions surrounding the presented simulator concept:

- '[…] when patients are unconscious, and the muscles are relaxed, the head and neck are really loose, looser and more flexible than people think. This neck captures this aspect well.'
- '[…] remarkably better than the other simulators we have used.'

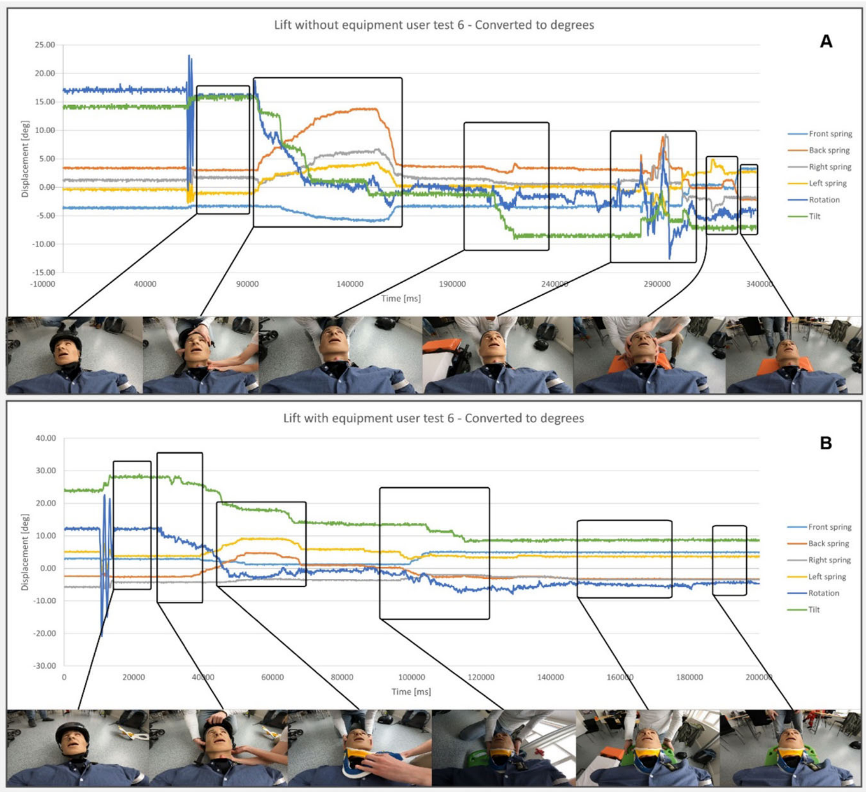


Figure 5. Recorded movement plotted from user test 6 in both without (A) and with (B) equipment conditions. Still frames are synchronized, and boxes are superimposed onto the relevant graph sections.

Table 3. Average results from the questionnaire, $n = 12$.

Question	Statement	Average rating	SD
1	The neck feels like an unconscious patient's neck	3.83	0.69
2	The patient lift from the ground to the stretcher felt realistic	2.42	1.04
3	The weight (of the head) feels realistic	3.42	1.19
4	The looseness of the neck feels realistic	3.92	1.11
5	The range of motion of the neck feels realistic	3.83	0.90

- 'It feels like it is missing an aspect of the natural resistance from the muscles in the neck.'
- 'The largest flexion is a bit too large.'
- 'The head falls naturally to the sides as it would with a real patient.'

It was stated that the status of patients is only measured properly from arrival at the hospital; pre-hospital stabilization is not scored or measured well currently. The potential of measuring how different immobilization techniques affect the neck in the pre-hospital scenarios was interesting and desirable. The body (meaning torso, arms, and legs) affects the neck more than it would in a real trauma scenario, due to the torso being too stiff.

Feedback and performance metrics

The ability to provide objective feedback during simulation of dextrous procedures is important, as this could enhance psychomotor skills from repeatedly performing procedures correctly. In suspected spinal injuries, it is important to keep the spine as still as possible. even though a specific limit for motion leading to further damage to the spine currently does not exist (Swartz et al. 2018; Gerling et al. 2000). Hence, large and sudden movements are undesirable and should be avoided during spinal immobilization, and handling. The momentaneous spinal deviation (MSD) metric is applicable to highlight such movements during simulation scenarios. This is exemplified by looking at both the testing conditions of user test 1. Large and sudden movements of the head are identifiable by applying the MSD, and both the magnitude and timing of the sudden movements are shown in Figure 6.

As there does not exist any threshold for an MSD value where further damage to the spine is likely (Gerling et al. 2000), the threshold can be set dynamically to vary the difficulty and scope of the exercise depending on the scenario or experience level of the trainee. An example of an MSD threshold of ±0.3 mm/sec applied to the recording of user test 4 without equipment is shown in Figure 7.

The participants in user test 4 exceeded the set threshold at 74, 104, and 107 seconds into the simulation. By retrieving a pair of still frames from these moments, the causality could be investigated. For the moments investigated in this test, the movements exceeding the threshold can be linked to the events of removing the ski-helmet (74 sec.), beginning of patient lift

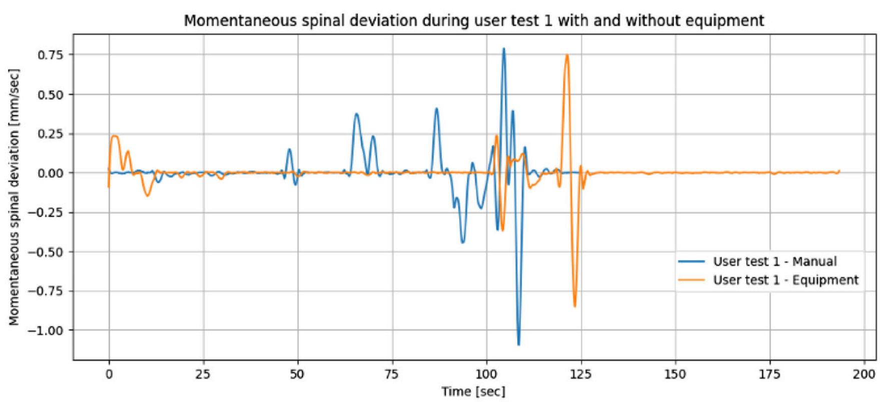


Figure 6. MSD of user test 1, with and without equipment.

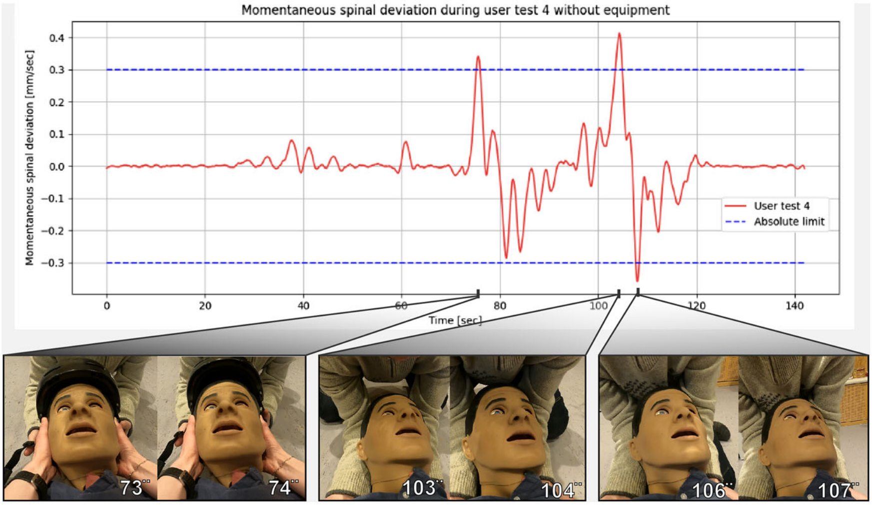


Figure 7. MSD of user test 4 without equipment, with a threshold set to ±0.3 mm/sec. Still frames from the events occurring when the threshold is exceeded are superimposed.

(104 sec.) and during lift (107 sec.). The MSD for all the tests both using and without using aiding equipment are shown in Figure 8.

Debrief and quality metrics

Using the MotionScore previously established, the six user tests can be analyzed and presented. During testing, however, the front spring malfunctioned and did not provide correct readings from user test 2 and onwards. The MotionScore for the rest of the user-tests disregards the front spring and considers the back spring instead, as it is the inverse of the front spring for a certain deflection range (within limits of what the prototype experiences during immobilization and careful handling).

To compare the neck movement between each user test, the MotionScore-analysis starts when the participants have positioned the mannequin to look straight forward with the helmet on. The analysis ends after the mannequin has been transported to the stretcher at rest. Using no equipment, user-test 1 registered an MS of 5.58. Meanwhile, the MS was 3.18 when a backboard and cervical collar were utilized. The MotionScore for each test is listed in Table 4, along with the mean angular deviation and standard deviation.

Discussion

Viability of the prototype

For users to leverage simulation-based training to transfer learning into clinical scenarios, simulators need to be sufficiently realistic. The results from the

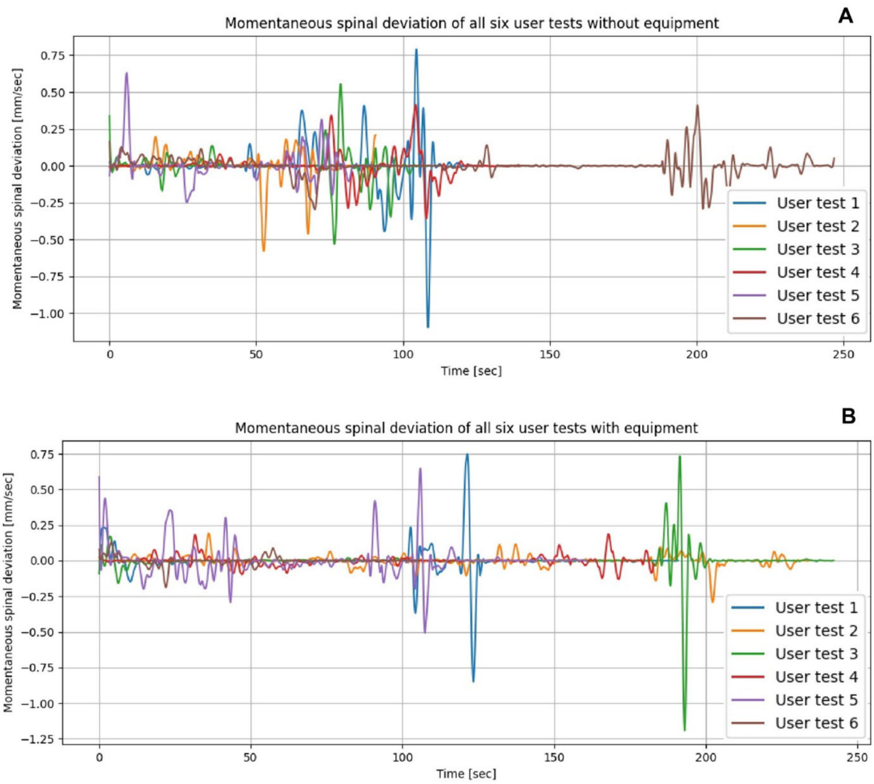


Figure 8. The MSD of all six user tests. Figure A shows the tests performed with equipment and B shows the tests performed without equipment.

Table 4. MotionScore and mean angular deviation and standard deviation of the tracked simulation events.

	MotionScore		Mean angular deviation and standard deviation	
User-test #	Without cervical collar and backboard	With cervical collar and backboard	Without cervical collar and backboard	With cervical collar and backboard
1	5.58	3.18	30.74 (8.67)	17.51 (4.70)
2	8.40	3.03	46.19 (10.40)	16.66 (7.09)
3	5.67	2.17	31.21 (15.80)	11.93 (6.84)
4	3.62	3.60	19.91 (7.68)	19.80 (8.19)
5	4.95	2.42	27.23 (9.63)	13.29 (8.72)
6	2.84	2.19	15.60 (6.42)	12.07 (3.23)

Likert scale and feedback during the free interaction point out apparent limitations of the current prototype. Aspects of the prototype did not facilitate realistic motion when lifting the patient as the results from statement 2 on the Likert scale clearly show. Pointed out was the torso, arms and legs, being both rigid and light compared to real patient. The torso was also anatomically incorrect, making the interaction less realistic. This furthermore led to the cervical collar not having a defined collar bone to rest against, thus affecting the support it provided the head. It was also noted the neck did not feel as naturally resistant to rotation, which is most likely due to the neck not being supported by tissue and compliant structures other than the spine and windpipe.

In terms of feedback concerning the included functionality of the neck prototype, the overall results were positive. Concerning the compliance and range of motion of the neck, it was considered adequate to what they would expect from a real patient. This corresponds well with the results from the Likert Scale, where these statements scored highly. The range of flexion was described as too large, and while the range should be adjusted, this does not affect the results from the experiment where the aim was to restrict the motion as much as possible. Hence, we see the presented prototype as a viable application in emergency care training of spinal injuries and should be developed further.

Possibilities for evaluating and comparing performance

Reviewing the obtained sensor data from the experiments, the spine's position, orientation, and the head's rotation could be further utilized depending on the objectives of the training scenario. The two metrics proposed in this study could enable trainees to get objective feedback on performance in pre-hospital immobilization and transportation training. This area of trauma care does not currently have any measurable parameters except for retrospective studies comparing the mortality and paralysis rates of real patients. Thus, having the opportunity to investigate simulated interventions through objective measures such as the movement data described in this paper could be of great value, also for evaluating and determining best practice.

The momentaneous spinal deviation introduces a direct and quantifiable performance metric for feedback during the interaction, and analysis of the event for debrief purposes. While the least amount of sudden and large motion of the neck to exacerbate spinal injuries is not known, it is argued that the general standard should be as little movement as possible (Gerling et al. 2000; Swartz et al. 2018). The MSD values shown in this paper are applicable for measuring the amount of relative motion a simulated patient endures during simulation. Hence, a threshold for allowed motion in training scenarios could be determined empirically by further research using the presented concept. By obtaining a large data set from highly skilled medical personnel, a threshold for acceptable relative motion of the head could be determined. This could enable a feedback response (audible, visual, or verbal) to be given trainees if the allowed threshold is exceeded. Such a threshold could, moreover, be tailored to specific trauma scenarios, patient characteristics, or the skill level of the participants. Hence, objective feedback and tailored training for managing spinal injuries could be enabled.

The suggested performance metric MotionScore, enables comparisons between entire immobilization and transportation scenarios. Like the MSD, a MS threshold for excessive motion is not defined as a result of lacking empirical data and clinical standards. However, in this study, the MS has been used to capture the difference between performing spinal immobilization with and without aiding equipment with regards to limiting relative head motion. The participants were not asked to perform the immobilization to be evaluated on their performance. Thus, the results presented for the user-test are not valid for evaluating the use of cervical collars and backboards compared to manual handling. The results do, however, call for further investigation as they all suggest a difference favouring the use of aiding equipment, but only in terms of limiting the movement of the spine and head.

Even though the data collected from the prototype could be used as valuable performance indicators in both formative and summative feedback, it can only describe the prototype's motion (McGaghie et al. 2010). The prototype does not capture the external factors of the interaction between the trainee and simulator, like the number of people partaking in the simulation and the techniques used. The performance indicators must be applied in a way that reflects the simulation case and the relevant learning objectives (Issenberg et al. 2005). This calls for further research on the implications of utilizing the concept in medical training, learning outcome, and simulated events.

On leveraging sensor data

From visual inspection, the sensor data clearly indicate parameters describing actions performed by participants. Not restricted to the case of spinal immobilization, the angular motion of the spine can be leveraged for a broad range of use-cases. For example, exploring how users position the head during simulated patient interactions could be further investigated. An example could be airway management training, where the users can get direct feedback on how the tilt and position of head correlate to an ideal position ensuring free airways.

Further exploration should also include better feedback during trauma scenario simulations, where there is currently a lack of objective measurements. Both the pre-hospital patient treatment and patient transportation to the hospital could benefit from measuring how different techniques and procedures affect the spine of the patients in terms of the motion it endures. Hence, the results presented in this study is merely a snapshot of the possibilities. Further research concerning medical validity is however required, but we suggest the concept described in this paper as a potential tool to facilitate such studies.

Insights on developing patient simulators

The development and testing of the prototype described in this study has shown the importance of purposeful involvement of experienced users at various stages of the project. For a patient simulator to be applicable in medical training, the functionality, fidelity, and clinical realism should be collectively sufficient to facilitate relevant and realistic interactions. To ensure the simulator correlates with these aspects, the users have been involved at critical junctions throughout the development process. To maximize learning from interactions with users, they should be made to target specific aspects of prototypes, by addressing specific design questions. This is exemplified by prototyped functionality, interactions, look or feel, and how these aspects could be improved. This presupposes the prototypes are made with the right intent and audience in mind. A prototype in the early product development phase should define itself as a potential answer to a question related to how the final product should perform. Thus, a user can be presented with the prototype and validate whether the function or behaviour etc., is correct or not and explain why. This interaction either validates the solution prototyped or provides valuable insight the developer often would not be able to acquire independently or as quickly.

The development of medical simulators, where there is often a substantial knowledge gap on the developer’s part, can benefit from utilizing the insights mentioned earlier regarding continuous user-involvement to ensure the final product aligns with the user-needs and requirements.

Conclusion

This paper has presented the development and testing of a new neck concept for human patient simulators to be used in trauma scenario training. Shortcomings in current patient simulators and the importance of training for suspected spinal injury scenarios have been identified. Training for these scenarios using simulation, the treatment and immobilization of the neck are important aspects to consider. Especially since the current solutions lack the realistic range of motion, compliance, and objective performance indicators for improving training and facilitating learning.

The team deployed an iterative and prototype-driven development approach by relying on frequent consultations with users. This enabled the solutions to align with users’ expectations and experience of handling an unconscious patient, as well as probing the needs for a solution to fit medical guidelines and current curricula. A proposed concept aims to replicate the behaviour of the neck of an unconscious patient by mimicking the human anatomy and vertebrae. By utilizing several sensors, the prototype also measures the relative motion it is subjected to during interventions and handling.

By performing experimental user-tests with medical professionals and students ($n = 12$), the interactions and handling of the prototype have been assessed qualitatively and quantitatively. The conceptual neck was rated by the participants to be a good representation of an unconscious patient's neck in terms of the looseness of the head and the range of motion during handling. Furthermore, the recorded motion of the spine and head during the simulations have been used to suggest performance metrics assessing the quality of handling and provide corrective feedback during the simulations. The generated data is analyzed and discussed as potential quality performance indicators for both formative and summative feedback in simulation-based trauma training.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research is supported by the Research Council of Norway (RCN) through its industrial Ph.D funding scheme, project number 290404.

Notes on contributors

Emil Matias Henriksen has a master's degree in mechanical engineering from the Department of Mechanical and Industrial Engineering at the Norwegian University of Science and Technology. His master's thesis is focused on early-stage development of medical training equipment. His research is part of TrollLabs, a prototyping and product development lab and community at NTNU in Trondheim, Norway.

Marius Auflem works at Laerdal Medical as an industrial PhD candidate through the Department of Mechanical and Industrial Engineering at the Norwegian University of Science and Technology. Marius does research in early-stage product development in the context of medical training and simulation equipment. His research is part of TrollLabs, a prototyping and product development lab and community at NTNU in Trondheim, Norway.

Martin Steinert is a professor of engineering design and innovation in the Department of Mechanical and Industrial Engineering at the Norwegian University of Science and Technology (NTNU). He teaches fuzzy front-end engineering for radical new product/service/system concepts and graduate research seminars for PhD students engaged in topics related to new product design and development. His various research projects are usually multidisciplinary (ME/CS/EE/Neuro- and cognitive science) and are often connected with the industry. The aim is to uncover, understand, and leverage early-stage engineering design paradigms with a special focus on human-machine/object interactions.

ORCID

Marius Auflem http://orcid.org/0000-0002-9297-2645
Martin Steinert http://orcid.org/0000-0002-8366-0201

References

ACS-COT. 2018. *ATLS Advanced Trauma Life Support: Student Course Manual.*. Chicago: American College of Surgeons.

Auflem, Marius, Jorgen Falck Erichsen, and Martin Steinert. 2019. "Exemplifying Prototype-Driven Development through Concepts for Medical Training Simulators." Procedia CIRP, 29th CIRP Design Conference 2019, 08–10 May 2019, Póvoa de Varzim, Portgal, vol. 84, 572–578.

Cant, Robyn P., and Simon J. Cooper. 2010. "Simulation-Based Learning in Nurse Education: Systematic Review." *Journal of Advanced Nursing* 66 (1): 3–15. doi:10.1111/j.1365-2648.2009.05240.x.

Chen, G., M. T. Pham, and T. Redarce. 2009. "Sensor-Based Guidance Control of a Continuum Robot for a Semi-Autonomous Colonoscopy." *Robotics and Autonomous Systems* 57 (6): 712–722. doi:10.1016/j.robot.2008.11.001.

Davies, G., C. Deakin, and A. Wilson. 1996. "The Effect of a Rigid Collar on Intracranial Pressure." *Injury* 27 (9): 647–649. doi:10.1016/S0020-1383(96)00115-5.

Domeier, Robert M., Rawden W. Evans, Robert A. Swor, Edgardo J. Rivera-Rivera, and Shirley M. Frederiksen. 1997. "Prehospital Clinical Findings Associated with Spinal Injury." *Prehospital Emergency Care* 1 (1): 11–15. doi:10.1080/10903129708958777.

Ege, Daniel Nygård, M. Auflem, O. Lilleløkken, and M. Steinert. 2021. "User Involvement in Early-Stage Design of Medical Training Devices – Case of a Palpation Task Trainer Prototype." *Design for Health* 5 (2): 214–232. doi:10.1080/24735132.2021.1980270.

Gerling, Michael C., Daniel P. Davis, Robert S. Hamilton, Gabrielle F. Morris, Gary M. Vilke, Steven R. Garfin, and Stephen R. Hayden. 2000. "Effects of Cervical Spine Immobilization Technique and Laryngoscope Blade Selection on an Unstable Cervical Spine in a Cadaver Model of Intubation." *Annals of Emergency Medicine* 36 (4): 293–300. doi:10.1067/mem.2000.109442.

Haut, Elliott R., Brian T. Kalish, David T. Efron, Adil H. Haider, Kent A. Stevens, Alicia N. Kieninger, Edward E. Cornwell, III, and David C. Chang. 2010. "Spine Immobilization in Penetrating Trauma: More Harm than Good?" *Journal of Trauma and Acute Care Surgery* 68 (1): 115–121. doi:10.1097/TA.0b013e3181c9ee58.

Houde, Stephanie, and Charles Hill. 1997. "Chapter 16 – What Do Prototypes Prototype?" In *Handbook of Human-Computer Interaction*, edited by Marting G. Helander, Thomas K. Landauer, and Prasad V. Prabhu, 367–381. Amsterdam: North-Holland.

Howard, Valerie Michele, Carl Ross, Ann M. Mitchell, and Glenn M. Nelson. 2010. "Human Patient Simulators and Interactive Case Studies: A Comparative Analysis of Learning Outcomes and Student Perceptions." *CIN: Computers, Informatics, Nursing* 28 (1): 42–48. doi:10.1097/NCN.0b013e3181c04939.

Issenberg, S. Barry, William C. Mcgaghie, Emil R. Petrusa, David Lee Gordon, and Ross J. Scalese. 2005. "Features and Uses of High-Fidelity Medical Simulations That Lead to Effective Learning: A BEME Systematic Review." *Medical Teacher* 27 (1): 10–28. doi:10.1080/01421590500046924.

Kato, Takahisa, Ichiro Okumura, Sang-Eun Song, Alexandra J. Golby, and Nobuhiko Hata. 2015. "Tendon-Driven Continuum Robot for Endoscopic Surgery: Preclinical

Development and Validation of a Tension Propagation Model." *IEEE/ASME Transactions on Mechatronics* 20 (5): 2252–2263. doi:10.1109/TMECH.2014.2372635.

Lemyze, Malcolm, Aurore Palud, Raphael Favory, and Daniel Mathieu. 2011. "Unintentional Strangulation by a Cervical Collar after Attempted Suicide by Hanging." *Emergency Medicine Journal* 28 (6): 532–532. doi:10.1136/emj.2010.106625.

McGaghie, William C., S. Barry Issenberg, Emil R. Petrusa, and Ross J. Scalese. 2010. "A Critical Review of Simulation. Based Medical Education Research: 2003–2009." *Medical Education* 44 (1): 50–63. doi:10.1111/j.1365-2923.2009.03547.x.

Nehring, Wendy M., Wayne E. Ellis, and Felissa R. Lashley. 2001. "Human Patient Simulators in Nursing Education: An Overview." *Simulation & Gaming* 32 (2): 194–204. doi:10.1177/104687810103200207.

Nolte, Philip Christian, Davut Deniz Uzun, David Häske, Jeronimo Weerts, Matthias Münzberg, Adrian Rittmann, Paul Alfred Grützner, and Michael Kreinest. 2021. "Analysis of Cervical Spine Immobilization during Patient Transport in Emergency Medical Services." *European Journal of Trauma and Emergency Surgery* 47 (3): 719–726. doi:10.1007/s00068-019-01143-z.

Papadopoulos, M. C., A. Chakraborty, G. Waldron, and B. A. Bell. 1999. "Exacerbating Cervical Spine Injury by Applying a Hard Collar." *British Medical Journal* 319 (7203): 171–172. doi:10.1136/bmj.319.7203.171.

Pritts, M. B., and C. D. Rahn. 2004. "Design of an Artificial Muscle Continuum Robot." In IEEE International Conference on Robotics and Automation, 2004, Proceedings ICRA 2004, New Orleans, LA, USA, vol. 5, 4742–4746. DOI: 10.1109/ROBOT.2004.1302467.

Rao, Priyanka, Quentin Peyron, Sven Lilge, and Jessica Burgner-Kahrs. 2020. "How to Model Tendon-Driven Continuum Robots and Benchmark Modelling Performance." *Frontiers in Robotics and AI* 7: 630245. 10.3389/frobt.2020.630245.

Santos, Thiago Bertolini dos., Carina Campese, Ricardo M. Marcacini, Roberta A. Sinoara, Solange O. Rezende, and Janaina Mascarenhas. 2021. "Prototyping for User Involvement Activities: How to Achieve Major Benefits." *CIRP Journal of Manufacturing Science and Technology* 33: 465–472. doi:10.1016/j.cirpj.2021.04.013.

Sekhon, H., S. Lali, and Michael G. Fehlings. 2001. "Epidemiology, Demographics, and Pathophysiology of Acute Spinal Cord Injury." *Spine* 26 (24S): S2.

Shinnick, Mary Ann, and Mary A. Woo. 2013. "The Effect of Human Patient Simulation on Critical Thinking and Its Predictors in Prelicensure Nursing Students." *Nurse Education Today* 33 (9): 1062–1067. doi:10.1016/j.nedt.2012.04.004.

Spooner, Brendan B., Jon F. Fallaha, Laura Kocierz, Christopher M. Smith, Sam C. L. Smith, and Gavin D. Perkins. 2007. "An Evaluation of Objective Feedback in Basic Life Support (BLS) Training." *Resuscitation* 73 (3): 417–424. doi:10.1016/j.resuscitation.2006.10.017.

Sutton, Robert M., Dana Niles, Peter A. Meaney, Richard Aplenc, Benjamin French, Benjamin S. Abella, Evelyn L. Lengetti, Robert A. Berg, Mark A. Helfaer, and Vinay Nadkarni. 2011. "Low-Dose, High-Frequency CPR Training Improves Skill Retention of in-Hospital Pediatric Providers." *Pediatrics* 128 (1): e145–51. doi:10.1542/peds.2010-2105.

Suzumori, K., S. Iikura, and H. Tanaka. 1992. "Applying a Flexible Microactuator to Robotic Mechanisms." *IEEE Control Systems Magazine* 12 (1): 21–27. doi:10.1109/37.120448.

Swartz, Erik E., W. Steven Tucker, Matthew Nowak, Jason Roberto, Amy Hollingworth, Laura C. Decoster, Thomas W. Trimarco, and Jason P. Mihalik. 2018. "Prehospital Cervical Spine Motion: Immobilization versus Spine Motion Restriction." *Prehospital Emergency Care* 22 (5): 630–636. doi:10.1080/10903127.2018.1431341.

Theodore, Nicholas, Mark N. Hadley, Bizhan Aarabi, Sanjay S. Dhall, Daniel E. Gelb, R. John Hurlbert, Curtis J. Rozzelle, Timothy C. Ryken, and Beverly C. Walters. 2013. "Prehospital

Cervical Spinal Immobilization after Trauma." *Neurosurgery* 72 (suppl_3): 22–34. doi:10.1227/NEU.0b013e318276edb1.

Vanderlan, Wesley B., Beverly E. Tew, and Norman E. McSwain. 2009. "Increased Risk of Death with Cervical Spine Immobilisation in Penetrating Cervical Trauma." *Injury* 40 (8): 880–883. doi:10.1016/j.injury.2009.01.011.

Webster, Robert J., and Bryan A. Jones. 2010. "Design and Kinematic Modeling of Constant Curvature Continuum Robots: A Review." *The International Journal of Robotics Research* 29 (13): 1661–1683. doi:10.1177/0278364910368147.

White, Chelsea C., IV, Robert M. Domeier, and Michael G. Millin. 2014. "EMS Spinal Precautions and the Use of the Long Backboard –Resource Document to the Position Statement of the National Association of EMS Physicians and the American College of Surgeons Committee on Trauma." *Prehospital Emergency Care* 18 (2): 306–314. doi:10.3109/10903127.2014.884197.

i1

On Iterative Prototyping, Testing, and Learning in New Product Development Projects - Accelerate Concept Development through Internal Prototyping Workshops

Internal Product Development Paper Distributed in Laerdal

Marius Auflem, Sindre Hegglund-Dalseth, Arild Eikefjord (2020),

This paper is not included due to copyright restrictions

ISBN 978-82-326-7136-6 (printed ver.)
ISBN 978-82-326-7135-9 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology