Understanding Innovation

Christoph Meinel
Larry Leifer *Editors*

Design Thinking Research

Translation, Prototyping, and Measurement

Springer

Understanding Innovation

Series Editors

Christoph Meinel, Potsdam, Germany

Larry Leifer, Stanford, USA

"Everyone loves an innovation, an idea that sells." Few definitions of innovation are more succinct. It cuts to the core. Yet in doing so, it lays bare the reality that selling depends on factors outside the innovation envelope. The "let's get creative" imperative does not control its own destiny. Expressed another way, in how many ways can we define innovation? A corollary lies in asking, in how many ways can the innovative enterprise be organized? For a third iteration, in how many ways can the innovation process be structured? Now we have a question worth addressing. "Understanding Innovation" is a book series designed to expose the reader to the breadth and depth of design thinking modalities in pursuit of innovations that sell. It is not our intent to give the reader a definitive protocol or paradigm. In fact, the very expectation of "one right answer" would be misguided. Instead we offer a journey of discovery, one that is radical, relevant, and rigorous.

More information about this series at http://www.springer.com/series/8802

Christoph Meinel · Larry Leifer
Editors

Design Thinking Research

Translation, Prototyping, and Measurement

Springer

Editors
Christoph Meinel
Hasso Plattner Institute for Digital Engineering
University of Potsdam
Potsdam, Germany

Larry Leifer
Stanford University
Stanford, CA, USA

ISSN 2197-5752 ISSN 2197-5760 (electronic)
Understanding Innovation
ISBN 978-3-030-76323-7 ISBN 978-3-030-76324-4 (eBook)
https://doi.org/10.1007/978-3-030-76324-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

Thousands of individuals around the world have been formally studying design thinking—the number of individuals informally practicing elements of design thinking far exceeds that number. The Hasso Plattner Institute (HPI) in Potsdam and the d.school at Stanford University train students and support them to become designers, leaders, and entrepreneurs. In turn, HPI and Stanford University attract universities, companies and organizations from around the world that wish to understand and adopt design thinking and change processes. Individuals do indeed find inspiration for their work. What we see in the last decade is a trend toward professionalization of design thinking in light of this increased interaction and heightened awareness. D-schools are being founded around the world based on the models at HPI and at Stanford University. Researchers are networked—interaction occurs on an individual basis or in broader forums such the Hasso Plattner Design Thinking Research Program (HPDTRP) and the Global Design Thinking Alliance.

The tools that design thinkers employ find broad interest outside of academia and in the workplace. Work is fundamentally changing. Change is occurring at the managerial level but also takes hold in multidisciplinary teams. Tools and techniques for improved team interaction often incorporate design thinking tools. This has been the case for years.

Therefore, the next obvious step should be to ascertain where we are going. The future of design thinking embraces the concept of translation. Translation is a practice and an analytical category. Translation provides us with a new lens to investigate and negotiate complex shifts in design thinking. Integral to progress in the field are prototyping and measurement. This volume therefore brings these three subject areas together for a unified look at the future of design thinking.

Our researchers in the Hasso Plattner Design Thinking Research Program innovate and inspire. They identify current challenges and use established frameworks to translate theory to practice. In small teams, they create human-centered innovation by identifying needs and using cooperative information sharing. Teams develop prototypes to understand the transfer of design thinking insights to other areas and create metrics to gage the impact of design thinking. The long-term explorative projects of small researcher teams push the boundaries of current thinking and include creative and complex prototypes for modeling, while yet other teams explore the impact and needed transformation of design thinking in organizations. As concerns design thinking, universities have become the location for knowledge creation in society.

Design thinking inspires. Our students and faculty—our design thinkers—take the building blocks offered and are indefatigable. They ideate. They innovate. They build frameworks, tools, systems and methods that help users understand human–human interactions with IT augmentation. The findings are discussed in academia, but they are also intended for the public at large. Every individual who struggles to drive innovation should be attuned to the results compiled by our design teams.

The joint HPI-Stanford Design Thinking Research Program strives to improve design practice in academia, in the workplace and in local communities. Design helps traditional practice give way to new practices that value and embrace alternative points of view. I applaud the work done by our research teams. We can all learn from studying human-human interaction. Each year it is with great pride that we highlight cutting-edge research in this research series. May it contribute to academic and professional dialog and stimulate new ideas and practical design solutions.

Palo Alto, CA, USA Hasso Plattner
Winter 2020/2021

Contents

Introduction

Larry Leifer and Christoph Meinel

Abstract Many researchers would say that this edition of the Design Thinking Research (Understanding Innovation) series addresses the need to know and understand prototyping.

Many researchers would say that this edition of the Design Thinking Research (Understanding Innovation) series addresses the need to know and understand prototyping. However, if we look deeper, we can detect two other important categories: translation and measurement**.** In engineering design thinking, not only the products and services are examined under the microscope, but also the innovators themselves. How do design thinkers reach their goals? And what are they doing differently this year? Let us see why these variables work together.

Prototyping—making one's ideas real and substantial—is the most critical phase in the design paradigm. First, one must specify the user being studied in exacting detail. Then, one must define a research protocol that will examine these users and how they interact with specific prototypes.

In this edition of our research book series, we report details of our study of remote collaboration "Innovation Teams" and the neurophysiological realities of the team members. These studies also look closely at the team organization structure on a global scale.

Details include definition of the nature of these design prototypes and the energy a team should invest in creating and testing the prototype with the designated users. Prototypes may be hardware, software, and/or both. We then address the machine learning and artificial intelligence dimensions of potential prototypes in the context

L. Leifer (✉)
Stanford Center for Design Research, Stanford University, Building 560, 424 Panama Mall, Stanford, CA 94305-2232, USA
e-mail: leifer@stanford.edu

C. Meinel
Hasso Plattner Institute for Digital Engineering, Campus Griebnitzsee, Prof.-Dr.-Helmert-Str. 2-3; 14482, Potsdam, Germany
e-mail: meinel@hpi.de

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_1

of their real-world applications and their expected impact. Addressing the nature of "Reflective Design Practice," we dive deeply into the need to understand both the user and our own cognitive biases.

Dominating all of our efforts, we must seek to understand why things are happening as they are. This search for the why is closely associated with the fact that most design research must seek to discover a hypothesis. This "hypothesis discover" focus contrasts dramatically to most engineering and science research, which seeks to validate an existing hypothesis.

But how is "hypothesis discover" operationalized? We know we have to progress from concept to functional proof of concept and incorporate hardware, software, and experience. That is a tall order. We do however have improvised tools at our fingertips. It is important to keep available resources in mind when moving toward a goal. For your prototype, use objects at hand, such as Lego, craft material, or other prefabricated modeling sets. Be creative, but also methodical. Make lists. When making the passage from concept to testing, think of your team as a team of teams. Disagree freely. The more diverse the team, the higher probability of producing breakthrough innovation. And never forget the basics in engineering design teams: research, redesign, and re-innovate. Remember you as a team have the technological know-how. Within the team, your emotional interaction dynamics predict team performance. And to explain why this is the case, we ask another question: why are you challenged? Of course the answer is that life is open, complex, dynamic, and networked.

The science paradigm asks how. The design paradigm asks why.

We can correct the imbalance between science and design. When you improve the balance, you improve the impact. To improve the balance, bring the how and why together. To break it down, this means that we take existing (validated) hypotheses in science as building blocks but we do not refrain from reaching for a "hypothesis discover." Team interaction dynamics pave the way to new solutions.

It is clear that prototyping is essential to design thinking. But why is this volume concerned with translation? Over the last 12 years of the Hasso Plattner Design Thinking Research Program, we have made great strides forward. Our researchers doing the cutting edge work have come from many different multidisciplinary backgrounds. Therefore, it is fitting to detail how a social science concept—translation—finds its place here. If this were a social science book, we would then trace the translational back to the linguistic. For our purposes, the term is introduced as a vehicle to understanding transfers between scientific cultures and as an analytical category. (Bachmann-Medick, 2016) Translation provides us with a new lens through which to investigate and negotiate complex shifts.

The pandemic has necessarily changed how society works. One reality—the in-person reality—is at present no longer an option. In 2020, we translated most of our activities to the virtual realm. The *Wissenschaftsrat* (German Science Council) has reported how educational institutions were forced to quickly transition to digital teaching without adequate preparation time and with many users lacking technical skills and sufficient access to technology and without significant guidance. (Wissenschaftsrat pp. 11–12). In 2020, a profound transformation came about: in-person to digital.

Our researchers focus their lens on something such as remote design technology and try to understand, for example, how interbrain synchrony functions in a zoom world. Yet another team focuses on decoding nonverbal actions. A third team attempts to understand how design thinking was leveraged across cultures—this was only possible through translation of the term design thinking.

It is fruitful to visualize translation as a descriptor for the analysis of complex shifts taking place and which are the object of study within design thinking research. The authors approach an understanding of everyday phenomena, contemporary issues, and human-centered design through application of the design thinking toolkit at hand. The questions being asked allow design teams to find answers through the study of complex shifts. Engineering design teams examine phenomena over extended periods of time, prototyping in more than one way, ranking the prototypes, and assessing potential. Translation is a good term to describe the passage and development, process and result of design team performance.

Whereas in prototyping, researchers search for notable hypotheses, and measurement helps us quantify developments. Contributing authors assess what improvements can be implemented in areas and fields through the application of design thinking and how the effectiveness of these improvements can be quantified. Measurement allows us to draw comparisons. Standardizing and unifying measurement are essential for the professionalization of Design Thinking. Researchers strive for replicable results, especially when "hypothesis discover" is new.

The design thinking process has established phases. It is clear that innovation is no longer a by-product of some other intention. It is the ultimate goal for design thinkers. The three variables that we focus on in this volume are all integral to the design thinking process. Of course prototyping is a phase in the design thinking process, so its importance is unquestionable. Translation and measurement are tools that help us in each phase along the way to the final results of testing. With the increasing professionalization of the field, translation, prototyping, and measurement are invaluable. It is our hope that this series volume will be a starting point for dialogue in these areas.

Please join in our search for notable hypotheses and in understanding why those hypotheses are critical to the implementation of Design Thinking and Design Thinking Research. Join us.

Road Map through this Book.

Researchers from HPI and Stanford University have conducted a wide range of research projects on design thinking. This annual publication is a compilation of their findings. The shared outcomes are arranged in three parts that illustrate the comprehensive approach of the program to design thinking research, namely translation, prototyping, and measurement.

Part 1 Translation in Design Thinking.

The first part of this book concerns itself with the category of translation in several senses. Translation provides us with a new lens through which to investigate and negotiate complex shifts. In the first text (Balters/Baker/Hawthorne/Reiss), the concept is applied to understand modernization in answering the question, how is design research translated to the digital sphere? COVID-19 is given as a cause for the rise in virtual interfacing. The assessment of collaborative outcomes is central as is the nuanced discussion of interbrain synchrony in virtual versus in-person interactions. The second team of researchers (Domingo/Gutzeit/Kim/Leifer/Auernhammer) also examines societal changes as a result of COVID-19. However, in this piece, the focus is on defining the challenges in design collaboration and design work. The team offers up questions such as whether engagement in physical movement might be a prerequisite to thinking while carrying out certain activities. In the third study (Traifeh/Abou Refaei/von Thienen/ von Schmieden/Mayer/Osman/Meinel), the text builds upon the assumption that there is an absence of a direct translation for DT in the Arab world. The authors successfully stake out the development of the DT concept in five sectors of the Arab world.

The stated goal in the fourth text (Park/Whiting/Shanks) in this section is to design social platforms for user–user interactions online. To accomplish this, the authors investigate nonverbal actions and underline in their work the importance of understanding rationales for nonverbal interactive experience on online platforms. The final chapter in this section (Plank/von Thienen/Meinel) raises the questions: what is design thinking empathy? And what sub-capacities are involved when a design thinker tries to understand a user? In order to understand empathy, the authors parse the term to find biases and pitfalls in empathy as well as to delineate empathy versus compassion.

The contributions in this section all rely on translation: the translation of DT to the virtual sphere, the translation of concepts across cultures, a scientific look at interbrain synchrony, physical movement, and nonverbal actions and how these activities translate to action in two different realities in-person and virtual. And in the final text, the authors take the assumptions of neuroscience and apply them to the conceptual distinctions of empathy.

Part 2 Creation of Models for Prototyping.

In this volume's second part, HPDTRP researchers identify challenges and in many instances create models for prototyping to meet design challenges. The first chapter (Taeumel/Rein/Hirschfeld) identifies four patterns in the field of software engineering and describes how the team drafted a pattern language to enable and control exploration. A second chapter (Jane L.E./Landay) tackles the subject of iterating in photography, drawing a fruitful comparison between the parallels in DT and the photographic process. The researchers created guided photography interfaces to aid iteration. The third text (Siu/Chase/Kim/Boadi-Agyemang/Gonzalez/Follmer) addresses improved design education for impaired people. The authors created PantoGuide, a low-cost system that provides audio and haptic guidance for students who are learning remotely.

The next chapter (Miller/Gutzman/Bailenson/Mabogunje/Sonalkar) concerns itself with Interaction Dynamics Notation (IDN). The research team explores the role of IDN and traces the path of IDN development over many years. In addition to information on how it inspired other research, the chapter also offers an outlook for two directions in future work. The fifth chapter (Edelman/Santuber/Owoyele) in this section introduces PretoVids, a method for structured prototyping without engineering and code in design. The research team places value on prototyping new (low cost and agile), digital products without code or software development. In the sixth chapter, contributing authors (Mullings/Utterback/Bernstein) developed the application Drawventure to enable self-directed learning in design sketching. The paper-prototyped and user-tested application reframes sketching lessons as micro-challenges. In the final chapter (Royalty) of this section, reflective design practice is associated with increased metacognitive awareness. This can then be harnessed to optimize the student experience and awareness of growth in design classes.

The chapters in this section all help the reader to visualize models and prototypes for understanding DT in various settings. Some researchers build upon existing models and others create new prototypes. This section, in particular, offers a glimpse of the technology that is currently being developed to understand the why. The purpose of a prototype is to learn from it. Our researchers are pushing forward and constantly refining. Read the chapters and understand the process and how it is possible to move to the final product.

Part 3 Measurement in DT: How to Improve Different Areas and Fields by Applying DT.

The third part of this volume is concerned with measurement in DT. Contributing authors assess what improvements can be implemented in areas and fields through the application of DT. Researchers evaluate and leverage assets for the enhancement of DT processes. Sometimes it is a small adjustment, and other times it is a drastic modification.

Section three begins with a chapter (Mayer/Schwemmle/Nicolai/Weinberg) proposing new foundations for the impact assessment of DT in organizations. The contributing authors argue this is the only avenue for DT to move forward. The second piece in this section looks to measurement as a mindset for transformation. The author (Haskamp) answers the question of how is it possible to quantify design thinking. He concludes that a differentiated understanding of DT is needed after staking out different frameworks: DT as a set of methods, DT as a process for innovation, and DT as a mindset for driving transformation. Chapter three of this section (Sheppard/Chen/Toye/Kempf/Elfiki) describes the process of designing a survey instrument to measure the engagement of Stanford alumni. By examining the relationship between the ME310 curriculum and alumni engagement in entrepreneurship and innovation, the authors attempt to understand how course-based training in DT can translate to professional endeavors and entrepreneurial outcomes. The fourth set of researchers in this section (Dobrigkeit/Matthies/Teusner/Perscheid) apply DT to scrum. Their findings reveal that agile practitioners changed scrum to improve collaboration by incorporating teamwork, innovation, and design activities. They find that design thinking techniques provide a fruitful addition to the scrum meeting toolkit. The final contribution of this section (Auernhammer/Sonalkar/Xie/Monlux/Bruno/Saggar) and of the series volume looks to create a foundation for training the next generation of entrepreneurs and managers. The authors investigate the cognitive basis of applied creativity in business—a very little known area—by comparing the neuropsychological profiles of entrepreneurs with those of administrators and managers.

Process improvement can improve not only team productivity but also produce better calibrated metrics. A whole range of products, services, and relationships can be augmented through measurement. Obtaining a more nuanced understanding of remaining shortfalls is a targeted goal. As processes improve, the actors involved are better equipped to make smarter decisions.

Outlook.

Since 2008, researchers in the Hasso Plattner Design Thinking Research Program have arrived at valuable insights on why and how design thinking works. Our researchers share with you formats, methods, and their evaluation. Concrete examples of successes and failures are identified in this book. We encourage you, the reader, to engage with us. At www.hpi.de/dtrp, you will find the latest information on all research conducted within our HPDTRP program and learn more about our contributors.

Another resource is the Web site https://thisisdesignthinking.net, which offers a comprehensive overview of current developments in design thinking. Rich descriptions including advantages and disadvantages of particular strategies are provided. The collection of examples and interviews is appropriate for all educators and design practitioners. Here, you will also gain a better understanding of the current challenges in the field of design thinking. Experiences, stories, and inquiries can be sent to thisisdesignthinking@hpi.de.

Notable hypotheses are critical to the implementation of Design Thinking and Design Thinking Research. Enter the dialogue. Help us ask why, and you too can influence future design thinking action to form a smarter world.

Acknowledgements We thank all authors for sharing their research results in this publication. Our special thanks go to Dr. Sharon Nemeth for her constant support in reviewing the contributions and to Jill Grinager for the publication's project management.

References

Bachmann-Medick, D. (2016). *Cultural Turns: New Orientations in the Study of Culture*. De Gruyter.

Wissenschaftsrat: Arbeitsprogramm des Wissenschaftsrats Januar bis Juli 2021: https://www.wissenschaftsrat.de/download/2021/Arbeitsprogramm.pdf?__blob=publicationFile&v=11 pp. 11–12. Accessed 19 Feb 2021.

Translation in Design Thinking

Inter-Brain Synchrony and Innovation in a Zoom World Using Analog and Digital Manipulatives

Stephanie Balters, Joseph M. Baker, Grace Hawthorne, and Allan L. Reiss

Abstract The ubiquity of technology in today's world is exemplified by our ability to connect with each other instantly all around the globe. Advances in video conferencing capabilities combined with dramatic socio-dynamic shifts brought about by COVID-19 have redefined the ways in which humans interact in modern society. Human reliance on effective virtual interfacing (e.g., zoom conferencing) is evermore present in today's COVID-19 world and will undoubtedly expand in the future. This unprecedented rise in digitalization has direct implications on the output and productivity of human interactions across all design (thinking) activities and practices. Working in a virtual environment limits access to traditional design thinking tools such as (analog) "artifacts" or "manipulatives" (e.g., physical prototypes, postits, etc.). As both neuroscientists and design researchers, we are interested in elucidating the neurobiological signatures that underlie these adapted human-to-human interactions. Our overarching goal is to understand and uncover the differences in collaborative outcomes (e.g., creativity) and inter-brain synchrony in virtual versus in-person interactions using both analog and digital manipulatives. We proposed an emergent technology in brain-imaging—hyperscanning (i.e., measuring two brains simultaneously to derive measures of inter-brain synchrony) with functional near-infrared spectroscopy (fNIRS)—as an ideal brain-imaging technique to tackle this challenge. A better understanding of how the nuances of these dynamics impact inter-brain synchrony during an innovation event will provide new insights for interventions or technology that can help optimize successful interaction in both scenarios. To inform the design of future fNIRS hyperscanning studies, we review the existing

S. Balters (✉) · J. M. Baker · A. L. Reiss
Department of Psychiatry and Behavioral Science, School of Medicine, Stanford University, Stanford, CA, USA
e-mail: balters@stanford.edu

G. Hawthorne
Hasso Plattner Institute of Design, School of Engineering, Stanford University, Stanford, CA, USA

A. L. Reiss
Department of Radiology, School of Medicine, Stanford University, Stanford, CA, USA

Department of Pediatrics, School of Medicine, Stanford University, Stanford, CA, USA

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_2

fNIRS hyperscanning literature in this book chapter. On the basis of the existing literature, we highlight the current gaps in research regarding virtual interactions. We provide insight into current hurdles regarding fNIRS hyperscanning hardware and methodology and give recommendations on how to advance the field of fNIRS hyperscanning relevant to design research in the digital age.

1 Introduction

The implications of social distancing as a result of COVID-19 immediately brought in-person interaction to a halt. Almost overnight, human-to-human interaction turned digital. Virtual meetings in the corporate world, telehealth, and digital education have become necessary to regular workflow. Recent press releases from the BBC ("The Reason Zoom calls drain your Energy"; April 22, 2020), the National Geographic ("Zoom fatigue is taxing the brain. Here's why that happens"; April 24, 2020), and the New York Times ("What We Lose When We Go from the Classroom to Zoom"; May 4, 2020) are only a few examples that demonstrate the pressing public need to understand the underlying (brain) mechanisms that operate during virtual interactions. This anecdotal evidence is first supported by empirical studies that demonstrate adverse effects from overuse of virtual platforms of communication for emotional and mental health (Holmes et al., 2020; Pfefferbaum & North, 2020), educational gains (Ahmed et al., 2020; Schwartz et al., 2020), and medical care services (Hollander and Carr, 2020; Pappot et al., 2020).

In like manner, the overuse of virtual platforms from enforced digitalization has significant potential implications for design thinking practices, specifically for our ability to solve problems and innovate effectively. Our (design thinking) understanding, concepts, and methods are built on the fundamental notion that groups of individuals work collaboratively to solve complex problems that lead to innovative ideas and products (Baruah and Paulus, 2009). In light of the central role that human-to-human interaction plays in the design thinking arena, it is of crucial importance to understand how virtual and in-person interactions impact innovation events. Communicating and collaborating virtually reduces our ability to utilize and share physical tools that more quickly and easily support our communication and work. The use of physical artifacts, usually comprised of common office supplies, is core to many design thinking processes like prototyping, need finding, ideation, and synthesis processes (Brereton and McGarry, 2000; Lande and Leifer, 2009; Gibson et al., 2004). For example, while in-person ideation sessions allow for sharing both analog artifacts (i.e., tangible prototype, post-its, etc.) and digital artifacts (e.g., drawing on sketchpad, "digital" post-its, etc.), the efficiency and productivity of this work might be hindered or reduced during virtual interactions that lack physical interactions with analog artifacts (see Fig. 1). Virtual design sessions are typically limited to two-dimensional screen sharing which reduces the transfer of information and the speed of communication and exchange (Brereton & McGarry, 2000; Lande and Leifer 2009; Gibson et al., 2004). The use of artifacts (analog or digital) affects

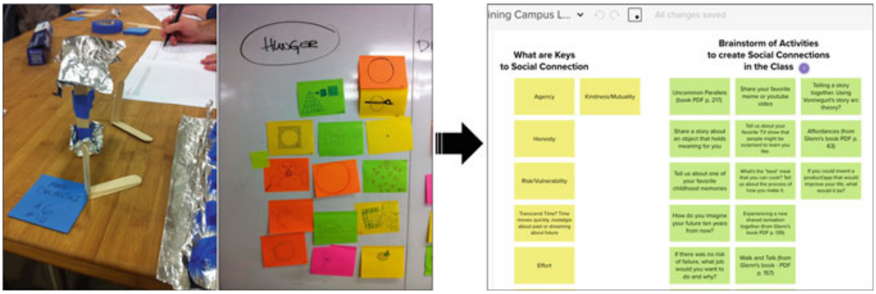


Fig. 1 Prototypes and brainstorming with foil, post-its, wood sticks, painter tape (*left*); Brainstorming with virtual post-its via Mural Online App (*right*)

practically every facet of the design thinking methodology and practice. The absence or curtailment of artifact usage and accessibility combined with the absence of face-to-face interaction together severely changes the quality of interaction and outcome between people during an innovation event (Fig. 2).

We argue that currently too little is understood about how enforced digitalization (i.e., virtual interactions with digital artifacts) impacts design thinking practices. To address this knowledge gap, we propose functional near-infrared spectroscopy (fNIRS) hyperscanning—simultaneous brain scanning of two or more people as they interact with one another—as a particularly promising approach to decode the underlying (brain) mechanisms that differ between in-person and virtual interactions. In this chapter, we review existing fNIRS hyperscanning literature with the goal of investigating how current knowledge can inform us about the differences between virtual and in-person interaction scenarios using analog and digital manipulatives. We discuss current hurdles in (fNIRS) technology and methodology and provide recommendations as to how to broadly advance this research domain.

Fig. 2 Example of an ultra-portable fNIRS imager: NIRSport by NIRX Medical Technologies, LLC

2 Background on Fnirs and Fnirs Hyperscanning

Functional near-infrared spectroscopy is a brain-imaging technique that has become increasingly popular during the last decade (Cutini and Brigadoi, 2014; Quaresima and Ferrari 2019). fNIRS devices have become small, portable, and relatively robust to motion artifacts (Baker et al., 2017). Given these attributes, brain-imaging with fNIRS allows researchers to observe neural activity in naturalistic environments, which is otherwise often not feasible with other modalities such as fMRI or EEG (Scholkmann et al., 2013; Gvirts and Perlmutter 2020; Quaresima and Ferrari 2019; Baker et al., 2017). In fact, fNIRS is currently the portable brain-imaging technique with the highest spatial resolution (Scholkmann et al., 2014). This characteristic allows scientists to make inferences about specific neurocognitive processes tied to particular cortical regions and networks. In other words, fNIRS provides key information about which cortical brain regions or cortical brain networks are active during a specific task.

Given its methodological qualities, fNIRS has recently allowed Design Science researchers to study brain function during design activities (Gero & Milovanovic, 2020). For example, Kato et al. used fNIRS to test for differences in prefrontal cortex activation while "pretending to sketch" versus "actually sketching" (Kato et al., 2017, 2018b), as well as when sketching by hand versus using computer operation (Kato, Okada, et al., 2018). Shealy et al. (2017) applied fNIRS to assess whether there are differences in prefrontal cortex activation between two groups of design students (college freshmen and seniors) who were engaged in an ideation session. In a follow-up study, the same authors tested for changes in cortical activation in and connectivity patterns between areas of the prefrontal and parietal cortex when performing concept mapping (i.e., labeling concepts and drawing directional connections between them) compared to concept listing (i.e., generating a list of concepts without indicating relationships) (Hu et al., 2019). Another research team focused on testing for differences in brain states during solution generation when using three different concept generation techniques, that is, an unstructured (brainstorming), partially structured (morphological analysis), and structured technique (TRIZ). They measured changes in blood oxygenation in the prefrontal cortex via fNIRS and executed activation as well as connectivity analysis (Hu et al., 2018; Shealy et al., 2018) and tested further for differences in cortical activation between the two hemispheres (Shealy and Gero, 2019).

About a decade ago, neuroscientists began to extend neuroimaging with fNIRS from scanning single brains to scanning multiple brain simultaneously ("i.e., hyperscanning") (Cui et al., 2012; Funane et al., 2011). Since then, the field of fNIRS hyperscanning has revolutionized our ability to observe neurobiological signatures of human-to-human interactions that are invisible to the naked eye. As a growing body of evidence indicates, interpersonal communication is frequently accompanied by synchronization of cortical activity (Cui et al., 2010), and such inter-brain synchrony (IBS) is often associated with enhanced behavioral metrics of interaction (Baker et al., 2016). Notably, some of the first fNIRS hyperscanning studies focused on

Fig. 3 Example of an fNIRS hyperscanning study within the field of NeuroDesign/ Design Neurocognition (Mayseless et al. 2019). Partners of a dyad are engaging in an ideation task (left), and in a problem-solving task (i.e., 3D model building) (right)

the field of "NeuroDesign" or "Design Neurocognition" (Mayseless et al., 2019; Lu et al., 2019, 2020). Mayseless et al. (2019) applied fNIRS hyperscanning to observe differences in IBS between partners engaging in ideation versus problem-solving activities (see Fig. 3). Lu et al. (2019) applied fNIRS hyperscanning to investigate how cooperative and competitive interaction modes affect the group creative performance; and whether the gender composition of the dyad impact these interactions (and related performance) (Lu et al., 2020).

The study examples presented above highlight how fNIRS has been successfully applied within naturalistic environments, including design thinking scenarios. We argue that fNIRS is currently the optimal tool to be used for understanding the underlying (brain) mechanisms that differ between in-person and virtual interactions using both analog and digital manipulatives. In order to inform the design of a study that fills this research gap, we review the existing fNIRS hyperscanning literature in the following section. The reader of this chapter should note that the review and analysis presented below extends our recently published perspective on the future of fNIRS hyperscanning (Balters et al., 2020).

3 Method: A Literature Review on Fnirs Hyperscanning

We executed a Google Scholar search and considered all peer-reviewed manuscripts that were published through May 15, 2020. Our search strategy included the following keywords: "fNIRS hyperscanning" and "NIRS hyperscanning." For each keyword, we inspected the first 250 entries and included all journal or conference articles in the English language. Additionally, we checked the reference lists of the included articles for any additional relevant articles. From the initial 69 fNIRS hyperscanning studies, we excluded nine infant-parent fNIRS hyperscanning studies (Leong et al., 2017; Reindl et al., 2018; Miller et al., 2019; Azhari et al. 2019, 2020; Piazza et al., 2020; Behrendt et al., 2020; Quinones-Camacho et al. 2019; Nguyen et al., 2020) and two more studies that included comparisons of temporally non-congruent fNIRS scans (Hou et al., 2020; Liu et al. 2017b). All resulting 58 studies are included in Table 1.

Table 1 List of 58 fNIRS hyperscanning studies—updated April 15 2020

Authors and year	Setup	Interaction scenarios	Interaction manipulatives	Cognitive function	Region
Duan et al. 2015	1 nonad, 1 scan cond	FTF	Tangible interface IM w/out verbal IM	Attention* Executive function* Motor	mPFC lPC(TPJ)
Jiang et al. 2015	12 triads (6fff, 6mmm), 1 scan cond	FTF	VeiballM	Attention* Executive function* Language* Social cognition*	IPFC lPC(TPJ)
Liu et al. 2015	10 dyads, 4 scan cond	SBS	Shared digital IM w/out verbal IM	Attention* Executive function* Motor Visuospatial function	IPFC lPC(TPJ)
Osaka et al. 2015	15 dyads (7ff, 8 mm), 4 scan cond	FTF + FTFv-b	Verbal IM	Attention* Executive function* Language*	Whole head
Baker et al. 2016	Ill dyads (38ff,34 fm,39mmf), 1 scan cond	FTF	Separate digital IM w/out verbal IM	Attention* Executive function* Motor	IPFC rPC(TPJ)
Liu N. et al. Liu, Mok, et al., 2016	9 dyads (2ff, 5 fm, 2 mm), 4 scan cond	FTF	Shared physical IM and verbal IM + VeiballM	Attention* Executive function* Motor Language Visuospatial function	rPFC rTC
Liu et al. 2016	10 dyads (2ff, 8 mm), 4 scan cond	SBS	Shared digital IM w/out verbal IM	Attention* Executive function* Motor Visuospatial function	PFC
Nozawa et al. 2016	12 quartets, 2 scan cond	FTF	Verbal IM	Attention* Executive function* Language	mPFC
Tang et al. 2016	101 dyads, 2 scan cond	FTF	Separate digital IM w/out verbal IM	Attention* Executive function* Social cognition*	mPFC rPC(TPJ)
Balconi & Vanutelli, 2017a	16 dyads, 2 scan cond	SBS v-b	Separate digital IM w/out verbal IM	Attention* Executive function* Social cognition	PFC

(continued)

Table 1 (continued)

Authors and year	Setup	Interaction scenarios	Interaction manipulatives	Cognitive function	Region
Balconi & Vanutelli, 2017b	14 dyads, 2 scan cond	SBS v-b	Separate digital IM w/out verbal IM	Attention* Executive function* Motor Social congition	PFC
Hirsch et al. 2017	19 dyads (6ff, lOfin, 3 mm), 4 scan cond	FTF	Non-verbal IM	Attention* Executive function*	PFC PC TC
Hu et al. 2017	35 dyads (all ff) 2 scan cond	FTF v-b	Separate digital IM w/out verbal IM	Attention* Executive function* Motor	PFC
Ikeda et al. 2017	4 groups of 24 or 25, 4 scan cond	FTB + BTB	Non-verbal IM	Attention* Executive function* Motor	mPFC
Liu, Piazza, et al., 2017	22 dyads (all mm) 4 scan cond	SBS	Shared digital IM w/out verbal IM	Attention* Executive function* Motor Visuospatial function	PC
Pan et al. 2017	49 dyads (all fin) 1 scan cond	FTF v-b	Separate digital IM w/out verbal IM	Attention* Executive function* Motor	rPFC rPC
Piva et al. 2017	20 dyads (4ff, 16 fm, 2 mm), 4 scan cond	FTF	Shared digital IM and verbal IM	Attention* Executive function* Motor Language* Social cognition*	PFC PC
Takeuchi et al. 2017	15 dyads (4ff, 3 fm,8 mm), 1 scan cond	SBS	Shared digital IM w/out verbal IM	Attention* Executive function* Motor Visuospatial cognition Social cognition	PFC

(continued)

Table 1 (continued)

Authors and year	Setup	Interaction scenarios	Interaction manipulatives	Cognitive function	Region
Zhang et al. 2017a	30 dyads, 2 scan cond	FTF	Shared physical IM and verbal IM	Attention* Executive function* Motor Language* Visuaspatial cognition Social cognition*	mFFC IPFC lPC(TPJ)
Zhang et al. 2017b	33 dyads, 2 scan cond	FTF	Shared digital IM w/out verbal IM	Attention* Executive function* Motor Language* Visuaspatial cognition Social cognition*	mPFC IPFC lPC(TPJ)
Zhao et al. 2017	48 dyads (24ff, 24 mm), 3 scan cond	BTB	Tangible interface IM w/out verbal IM	Attention* Executive function* Motor	mFFC rPFC
Dai et al. Zhang, Ding, et al., 2018	48 dyads (24ff, 24 mm), 3 scan cond	BTB	Tangible interface IM w/out verbal IM	Attention* Executive function* Motor	IPFC IPC 1TC
Dai et al. Dai, Liu, et al., 2018	22 triads (all same sex), 4 scan cond	FTF + BtB	Verbal IM	Attention* Executive function* Language*	IPFC IPC 1TC
Fishburn et al. 2018	20 triads, 5 scan cond	FTF	Shared physical IM and verbal IM + Shared digital IM w/out verbal IM	Attention* Executive function* Motor Language Visuospatial function	rPFC IPFC
Hirsch et al. 2018	27 dyads (10ff,12 fm,5 mm), 4 scan cond	FTFv-b	Verbal IM	Attention* Executive function* Language*	PFC PC
Pan et al. 2018	12 dyads, 2 scan cond	FTF	Verbal IM	Attention* Executive function* Memory Language*	IPFC IPC 1TC

(continued)

Table 1 (continued)

Authors and year	Setup	Interaction scenarios	Interaction manipulatives	Cognitive function	Region
Xue et al. 2018	45 dyads, 1 scan cand	FTF	Verbal TM	Attention* Executive function* Social cognition*	PFC rPC(TPJ)
Zhang, Y. et al. Zhang, Ding, et al., 2018b	17 dyads, 2 scan cand	FTF	Verbal IM	Attention* Executive function* Language Memory Social cognition*	PFC rPC(TPJ)
Zhang et al. 2018a	31 dyads, 1 scan cond	SBS	Shared digital IM w/out verbal IM	Attention* Executive function* Social cognition	PFC
Zheng et al. 2018	32 dyads, 2 scan cond	SBS	Shared digital IM and verbal IM	Attention* Executive function* Language* Memory Social cognition*	PFC PC TC
Balconi et al. 2019	31 dyads 16 dyads (all fl), 2 scan cond	SBS v-b	Separate digital IM w/out verbal IM	Attention* Executive function* Memory Social cognition	PFC pMC
Cheng et al. 2019	31 dyads (16ff, 15 fm), 2 scan cond	FTFv-b	Separate digital IM w/out verbal IM	Attention* Executive function* Motor	PFC
Liu J. et al. 2019	21 dyads, 4 scan cond	FTF + BTB	Shared digital IM and verbal IM	Attention* Executive function* Language Memory Social cognition*	PFC rPC(TPJ)
Lu et al. 2019	52 dyads, 4 scan cond	FTF	Verbal IM	Attention* Executive function* Language	PFC rPC(TPJ)

(continued)

Table 1 (continued)

Authors and year	Setup	Interaction scenarios	Interaction manipulatives	Cognitive function	Region
Mayseless et al. 2019	25 dyads (8ff, 8 fm,9 mm), 2 scan cond	FTF	Shared physical IM and verbal IM	Attention* Executive function* Language* Memory Motor Social cognition*	IPFC lPC(TPJ) ITC
Niu et al. 2019	20 dyads (Iff, 9 mm), 4 scan cond	SBS	Non-verbal IM	Attention* Executive function* Motor	rPFC rPC
Nozawa et al. 2019	32 dyads (9ff, 23 mm), 4 scan cond	FTF	Verbal IM + Non-verbal IM	Attention* Executive function* Language Memory Social cognition	PFC
Sarinasadat, Hattori, et al., 2019	IS dyads, 2 scan cond	FTF	Shared digital IM and verbal IM	Attention* Executive function* Language Memory Social cognition	PFC
Sarinasadat, Miyake, et al., 2019	15 dyads, 2 scan cond	FTF	Shared digital IM and verbal IM	Attention* Executive function* Language Memory Social cognition	PFC
Vanzella et al. 2019	5 dyads, 4 scan cond	SBS	Tangible interface IM w/out verbal IM	Attention* Executive function* Memory Motor*	dPFC MC TC
Balconi et al. 2020	15 dyads (all fl), 2 scan cond.)	SBS v-b	Separate digital IM w/out verbal IM	Attention* Executive function* Memory Social cognition	PFC pMC

(continued)

Table 1 (continued)

Authors and year	Setup	Interaction scenarios	Interaction manipulatives	Cognitive function	Region
Feng et al. 2020	120 dyads (60S, 60 mm), 2 scan cond	FTFv-b	Separate digital IM w/out verbal IM	Attention* Executive function* Memory Language Social cognition	PFC
Lu et al. 2020	66 dyads (26ff,22 fm,18 mm), 2 scan cond	FTF	Verbal IM	Attention* Executive function* Memory Language	PFC rPC(TPJ)
Noah etal. 2020	15 dyads, 2 scan cond	FTF	Non-verbal IM	Attention* Executive function* Social cognition*	PFC PC
Pan et al. 2020a	24 dyads (all fl), 4 scan cond.)	FTF	VeiballM	Attention* Executive function* Memory Language* Social cognition*	PFC IPC ITC
Pan, Guyon, et al., 2020	16 dyads (all fl), 1 scan cond.)	SBS	Shared digital IM and verbal IM	Attention* Executive function* Memory Language* Social cognition*	PFC PC TC
Sun et al. 2020	34 dyads (27ff,7 mm), 2 scan cond	FTFv-b	Separate digital IM w/out verbal IM	Attention* Executive function* Motor	PFC
Yang et al. 2020	93 sextets (5 l£6EEEf,42mmmmmm), 3 scan cond	FTF	Verbal IMH-Shared digital IM and verbal IM	Attention* Executive function* Memory Motor Language Social cognition*	PFC rPC(TPJ)

(continued)

Table 1 (continued)

Authors and year	Setup	Interaction scenarios	Interaction manipulatives	Cognitive function	Region
Zheng et al. 2020	32 dyads, 2 scan cond	SBS	Shared digital IM and verbal IM	Attention* Executive function* Memory Language* Social cognition*	PFC PC TC

Abbreviations comprise ff (female-female), fm (female-male), mm (male-male), FTF (face-to-face), SBS (side-by-side), BTB (back-to-back), PFC (prefrontal cortex), PC (parietal cortex), TC (temporal cortex), l (left), r (right), m (medial), TPJ (temporoparietal junction). Shielded refers to a setup in which participants interaction is shielded by a physical divider, and cond. is the abbreviation for condition(s). We marked those studies that included wavelet coherence analysis "WTC." We further included cognitive functions that were required to execute the experimental task and highlighted those cognitive functions that were investigated with an "*".

4 Analysis

To investigate to what extent the differences between virtual and in-person interaction scenarios using analog and digital manipulatives have been studied, we extracted all experimental conditions (i.e., "hyperscans") that were executed across all 58 fNIRS hyperscanning studies. This resulted in a total of 151 hyperscans. We then executed a thematic analysis to cluster all hyperscans across a consistent methodological structure. As shown in Fig. 4, a 7 x 8 matrix naturally occurred across two dimensions. First, for *interaction scenario,* the fNIRS hyperscanning studies could be clustered into seven different categories. These categories included settings in which the participants interacted:

1. face-to-face *(FTF)*
2. side-to-side *(SBS)*
3. face-to-back *(FTB)*
4. back-to-back *(BTB)*
5. face-to-face with visual barriers between subjects *(FTF v-b)*
6. side-by-side with visual barriers between subjects *(SBS v-b)*
7. via a virtual divide *(Virtual).*

Second, for *interaction manipulative (IM),* we clustered all experimental tasks into categories that included

1. participants sharing physical objects such as 3D puzzles or post-it notes and verbally communicating while interacting with the objects (i.e., *shared physical IM and verbal IM*)
2. participants interacting with (non-shared) tangible interfaces such as buttons, but also with music instruments such as a drum kit, without verbal communications (i.e., s*physical IM w/out verbal IM*)
3. participants interacting solely verbally (i.e., *verbal IM*)

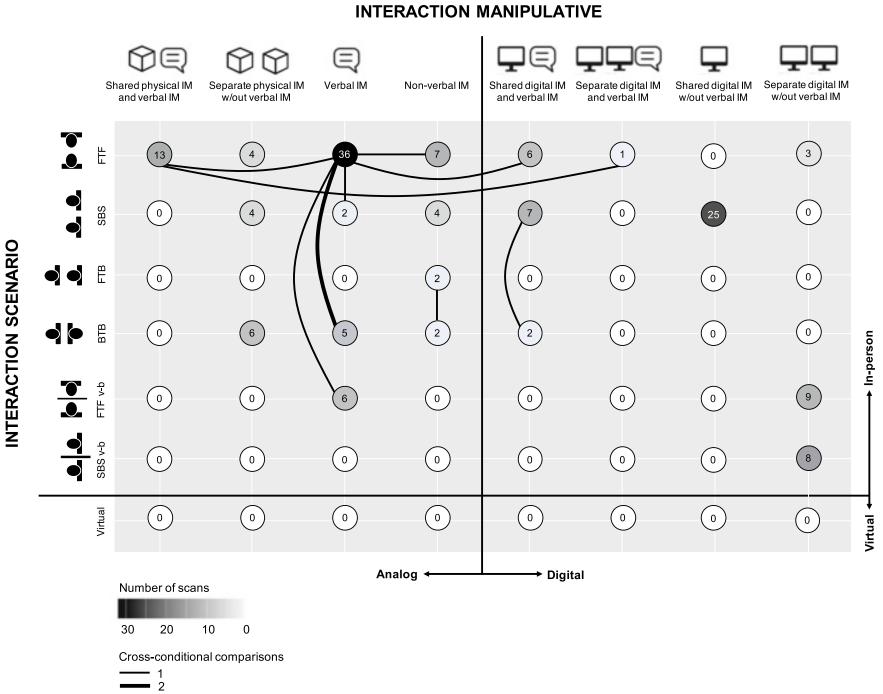


Fig. 4 Cross-sectional distribution of all 151 conducted hyperscan conditions across the *interaction scenario* and *interaction manipulative* axes. The shade of each circle provides the number of scans that belong to each cross-sectional condition. Light colors indicate fewer scans and darker colors indicate more scans. The width of each line provides an illustration of the number of scans conducted within each cross-condition comparison

4. participants solely interacting non-verbally, such as looking at one another or synchronizing limb movements while observing one another (i.e., *non-verbal IM*)
5. participants interacting together on one shared computer screen while also engaging in verbal interaction (i.e., *shared digital IM and verbal IM*)
6. participants interacting together on separate computer screens while also engaging in verbal interaction (i.e., *separate digital IM and verbal IM*)
7. participants interacting together on one shared computer screen without verbally communicating (i.e., *shared digital IM w/out verbal IM*
8. participants interacting on separate digital task manipulatives without interacting verbally (i.e., *separate digital IM w/out verbal IM*).

As shown in Fig. 4, the distribution of hyperscan conditions across the 7 x 8 matrix is highly unequal and sparse. Almost half of all hyperscans (46.1%) were conducted in *FTF* interaction (Baker et al., 2016; Duan et al., 2015; Fishburn et al., 2018; Funane et al., 2011; Hirsch et al., 2017; Holper et al., 2012, 2013; Jiang et al. 2012, 2015; Liu et al., 2016a; Lu et al., 2019, 2020; Mayseless et al., 2019; Noah et al., 2020; Nozawa et al., 2016, 2019; Osaka et al., 2014, 2015; Pan et al., 2018, 2020a; Sarinasadat, Miyake, et al., 2019; Tang et al., 2016; Xue et al., 2018; Zhang et al. 2017a, 2017b, 2018b; Dai et al., 2018a; Yang et al., 2020), while another 27.6% of hyperscans were conducted in *SBS* interaction (Cheng et al., 2015; Cui et al., 2012; Dommer et al., 2012; Duan et al., 2013; Liu et al. 2015, 2019; 2016b, 2017a; Niu et al., 2019; Osaka et al., 2014; Pan, Guyon, et al., 2020; Takeuchi et al., 2017; Vanzella et al., 2019; Zhang, Ding, et al., 2018; Zheng et al., 2018, 2020). During the remaining scans, participants interacted *FTB* ($N = 2$) (Ikeda et al., 2017), *BTB* ($N = 15$) (Dai et al., 2018b; Ikeda et al., 2017; Jiang et al., 2012; Liu et al., 2017b; Zhao et al., 2017; Dai et al., 2018a), *FTF v-b* ($N = 15$) (Cheng et al., 2019; Feng et al., 2020; Hirsch et al., 2018; Hu et al., 2017; Osaka et al., 2015; Pan et al., 2017; Sun et al., 2020), and *SBS v-b* ($N = 8$) (Balconi and Vanutelli, 2017a,b; Balconi et al., 2019; Balconi and Fronda, 2020). Notably, no fNIRS hyperscanning study included a virtual interaction scenario.

A more equal empirical focus currently exists for *interaction manipulatives*. With an exception of *verbal IM* that reached a total of 49 scans (32.4%), the remaining IM categories shared more equal distribution ranging from 8.6% for *shared physical IM and verbal IM* (Fishburn et al., 2018; Liu et al., 2016a; Mayseless et al., 2019; Zhang et al. 2017a, 2017b), 9.3% for *separate physical IM w/out verbal IM* (Dai et al., 2018b; Duan et al., 2015; Funane et al., 2011; Holper et al., 2012; Vanzella et al., 2019; Zhao et al., 2017), 9.9% for non-verbal IM (Hirsch et al., 2017; Ikeda et al., 2017; Niu et al., 2019; Noah et al., 2020; Nozawa et al., 2019), 9.3% for *shared digital IM and verbal IM* (Liu et al., 2019; Pan, Guyon, et al., 2020; Sarinasadat, Miyake, et al., 2019; Zheng et al., 2018, 2020; Piva et al., 2017; Yang et al., 2020), 17.2% for *shared digital IM w/out verbal IM* (Cheng et al., 2015; Cui et al., 2012; Dommer et al., 2012; Duan et al., 2013; Fishburn et al., 2018; Liu et al., 2015, 2016, 2017a; Takeuchi et al., 2017; Zhang, Ding, et al., 2018), and 13.1% for *separate digital IM w/out verbal IM* (Baker et al., 2016; Balconi and Vanutelli, 2017a, 2017b; Balconi et al., 2019; Balconi and Fronda, 2020; Cheng et al., 2019; Feng et al., 2020; Hu et al., 2017; Pan et al., 2017; Sun et al., 2020; Tang et al., 2016). Only one study included *separate digital IM and verbal IM* (Yang et al., 2020).

The lines in Fig. 4 represent the frequency of cross-condition comparison reported. A cross-condition comparison occurred when the *interaction scenario* or *interaction manipulative* during a hyperscan differed between experimental tasks. A total of ten papers included in our analysis included one cross-condition comparison (Fishburn et al., 2018; Ikeda et al., 2017; Jiang et al., 2012; Liu et al. 2019, 2016a; Nozawa et al., 2019; Osaka et al. 2014, 2015; Dai et al., 2018a; Yang et al., 2020). Seven studies included the condition *FTF/verbal IM* in their comparison. Specifically, they compared *FTF/verbal IM* with *FTF/shared physical IM and verbal IM* (Liu et al., 2016a), with *FTF v-b/ verbal IM* (Osaka et al., 2015), with *BTB/verbal IM* (Dai et al., 2018a; Jiang et al., 2012), with *SBS/ verbal IM* (Osaka et al., 2014), with *FTF/ non-verbal IM* (Nozawa et al., 2019), and with *separate digital IM and verbal IM* (Yang et al., 2020). The remaining three comparisons have been conducted between *FTB/ non-verbal IM* and *BTB/ non-verbal IM* (Ikeda et al., 2017), between *SBS/ shared digital IM and verbal IM* and *BTB/ shared digital IM and verbal IM* (Liu et al., 2019), and between *FTF/ shared physical IM and verbal IM* and *FTF/ shared digital IM w/out verbal IM* (Fishburn et al., 2018).

5 Discussion and Future Direction

This review highlights areas of fNIRS hyperscanning that have received the most attention to date as well as topics that have received little to no attention. For example, no fNIRS hyperscanning study has, to date, accessed inter-brain synchrony within a textitVirtual Interaction Scenario, nor has any study assessed the difference between a textitVirtual Interaction Scenario and an *in-person interaction scenario*. With respect to *interaction manipulative*, the distribution between *analog interaction manipulative* ($N = 91$) and *digital interaction manipulative* ($N = 61$) is more equal. Though two studies exist that cross the "analog/digital barrier" and include tasks from both domains (Fishburn et al., 2018; Yang et al., 2020) (see Fig. 4), none of the studies has yet compared an "analog" manipulative with its "digital" representation (e.g., physical puzzle versus digital puzzle). Instead, Fishburn et al. (2018) included experimental tasks that compared physical tangram puzzling with watching other individuals "puzzling"; and Yang et al. (2020) included a computer task within their priming activity during the experiment. The current gaps in fNIRS hyperscanning research within and across all four quadrants (see Fig. 5 left) present potential challenges for designing new studies (see Fig. 5 right). These are discussed below.

fNIRS hyperscanning systems for virtual interaction scenarios. Conducting an fNIRS hyperscan when the participants are completely separated is by no means trivial. Indeed, the majority of studies reviewed above ($N = 45$) employed a traditional hyperscanning approach in which the optodes from a single NIRS device were split, such that half of the optodes were placed on one participant and the remaining half on the other. This simple approach is effective and reduces technical issues related to time synchronizing and trigger placement within the time series. However, this approach also limits the possible distance between hyperscanning participants

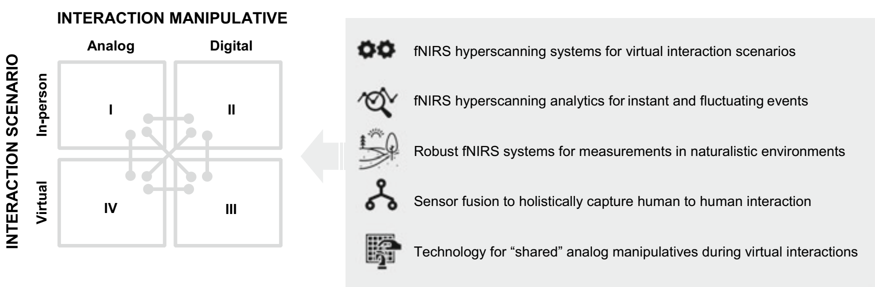


Fig. 5 Potential future of fNIRS hyperscanning as focused on increasing research within and across all four quadrants (*left*). List of specific recommendations that could support future research efforts (*right*)

since the optodes are physically tethered to the same fNIRS device. In order to overcome the physical limitations of using a single fNIRS device, researchers can employ multiple devices that are not physically connected ($N = 13$ within this review). As mentioned above, this strategy poses many technical complications including but not limited to accurate synchronizing of multiple time series as well as accurate placement of triggers within each time series. Thus, researchers interested in conducting hyperscanning across a virtual divide must consider solutions to these issues. Unfortunately, current solutions are often derived "in-house" and are not widely shared among the fNIRS hyperscanning community. We argue that future efforts be made to develop and disseminate these tools.

fNIRS hyperscanning analytics to capture instant and fluctuating collaborative interactions. Another potential hurdle is related to limitations in current analysis techniques. Collaborative interactions are instant and fluctuating. fNIRS hyperscanning analytics are, to date, however rather static. For example, researchers have applied more traditional statistical approaches such as block averaging (e.g., Holper et al. 2013), analysis of co-variance (e.g., Funane et al. 2011), and correlation analysis (e.g., Duan et al. 2013) to quantify and analyze inter-brain coherence (i.e., "synchrony between the brains"). Another technique, which is highly popular within the fNIRS hypersanning research community (i.e., roughly 70% of all studies included this analysis) is the "wavelet transform analysis" or "WTC" (Cui et al., 2012). WTC allows the analysis of both coherence and phase lag in two time series across time and frequency domains. However, this method was originally developed for block-design studies, and researchers have not yet developed extensions to study fluctuating events within time series. Lastly, the Granger causality method has been applied to fNIRS hyperscanning data sets to derive the directionality of synchrony between two time series (Holper et al., 2012). However, like WTC, Granger causality does not take fluctuating events into account. For future research, it will thus be of crucial importance to advance fNIRS hyperscanning analytics to analyze more real-time (i.e., instant and fluctuating) human-to-human interaction.

Robust fNIRS systems for reliable measurements in naturalistic environments. Though portable fNIRS devices are commercially available, and fNIRS (hyperscanning) research within naturalistic environments has accelerated, it is important to note that these developments are relatively recent. Indeed, there remains much potential to increase fNIRS's robustness for naturalistic experimentation. For example, developments in hardware could increase the tool's robustness for movements and ambient light. A standardization to integrate short-channels that measure extra-cerebellar blood flow (i.e., noise) could allow researchers to reduce the signal to noise ratio (Brigadoi & Cooper, 2015). None of the papers included in this review include short-channels in their experiment. Further, researchers to date have uniformly had to compromise and limit their regions of interest under observation due to limitations of their equipment; the number of measurement channels ranged between 1 and 54 ($M = 18.3$, $SD = 12.6$). Thus, an increase in the number of measurement channels will permit a more holistic understanding of brain networks of interest. Finally, fNIRS configuration, pre-processing, and analysis includes many steps with, to date, exhibit a considerable degree of variation. This ultimately hampers the comparison of results across studies. (It is important to note that many fNIRS researchers have called for a standardization of methodology and procedures during fNIRS research to allow for meta-analyses and confirmatory science (Cutini and Brigadoi, 2014; Herold et al., 2017; Balters et al., 2020)). Such standardizations could include methods for the placement of optodes (Herold et al., 2017), data processing steps (Brigadoi et al., 2014; Di Lorenzo et al., 2019), and choice of fNIRS signals (i.e., oxy- vs. deoxygentated hemoglobin) (Tachtsidis and Scholkmann, 2016), to name a few. These standardizations could be further integrated into the existing open-source fNIRS-specific data analysis packages (e.g., HOMER2, NIRS SPM, nirsLAB, open-potato, etc.).

Sensor fusion to holistically capture human-to-human interaction. Given the complex nature of human-to-human interaction, we argue that multi-modal approaches will be required to determine which parameters (e.g., behavioral, environment, and/or technological) are most explanatory with respect to potential differences in neurocognitive signatures between virtual and in-person interactions using analog and digital manipulatives. For example, to accomplish higher temporal resolution, electroencephalography (EEG) could be added to fNIRS hyperscanning scenarios. Congruent fNIRS-EEG measurement has been successfully applied for single brain assessments (Li et al., 2017). Physiological metrics (e.g., heart rate, heart rate variability, breathing rate, galvanic skin response, etc.) as well as behavior measures (e.g., eye tracking, body motion tracking, analysis of voice, emotional face tracking, etc.) will provide important information about psych-physiological states of the interacting partners. To account for potential external biases, it will be important to capture environmental information (e.g., ambient noise, reflecting light on reading glasses, etc.) and technological parameters (e.g., computer frame-rate, computer audio, internet speed, etc.).

Technology for "shared" analog manipulatives during virtual interactions. Another hurdle will be to create (robotic) technology which allows for the assessment of brain activity and inter-brain coherence during a *virtual interaction scenario* using a *shared analog interaction manipulative* (e.g., two individuals playing chess remotely with the physical chess figures moving on both chess boards). Such technology could be an important advance for the design thinking arena and allow for resumption of some aspects of "tangible" interactions across the virtual divide such as prototyping. Design thinking researchers could investigate to what extent such tangible interaction facilitates the communication of ideas and to what extent actual human physical presence is needed. Scientists and engineers have begun to tackle this technological challenge (e.g., miniature swarm robots—http://shape.stanford.edu/research/swarm/).

6 Conclusion

In this book chapter, we discussed the need to better understand the impact and consequences that "COVID19 enforced digitalization" (e.g., virtual meetings with digital design artifacts) has on innovative design outcomes. This chapter has attempted to emphasize the potential value in conducting hyperscanning experiments (i.e., measuring two or more brains simultaneously) with functional near-infrared spectroscopy (fNIRS) to derive the underlying (brain) mechanisms that differ between in-person and virtual interactions using both analog and digital manipulatives. At present, no study has yet directly addressed this issue. It is our hope that advances in technology (i.e., hyperscanning devices and manipulatives) will allow easier access for scientists. To move forward, we presented five different domains that could support future research. Overall, we argue that the proposed fNIRS hyperscanning research has a holistic and eminently important relevance to any current or future scenario in which two (or more) humans are interacting via a virtual divide. As discussed above, human reliance on digital interfacing is evermore present in today's COVID-19 world and will undoubtedly persist or even expand in the future. We argue that it is imperative that we understand the influence of digital interfacing on inter-brain coherence and related human-to-human behavior in the context of design thinking activities (e.g., joint task performance, innovation, empathy, etc.). By assessing the impact of environmental conditions (i.e., virtual/ in-person) and necessary task-specific artifacts/tools (e.g., digital/analog manipulatives) on the brain and behavioral bases of collaboration during an innovation event, we will be able to expand our knowledge of the role of IBS in teams as well as suggest ways to improve innovation practices—in both virtual and in-person team interactions. Outside of addressing, these issues for identifying successful interventions (e.g., coherence-increasing neurofeedback), efforts in this domain will provide useful information that will assist in the development of future technologies to improve virtual experience. Improvements in effectiveness (e.g., coherence-increasing) may have great relevance to any current or future scenario in which two (or more) people are interacting via a virtual divide. We contend that the generated findings will be translatable into a vast amount of complementary contexts, ranging from any kind of virtual meeting in a professional setting, to telehealth and mission critical scenarios (e.g., ISS to ground station), to digital education and personal well-being activities (e.g., classroom yoga, catching up with friends, etc.). Furthermore, this investigation is further critical in gaining a better understanding of which tools and conditions can optimize the outcome of an innovation event in a (post-) pandemic world. Interfacing with artifacts is essential to the design thinking process as it is the cornerstone to the prototyping, need finding and synthesis processes. Uncovering the advantages of various scenario combinations (e.g., virtual vs. in-person; digital vs. analog) can help us identify user needs for such scenarios. Imagine technology that allows two virtually connected humans to interact with the same physical object or educators who are able to physically interact with their students' tangible prototype—all remotely. The development of such tools will facilitate these and other cross-conditional comparisons while providing novel ways to enhance and optimize the way humans interact with virtual constraints.

References

Ahmed, H., Allaf, M., and Elghazaly, H. (2020). Covid-19 and medical education. *The Lancet Infectious Diseases.*

Azhari, A., Leck, W., Gabrieli, G., Bizzego, A., Rigo, P., Setoh, P., et al. (2019). Parenting stress undermines mother-child brain-to-brain synchrony: A hyperscanning study. *Scientific Reports, 9*, 1–9.

Azhari, A., Gabrieli, G., Bizzego, A., Bornstein, M. H., and Esposito, G. (2020). Maternal anxious attachment style is associated with reduced mother-child brain-to-brain synchrony. bioRxiv

Baker, J. M., Liu, N., Cui, X., Vrticka, P., Saggar, M., Hosseini, S. H., et al. (2016). Sex differences in neural and behavioral signatures of cooperation revealed by fnirs hyperscanning. *Scientific Reports, 6*, 26492.

Baker, J. M., Rojas-Valverde, D., Gutierrez, R., Winkler, M., Fuhrimann, S., Eskenazi, B., et al. (2017). Portable functional neuroimaging as an environmental epidemiology tool: A how-to guide for the use of fnirs in field studies. *Environmental Health Perspectives, 125*, 094502.

Balconi, M., & Fronda, G. (2020). The "gift effect" on functional brain connectivity. inter-brain synchronization when prosocial behavior is in action. *Scientific Reports, 10*, 1–10.

Balconi, M., & Vanutelli, M. E. (2017a). Brains in competition: Improved cognitive performance and inter-brain coupling by hyperscanning paradigm with functional near-infrared spectroscopy. *Frontiers in Behavioral Neuroscience, 11*, 163.

Balconi, M., & Vanutelli, M. E. (2017b). Interbrains cooperation: Hyperscanning and self-perception in joint actions. *Journal of Clinical and Experimental Neuropsychology, 39*, 607–620.

Balconi, M., Fronda, G., & Vanutelli, M. E. (2019). Donate or receive? Social hyperscanning application with fnirs. *Current Psychology, 38*, 991–1002.

Balters, S., Baker, J. M., Hawthorne, G., & Reiss, A. L. (2020). Capturing human interaction in the virtual age: A perspective on the future of fnirs hyperscanning. *Frontiers in Human Neuroscience, 14*, 458.

Baruah, J., & Paulus, P. B. (2009). Enhancing group creativity: The search for synergy. *Research on Managing Groups and Teams, 12*, 29–56.

Behrendt, H. F., Konrad, K., Perdue, K. L., & Firk, C. (2020). Infant brain responses to live face-to-face interaction with their mothers: Combining functional near-infrared spectroscopy (fnirs) with a modified still-face paradigm. *Infant Behavior and Development, 58*, 101410.

Brereton, M., McGarry, B. (2000). An observational study of how objects support engineering design thinking and communication: implications for the design of tangible media. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems. pp. 217–224.

Brigadoi, S., & Cooper, R. J. (2015). How short is short? optimum source–detector distance for short-separation channels in functional near-infrared spectroscopy. *Neurophotonics, 2*, 025005.

Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., et al. (2014). Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data. *NeuroImage, 85*, 181–191.

Cheng, X., Li, X., & Hu, Y. (2015). Synchronous brain activity during cooperative exchange depends on gender of partner: A fnirs-based hyperscanning study. *Human Brain Mapping, 36*, 2039–2048.

Cheng, X., Pan, Y., Hu, Y., Hu, Y. (2019). Coordination elicits synchronous brain activity between co-actors: Frequency ratio matters. *Frontiers in Neuroscience*, *13*

Cui, X., Bray, S., & Reiss, A. L. (2010). Speeded near infrared spectroscopy (nirs) response detection. *PLoS ONE, 5*, e15474.

Cui, X., Bryant, D. M., & Reiss, A. L. (2012). Nirs-based hyperscanning reveals increased inter-personal coherence in superior frontal cortex during cooperation. *NeuroImage, 59*, 2430–2437.

Cutini, S., & Brigadoi, S. (2014). Unleashing the future potential of functional near-infrared spectroscopy in brain sciences. *Journal of Neuroscience Methods, 232*, 152–156.

Dai, B., Chen, C., Long, Y., Zheng, L., Zhao, H., Bai, X., et al. (2018). Neural mechanisms for selectively tuning in to the target speaker in a naturalistic noisy situation. *Nature Communications, 9*, 1–12.

Dai, R., Liu, R., Liu, T., Zhang, Z., Xiao, X., Sun, P., et al. (2018). Holistic cognitive and neural processes: a fnirs-hyperscanning study on interpersonal sensorimotor synchronization. *Social Cognitive and Affective Neuroscience, 13*, 1141–1154.

Di Lorenzo, R., Pirazzoli, L., Blasi, A., Bulgarelli, C., Hakuno, Y., Minagawa, Y., et al. (2019). Recommendations for motion correction of infant fnirs data applicable to data sets acquired with a variety of experimental designs and acquisition systems. *NeuroImage*

Dommer, L., Jager, N., Scholkmann, F., Wolf, M., & Holper, L. (2012). Between-brain coherence during¨ joint n-back task performance: A two-person functional near-infrared spectroscopy study. *Behavioural Brain Research, 234*, 212–222.

Duan, L., Liu, W.-J., Dai, R.-N., Li, R., Lu, C.-M., Huang, Y.-X., et al. (2013). Cross-brain neurofeedback: Scientific concept and experimental platform. *PloS One*, *8*

Duan, L., Dai, R., Xiao, X., Sun, P., Li, Z., & Zhu, C. (2015). Cluster imaging of multi-brain networks (cimbn): a general framework for hyperscanning and modeling a group of interacting brains. *Frontiers in Neuroscience, 9*, 267.

Feng, X., Sun, B., Chen, C., Li, W., Wang, Y., Zhang, W., et al. (2020). Self-other overlap and interpersonal neural synchronization serially mediate the effect of behavioral synchronization on prosociality. *Social Cognitive and Affective Neuroscience*

Fishburn, F. A., Murty, V. P., Hlutkowsky, C. O., MacGillivray, C. E., Bemis, L. M., Murphy, M. E., et al. (2018). Putting our heads together: interpersonal neural synchronization as a biological mechanism for shared intentionality. *Social Cognitive and Affective Neuroscience, 13*, 841–849.

Funane, T., Kiguchi, M., Atsumori, H., Sato, H., Kubota, K., & Koizumi, H. (2011). Synchronous activity of two people's prefrontal cortices during a cooperative task measured by simultaneous near-infrared spectroscopy. *Journal of Biomedical Optics, 16*, 077011.

Gero, J. S., Milovanovic, J. (2020). A framework for studying design thinking through measuring designers' minds, bodies and brains. *Design Science*

Gibson, I., Gao, Z., & Campbell, I. (2004). A comparative study of virtual prototyping and physical prototyping. *International Journal of Manufacturing Technology and Management, 6*, 503–522.

Gvirts, H. Z., & Perlmutter, R. (2020). What guides us to neurally and behaviorally align with anyone specific? A neurobiological model based on fnirs hyperscanning studies. *The Neuroscientist, 26*, 108–116.

Herold, F., Wiegel, P., Scholkmann, F., Thiers, A., Hamacher, D., & Schega, L. (2017). Functional near-infrared spectroscopy in movement science: a systematic review on cortical activity in postural and walking tasks. *Neurophotonics, 4*, 041403.

Hirsch, J., Zhang, X., Noah, J. A., & Ono, Y. (2017). Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. *NeuroImage, 157*, 314–330.

Hirsch, J., Adam Noah, J., Zhang, X., Dravida, S., & Ono, Y. (2018). A cross-brain neural mechanism for human-to-human verbal communication. *Social Cognitive and Affective Neuroscience, 13*, 907–920.

Hollander, J. E., & Carr, B. G. (2020). Virtually perfect? Telemedicine for covid-19. *New England Journal of Medicine, 382*, 1679–1681.

Holmes, E.A., O'Connor, R.C., Perry, V.H., Tracey, I., Wessely, S., Arseneault, L., et al. (2020). Multidisciplinary research priorities for the covid-19 pandemic: A call for action for mental health science. *The Lancet Psychiatry.*

Holper, L., Scholkmann, F., & Wolf, M. (2012). Between-brain connectivity during imitation measured by fnirs. *NeuroImage, 63*, 212–222.

Holper, L., Goldin, A. P., Shalom, D. E., Battro, A. M., Wolf, M., & Sigman, M. (2013). The teaching´ and the learning brain: A cortical hemodynamic marker of teacher–student interactions in the socratic dialog. *International Journal of Educational Research, 59*, 1–10.

Hou, Y., Song, B., Hu, Y., Pan, Y., & Hu, Y. (2020). The averaged inter-brain coherence between the audience and a violinist predicts the popularity of violin performance. *NeuroImage, 211*, 116655.

Hu, Y., Hu, Y., Li, X., Pan, Y., & Cheng, X. (2017). Brain-to-brain synchronization across two persons predicts mutual prosociality. *Social Cognitive and Affective Neuroscience, 12*, 1835–1844.

Hu, M., Shealy, T., Gero, J.S. (2018). Neuro-cognitive differences among engineering students when using un-structured, partially structured, and structured design concept generation techniques.

Hu, M., Shealy, T., Grohs, J., & Panneton, R. (2019). Empirical evidence that concept mapping reduces neurocognitive effort during concept generation for sustainability. *Journal of Cleaner Production, 238*, 117815.

Ikeda, S., Nozawa, T., Yokoyama, R., Miyazaki, A., Sasaki, Y., Sakaki, K., et al. (2017). Steady beat sound facilitates both coordinated group walking and inter-subject neural synchrony. *Frontiers in Human Neuroscience, 11*, 147.

Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., & Lu, C. (2012). Neural synchronization during face-to-face communication. *Journal of Neuroscience, 32*, 16064–16069.

Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., et al. (2015). Leader emergence through interpersonal neural synchronization. *Proceedings of the National Academy of Sciences, 112*, 4274–4279.

Kato, T., Okada, H., Izu, Y., et al. (2018a). Measurement of brain activities of idea generation (sketch). In DS 92: Proceedings of the DESIGN 2018 15th international design conference. pp. 2027–2034.

Kato, T., Okada, H., Izu, Y. (2017). Brain activities of idea generation using sketches.

Kato, T., Otagiri, S., Nagamori, Y., & Izu, Y. (2018). Comparison of brain activities between hand and computer drawings in finke's pattern generation task. *Journal of the Science of Design, 2*, 243–252.

Lande, M., Leifer, L. (2009). Prototyping to learn: Characterizing engineering students' prototyping activities and prototypes. In DS 58-1: Proceedings of ICED 09, the 17th international conference on engineering design, Vol. 1, Design Processes, Palo Alto, CA, USA, 24.-27.08. 2009.

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences, 114*, 13290–13295.

Liu, T., Saito, H., & Oi, M. (2015). Role of the right inferior frontal gyrus in turn-based cooperation and competition: a near-infrared spectroscopy study. *Brain and Cognition, 99*, 17–23.

Liu, N., Mok, C., Witt, E. E., Pradhan, A. H., Chen, J. E., & Reiss, A. L. (2016). Nirs-based hyperscanning reveals inter-brain neural synchronization during cooperative jenga game with face-to-face communication. *Frontiers in Human Neuroscience, 10*, 82.

Liu, T., Saito, H., & Oi, M. (2016). Obstruction increases activation in the right inferior frontal gyrus. *Social Neuroscience, 11*, 344–352.

Liu, T., Saito, G., Lin, C., & Saito, H. (2017). Inter-brain network underlying turn-based cooperation and competition: A hyperscanning study using near-infrared spectroscopy. *Scientific Reports, 7*, 1–12.

Liu, Y., Piazza, E. A., Simony, E., Shewokis, P. A., Onaral, B., Hasson, U., et al. (2017). Measuring speaker–listener neural coupling with functional near infrared spectroscopy. *Scientific Reports, 7*, 43293.

Liu, J., Zhang, R., Geng, B., Zhang, T., Yuan, D., Otani, S., et al. (2019). Interplay between prior knowledge and communication mode on teaching effectiveness: interpersonal neural synchronization as a neural marker. *NeuroImage, 193*, 93–102.

Li, R., Potter, T., Huang, W., & Zhang, Y. (2017). Enhancing performance of a hybrid eeg-fnirs system using channel selection and early temporal features. *Frontiers in Human Neuroscience, 11*, 462.

Lu, K., Xue, H., Nozawa, T., & Hao, N. (2019). Cooperation makes a group be more creative. *Cerebral Cortex, 29*, 3457–3470.

Lu, K., Teng, J., Hao, N. (2020). Gender of partner affects the interaction pattern during group creative idea generation. *Experimental Brain Research.*

Mayseless, N., Hawthorne, G., & Reiss, A. L. (2019). Real-life creative problem solving in teams: Fnirs based hyperscanning study. *NeuroImage, 203*, 116161.

Miller, J. G., Vrticka, P., Cui, X., Shrestha, S., Hosseini, S. H., Baker, J. M., et al. (2019). Inter-brainˇ synchrony in mother-child dyads during cooperation: An fnirs hyperscanning study. *Neuropsychologia, 124*, 117–124.

Nguyen, T., Schleihauf, H., Kayhan, E., Mathes, D., Vrticka, P., Hoehl, S. (2020). Neural synchrony in mother-child conversation: Exploring the role of communicative features. *Social Cognitive and Affective Neuroscience.*

Niu, R., Yu, Y., Li, Y., Liu, Y. (2019). Use of fnirs to characterize the neural mechanism of inter-individual rhythmic movement coordination. *Frontiers in Physiology*, *10*.

Noah, J. A., Zhang, X., Dravida, S., Ono, Y., Naples, A., McPartland, J. C., et al. (2020). Real-time eye-to-eye contact is associated with cross-brain neural coupling in angular gyrus. *Frontiers in Human Neuroscience, 14*, 19.

Nozawa, T., Sasaki, Y., Sakaki, K., Yokoyama, R., & Kawashima, R. (2016). Interpersonal frontopolar neural synchronization in group communication: An exploration toward fnirs hyperscanning of natural interactions. *NeuroImage, 133*, 484–497.

Nozawa, T., Sakaki, K., Ikeda, S., Jeong, H., Yamazaki, S., dos Santos Kawata, K. H., et al. (2019). Prior physical synchrony enhances rapport and inter-brain synchronization during subsequent educational communication. *Scientific Reports, 9*, 1–13.

Osaka, N., Minamoto, T., Yaoi, K., Azuma, M., & Osaka, M. (2014). Neural synchronization during cooperated humming: A hyperscanning study using fnirs. *Procedia-Social and Behavioral Sciences, 126*, 241–243.

Osaka, N., Minamoto, T., Yaoi, K., Azuma, M., Shimada, Y. M., & Osaka, M. (2015). How two brains make one synchronized mind in the inferior frontal cortex: Fnirs-based hyperscanning during cooperative singing. *Frontiers in Psychology, 6*, 1811.

Pan, Y., Cheng, X., Zhang, Z., Li, X., & Hu, Y. (2017). Cooperation in lovers: An f nirs-based hyperscanning study. *Human Brain Mapping, 38*, 831–841.

Pan, Y., Novembre, G., Song, B., Li, X., & Hu, Y. (2018). Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song. *NeuroImage, 183*, 280–290.

Pan, Y., Dikker, S., Goldstein, P., Zhu, Y., Yang, C., & Hu, Y. (2020). Instructor-learner brain coupling discriminates between instructional approaches and predicts learning. *NeuroImage, 211*, 116657.

Pan, Y., Guyon, C., Borragan, G., Hu, Y., Peigneux, P. (2020b). Interpersonal brain synchronization with instructor compensates for learner's sleep deprivation in interactive learning. Biorxiv

Pappot, N., Taarnhøj, G. A., Pappot, H. (2020). Telemedicine and e-health solutions for covid-19: Patients' perspective. *Telemedicine and e-Health.*

Pfefferbaum, B., North, C.S. (2020). Mental health and the covid-19 pandemic. *New England Journal of Medicine.*

Piazza, E. A., Hasenfratz, L., Hasson, U., & Lew-Williams, C. (2020). Infant and adult brains are coupled to the dynamics of natural communication. *Psychological Science, 31*, 6–17.

Piva, M., Zhang, X., Noah, J. A., Chang, S. W., & Hirsch, J. (2017). Distributed neural activity patterns during human-to-human competition. *Frontiers in Human Neuroscience, 11*, 571.

Quaresima, V., & Ferrari, M. (2019). Functional near-infrared spectroscopy (fnirs) for assessing cerebral cortex function during human behavior in natural/social situations: A concise review. *Organizational Research Methods, 22*, 46–68.

Quinones-Camacho, L.E., Fishburn, F.A., Camacho, M.C., Hlutkowsky, C.O., Huppert, T.J., Wakschlag,˜ L.S., et al. (2019). Parent–child neural synchrony: A novel approach to elucidating dyadic correlates of preschool irritability. *Journal of Child Psychology and Psychiatry*

Reindl, V., Gerloff, C., Scharke, W., & Konrad, K. (2018). Brain-to-brain synchrony in parent-child dyads and the relationship with emotion regulation revealed by fnirs-based hyperscanning. *NeuroImage, 178*, 493–502.

Sarinasadat, H., Hattori, Y., Miyake, Y., Nozawa, T. (2019a). Music valence and genre influence group creativity. In *International Conference on Human-Computer Interaction* (Springer), pp. 410–422

Sarinasadat, H., Miyake, Y., Nozawa, T. (2019b). Combined effects of background music and nonverbal synchrony measures on group creativity—a multiple regression approach. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC) (IEEE)*, pp. 3752–3757.
Scholkmann, F., Holper, L., Wolf, U., & Wolf, M. (2013). A new methodical approach in neuroscience: Assessing inter-personal brain coupling using functional near-infrared imaging (fniri) hyperscanning. *Frontiers in Human Neuroscience, 7*, 813.
Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Pavia, J. M., Wolf, U., et al. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *NeuroImage, 85*, 6–27.
Schwartz, A. M., Wilson, J. M., Boden, S. D., Moore, T. J., Jr., Bradbury, T. L., Jr., & Fletcher, N. D. (2020). Managing resident workforce and education during the covid-19 pandemic: Evolving strategies and lessons learned. *JBJS Open Access, 5*, e0045.
Shealy, T., Gero, J. (2019). The neurocognition of three engineering concept generation techniques. In *Proceedings of the Design Society: International Conference on Engineering Design* (Cambridge University Press), vol. 1, 1833–1842.
Shealy, T., Grohs, J., Hu, M., Maczka, D., & Panneton, R. (2017). *Investigating design cognition during brainstorming tasks with freshmen and senior engineering students using functional near infrared spectroscopy*. ASEE.
Shealy, T., Hu, M., Gero, J. (2018). Patterns of cortical activation when using concept generation techniques of brainstorming, morphological analysis, and triz. In *International design engineering technical conferences and computers and information in engineering conference (American Society of Mechanical Engineers)*, vol. 51845, V007T06A035.
Sun, B., Xiao, W., Feng, X., Shao, Y., Zhang, W., & Li, W. (2020). Behavioral and brain synchronization differences between expert and novice teachers when collaborating with students. *Brain and Cognition, 139*, 105513.
Tachtsidis, I., & Scholkmann, F. (2016). False positives and false negatives in functional near-infrared spectroscopy: Issues, challenges, and the way forward. *Neurophotonics, 3*, 031405.
Takeuchi, N., Mori, T., Suzukamo, Y., & Izumi, S.-I. (2017). Integration of teaching processes and learning assessment in the prefrontal cortex during a video game teaching–learning task. *Frontiers in Psychology, 7*, 2052.
Tang, H., Mai, X., Wang, S., Zhu, C., Krueger, F., & Liu, C. (2016). Interpersonal brain synchronization in the right temporo-parietal junction during face-to-face economic exchange. *Social Cognitive and Affective Neuroscience, 11*, 23–32.
Vanzella, P., Balardin, J. B., Furucho, R. A., Zimeo Morais, G. A., Braun Janzen, T., Sammler, D., et al. (2019). fnirs responses in professional violinists while playing duets: Evidence for distinct leader and follower roles at the brain level. *Frontiers in Psychology, 10*, 164.
Xue, H., Lu, K., & Hao, N. (2018). Cooperation makes two less-creative individuals turn into a highly-creative pair. *NeuroImage, 172*, 527–537.
Yang, J., Zhang, H., Ni, J., De Dreu, C. K., Ma, Y. (2020). Within-group synchronization in the prefrontal cortex associates with intergroup conflict. *Nature Neuroscience*, pp. 1–7.
Zhang, M., Liu, T., Pelowski, M., Jia, H., & Yu, D. (2017a). Social risky decision-making reveals gender differences in the tpj: A hyperscanning study using functional near-infrared spectroscopy. *Brain and Cognition, 119*, 54–63.
Zhang, M., Liu, T., Pelowski, M., & Yu, D. (2017b). Gender difference in spontaneous deception: A hyperscanning study using functional near-infrared spectroscopy. *Scientific Reports, 7*, 1–13.
Zhang, Y., Meng, T., Hou, Y., Pan, Y., & Hu, Y. (2018a). Interpersonal brain synchronization associated with working alliance during psychological counseling. *Psychiatry Research: Neuroimaging, 282*, 103–109.
Zhang, M., Ding, K., Jia, H., Yu, D. (2018b). Brain-to-brain synchronization of the expectation of cooperation behavior: A fnirs hyperscanning study. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (IEEE), pp. 546–549.

Zhao, Y., Dai, R.-N., Xiao, X., Zhang, Z., Duan, L., Li, Z., et al. (2017). Independent component analysisbased source-level hyperlink analysis for two-person neuroscience studies. *Journal of Biomedical Optics, 22*, 027004.

Zheng, L., Chen, C., Liu, W., Long, Y., Zhao, H., Bai, X., et al. (2018). Enhancement of teaching outcome through neural prediction of the students' knowledge state. *Human Brain Mapping, 39*, 3046–3057.

Zheng, L., Liu, W., Long, Y., Zhai, Y., Zhao, H., Bai, X., et al. (2020). Affiliative bonding between teachers and students through interpersonal synchronisation in brain activity. *Social Cognitive and Affective Neuroscience*

Contemporary Issues in Remote Design Collaboration

Lawrence Domingo, Marius Gutzeit, Larry Leifer, and Jan Auernhammer

Abstract The COVID-19 Pandemic has exacerbated challenges in remote design collaboration during the past year. Public policy has shifted almost all activities into a remote format. This separation in physical space results in challenges in design collaboration and design work. Three topics that have become more complicated during our study during the pandemic and are explored here. These topics are the Mind–Body Dichotomy, Perspective, and Wicked Problems.

The COVID-19 Pandemic has exacerbated challenges in remote design collaboration during the past year. Public policy has shifted almost all activities into a remote format. This separation in physical space results in challenges in design collaboration and design work. Three topics that have become more complicated during the duration of our study during the pandemic and are explored here. These topics are the Mind–Body Dichotomy, Perspective, and Wicked Problems.

1 Mind–Body Dichotomy

The name *Design Thinking* has led to design courses that overemphasize conceptual work at the cost of physical expression. This Mind–Body separation is similar to the Cartesian philosophy founded by Rene Descartes. The Cartesian philosophy separates things for analysis. For example, physics kinematics problems are often subdivided into three spatial directions: x, y, and z, in order to analyze a system behavior.

L. Domingo (✉) · M. Gutzeit · L. Leifer · J. Auernhammer
Stanford Center for Design Research, Stanford University, Building 560, 424 Panama Mall, Stanford, CA 94305-2232, USA
e-mail: LDomingo@stanford.edu

L. Leifer
e-mail: leifer@stanford.edu

J. Auernhammer
e-mail: jan.auernhammer@stanford.edu

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_3

In the case of design thinking, mind (thinking) and body (physical manifestation) are often divorced into separate steps of the design process. While this works in the theoretical physics space, challenges emerge in the design space. A counter philosophy outlined by McKim (1972, 1980), and inspired by Esalen Institute's founders, Mike Murphy and Dick Price, sought to instead explore the integration of mind and body. Esalen's philosophies influenced Stanford's Design Division/Group resulting in a pedagogical approach of designing for the whole person—and not just for the thinking mind.

Curriculum, to fit into one class session, can inadvertently divorce design concepts or processes, such as thinking and acting, into separate buckets without actively interconnecting the parts. A consequence of this separation is that individuals can engage in one part of the design process without experiencing the whole breadth of designing a solution through the entire design process. As a consequence, individuals may not build empathy for the complimenting design processes.

Donald Schön (1983) studied how expert designers work as reflective practitioners. Designers interact with the physical world and then reflect on what the next course of action would be, often in implicit ways. An example of an activity that requires both reflection and action is wood whittling. Certain aspects of wood whittling can be taught as verbal instruction, but the act of wood whittling itself requires kinesthetic knowledge. If a whittler applies too much force, the wood will give and too much will be carved away. If a whittler applies too little force, the wood will not de-form and shave material away. The body needs to learn what force and at what rate the wood can be whittled without damaging its grain. This kinesthetic knowledge could then be used to manage project timelines acknowledging how long certain activities can take that cannot be rushed. For some undergraduate design courses 3D printers were sent to students to augment the manufacturing experience and to develop kinesthetic knowledge.

Movement to think is supported by more recent work conducted by Barbara Tversky (2019) on the nature of spatial knowledge and spatial reasoning. Tversky (2019) argues that gestures occur more instinctively and quickly than words since words are abstractions used to describe things in the world. Gestures on the other hand are direct interactions in the world. Tversky (2019) describes a study where participants had to sit on their hands to solve math problems in order to limit gesturing. Participants experienced higher cognitive loads when they were forced to not move their hands to help solve math problems. During our study, team members were not able to co-engage due to the lack of gestural communication. Participants could not express themselves through physical movement. Participants could not move to think that same way as in-person.

Furthermore, the deictic gesture of pointing can help bring about shared attention in a 3-dimensional world. Movement to think goes beyond hand gestures. Thinking can also be facilitated with walking, a reflective activity that can help us escape our ubiquitous digital screens, which has decreased in frequency due to the pandemic. Furthermore, lab activities, such as in manufacturing courses, have significantly decreased, forcing design to become almost entirely in the head as a thinking activity. The long-term consequences of the reduction in hands-on learning are yet to be observed. Schön's (1983) and Tversky's (2019) work shows us that for certain activities people must engage in physical movement to think. Knowledge work is not solely in the head. When we limit ourselves to the confines of our desk, we limit our means to think in the physical world.

We observed in our study that communication was limited to verbal discourse. What could be explained in a shared annotated sketch or gestural enactment was explained in long back-and-forth explanations with ideas that were built upon using lists. During our study, participants seemed limited by their shared mental models which slowed the brainstorming process down. If team meetings did not have time limits, this is not necessarily an issue. However, time is the only real limited commodity in a design project or in a meeting, and once time is exhausted, the project is "done." Time spent giving explanations, takes away from time available for developing ideas or exploring alternatives. When more efficient means for explaining, such as sketching or gesturing, are limited, time spent for developing is subtracted even further. Laying out shared mind maps can reveal multi-dimensional interconnections to design teams that are difficult to capture through linear lists. These design sessions were unnecessarily cognitively demanding knowledge work that could be facilitated by shared spatial work.

Physical prototypes can function as a 3-dimensional map of reality that helps teams orient themselves and their shared mental models to facilitate collaboration. When teams are physically separated in remote meetings, 3D objects are limited to 2D expressions in modeling software, and 2D mind maps may be limited to 1D document lists. The physical act of pointing at a prototype or conceptual map aids in the cognitive collaboration work of brainstorming. Future work should use neuroscientific instruments to measure the cognitive load that teams experience when the use of shared spatial media is limited.

2 Perspective

In response to ethnic conflicts in the United States, academic communities took time to reflect on ways their disciplines can become more inclusive. Based on the present lack of diversity in STEM fields, various academic communities have sought out ways to support broader social integration. However, building empathy with marginalized groups proves a significant challenge when done through remote, digital media. Designers must build empathy with a diverse group of users when creating or scaling products including biological sex in product design or infrastructure for lower socioeconomic groups.

For example, the biomechanics community reflected on how aggregated data sets do not always accurately represent all humans. Caroline Perez (2019) gives an example of how seat belts in vehicles are not designed for pregnant women. The kinematics of a pregnant woman is quite different from the kinematics of the average male. A pregnant woman has a much lower center of mass, wider hips, and a different mass distribution. These changes in inertia require a different seat belt design that does not constrict a pregnant woman's abdomen, and thus provides equal seat belt safety efficacy. Design teams often collect data based on what is available in their networks. Therefore, if an engineering team is made up exclusively of males, the perspective of women could be excluded. The lack of perspective that leads to data disparities and informs product design can put users at a higher risk.

Stanford University has a long history of encouraging students to start their own companies that help fill unmet needs. The startups are often supported by venture capitalists who occupy the bordering Sand Hill Road. Stanford also neighbors East Palo Alto (EPA), a city that has historically struggled economically. While Stanford students have access to computer labs which can supplement their personal laptops, some community members in EPA's StreetCode, a computer education group, were unable to own a laptop before remote school became mandatory. During the pandemic, in a different design project, one of the authors interviewed StreetCode and found out that they gave laptops to families from lower socioeconomic backgrounds that could not afford a laptop. The group would later learn that merely handing out new technology is unhelpful unless the user knows how to use the technology. For example, some community members who received free laptops struggled to locate the "Start" menu in Windows on their computer. In response to this insight, StreetCode developed new courses in the community that would introduce the fundamentals of using computers—from typing up a document to browsing the Internet. StreetCode expanded their curriculum beyond coding fundamentals by considering and working with user groups who had not owned a computer in the past. This disparity in technical knowledge can be found no further than one freeway overpass away. Although education is often touted as the "great equalizer" this is not always the case, and without giving careful attention to the specific situation it can even exacerbate disparities.

Perspective is important on a design project. Nicholas Nassim Taleb (2007) argues in *The Black Swan* and Paul Feyerabend (1993) in *Against Method* that perspective helps us gain a new story. A new story helps us invest in collecting new data through new tool development or time to explore. New data helps to inform our (design) decisions. Without the proper perspective, we can have a blind spot in our perspectives and subsequently in our data. However, perspective is not enough in design teams, organizations, and entire regions, a culture of psychological safety is required to acknowledge and name uncomfortable problems, such as the impact of age in triage or sharing body fluids and infections in product design.

Some design courses today emphasize interview methods over observational techniques. While the reasons vary with each discipline, the consequences are noticeable as poorer insights during design reviews. The pandemic has exacerbated these issues because of safety concerns with conducting live observational studies with people, and instead limiting interaction to online interviews. Methods beyond interviews can quickly reveal contradictions between intended reasons and actual behavior. Interviewees are not always aware of the ways they may unintentionally fib to others or themselves. Such discrepancies can be difficult to capture in a scripted story alone. As designers, we must capture diverse perspectives, and this goes beyond just building empathy through interviews and stories. In the past, designers were able to participate in activities that would help build empathy along with users. For example, visiting an elderly person's home who has difficulty walking can reveal that the front door has four steps, which make it difficult to enter the house. This might not necessarily be revealed in an interview conducted remotely since such obstacles can eventually become normalized. During this pandemic, we are limited in how we visit people and are restricted to remote interaction. However, we can still capture perspectives and build empathy by diversifying our teams to include team members who have built empathy through lived experience.

3 Wicked Problems

Project-based courses afford themselves a pedagogy that ground students in reality. For example, students actively apply the materials science of metal crystal structures in a manufacturing course while turning a metal rod on a lathe using cutting tools with different hardness specifications for either aluminum or steel. The psychology of user experience design can be applied in computer science courses. Semester-long and quarter-long courses limit the scope of project prompts to make the time spent manageable. Wicked problems are dynamic, interconnected, and complex where today's solutions become tomorrow's problems (Dorst, 2015; Rittel & Webber, 1973). Solutions to wicked problems cannot be myopic to short time scales at the consequence of the future and must take the long-term problem evolution into account.

Our twentieth century mindset has helped solve our twentieth century wicked problems. However, those solutions have led to our twenty-first century problems and we need a new paradigm to solve our twenty-first century problems. Kate Raworth (2017) in her book *Doughnut Economics* describes characteristics of how we think in this new paradigm. First, we need to shift our design goals and reevaluate our metrics of success to adapt to constantly evolving and dynamic wicked problems. Economic growth as measured by gross-domestic product (GDP) should not be our design goal. The pandemic has shown us that GDP growth can accelerate despite severe job loss in certain sectors. GDP was originally used in the twentieth century to measure economic progress during the Great Depression in the US. However, GDP should not be the only metric we use to reflect on the goals we are trying to attain and to determine the new metrics for what added value means—whether it be happiness, good health, or well-being. We must keep a balance between economic growth and environmental sustainability. Environmental sustainability is as directly linked to well-being as the air we breathe (Stoknes, 2015).

Designers must recognize their role as actors in a broader system and how the decisions they make can have far-reaching effects distant in space and time. Designers must acknowledge that rather than being rational creatures, humans are rationalizing creatures. Humans do not always make the optimal decision since we cannot predict all possible futures and wrong decisions can irreversibly affect the future (Rittel & Webber, 1973; Simon, 2019). This is especially the case when infrastructure or policy are involved. We must solve the problems of today without compromising the needs of the future. Furthermore, we need to solve our local problems without inadvertently hurting others who are far away, whether in space or time. Designers need to learn to think nonlinearly. For instance, engineers love linear equations like Newton's Second Law: F = ma. However, not all phenomena follow linear models including Newton's Second Law. One of Albert Einstein's contributions to science was his adjustment to Newton's Second Law as objects approached the speed of light. Objects do not accelerate infinitely and instead, at least in our current observations, approach the speed of light (Born, 2012). Many social system phenomena behave in similar ways beyond linear relationships. Systems can grow exponentially, but also reach equilibria. Hockey stick graphs are almost always S-curves in disguise (Brand, 2008). The linear thinking persists which does not solve the wicked nature of design problems.

The Black Plague from the mid-1300's, much like the COVID-19 pandemic, grew exponentially and when momentum of spreading built up, it was difficult to slow that momentum down. Furthermore, the Black Plague, as theorized in some anthropologic communities, eventually reached an equilibrium in spread because it ran out of people to infect and the population thereby achieved herd immunity. However, this herd immunity was attained with a great loss of life. Mental models of the past that model rising growth do not model the slowing equilibria of the later stages of an S-curve. As designers, we must recognize that hockey stick growth or behavior is not the only mode and that change is inevitable.

Wicked problem-themed prompts can be built upon throughout a student's academic career. The pandemic has revealed that students are interested in wicked problems and applying their knowledge to current events. In this way, students can develop their own expertise by applying what they learn as reflective practitioners. Following the WHO's declaration of a pandemic, university communities scrambled to contribute to the mitigation and spread of disease. University-organized hackathons during the pandemic attracted students from around the globe who wanted to try and help by applying what they had learned.

Instead of limiting the scope of projects from the beginning, project prompts can be left as open-ended wicked problem themes. In a follow-up coaching session, Student design teams can focus on a project scope for that quarter or semester that applies the techniques learned in the course. The problem theme can persist in different courses and students can then apply what they learned from past courses in more meaningful ways. The details of how the projects can be graded are left for the course instructor, but students can then reflect on the most important question about any design problem—what to do next? If through the course's activities and project feedback the student can discern the next steps for value added work, then the lesson for how to persist in working on a wicked problem is learned. Wicked problems are never truly solved but evolve into new problems.

4 Conclusion

The COVID-19 pandemic resulted in emerging design obstacles. New challenges will emerge as remote collaboration becomes the new status quo. The Mind–Body dichotomy, diverse perspectives, and wicked problems will be major challenges in remote collaborative Design Thinking. In design teams we need to socialize to develop personal relationships and psychological safety to enable multiple perspectives to tame the wicked nature of design problems. This will be a major challenge in remote design collaboration. The Mind–Body dichotomy is not the only problem: individuals are physically separated. While we remain connected through the Internet, tribal groups in society based on ideology or interests are inadvertently clustered. This separates diverse perspectives. We must foster a culture of psychological safety to engage in a dialogue to foster open thinking in design and that allows the taming of wicked problems. To redesign our stay-at-home lifestyle, we should seek ways to be more intentional in integrating knowledge work with the physical world so that we are more grounded in reality and communicate more effectively. To augment our inaccessibility to interact with people, we should intentionally form teams that can provide a perspective for our blind spots. To prepare us for a world of wicked problems, we alter our pedagogies to engage with wicked problems.

References

Born, M. (2012). *Einstein's theory of relativity*. Dover Publications.

Brand, S. (2008). *The Clock Of The Long Now: Time and Responsibility*: Basic Books.

Dorst, K. (2015). *Frame innovation: Create new thinking by design.* MIT Press.

Feyerabend, P. (1993). *Against method*. Verso.

McKim, R. H. (1972). *Experiences in visual thinking*. Wadsworth Publishing Company Inc.

McKim, R. H. (1980). *Experiences in Visual Thinking*: Brooks/Cole Publishing Company.

Perez, C. C. (2019). *Invisible women: Data bias in a world designed for men.* ABRAMS.

Raworth, K. (2017). *Doughnut economics. Seven Ways to Think Like a 21st-Century Economist*: Chelsea Green Publishing.

Rittel, H., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy Sciences, 4*(2), 155–169. https://doi.org/10.1007/BF01405730

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action.* Basic Books.

Simon, H. A. (2019). *The sciences of the artificial*. MIT press.

Stoknes, P. E. (2015). *What we think about when we try not to think about global warming: Toward a new psychology of climate action.* Chelsea Green Publishing.

Taleb, N. N. (2007). *The black swan: The impact of the highly improbable.* Random House Publishing Group.

Tversky, B. (2019). *Mind in motion: How action shapes thought*. Basic Books.

Mapping Design Thinking in the Arab World

Hanadi Traifeh, Reem Abou Refaei, Julia von Thienen, Karen von Schmieden, Lena Mayer, Sherif Osman, and Christoph Meinel

Abstract Design Thinking is a human-centered approach to innovation that has become increasingly popular globally over the last decade. While the spread of Design Thinking is well understood and documented in Europe and North America, this is not the case when it comes to the Arab world. This study explores the history, application, and understanding of Design Thinking in the Arab world through a systematic literature review based on the guidelines of the PRISMA framework. Analysis of collected data revealed that there is no widely accepted Arabic translation of the term "Design Thinking." This study identified the first appearance of an Arabic translation of "Design Thinking," coined in 2010 as "التفكير التصميمي." Another major contribution of the study is the aggregation of published articles associated with the different Arabic translations of "Design Thinking," documenting findings and mapping out the spread of Design Thinking in the Arab region in five sectors: education, development, private, entrepreneurship, and the public sector. The results reported in this paper are part of a large-scale study that connects a range of research methods to track the development of Design Thinking in the Arab world.

H. Traifeh (✉) · R. Abou Refaei · J. von Thienen · K. von Schmieden · L. Mayer · S. Osman · C. Meinel
Hasso Plattner Institute for Digital Engineering, Campus Griebnitzsee, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany
e-mail: hanadi.traifeh@hpi.de

R. Abou Refaei
e-mail: Reem.AbouRefaie@hpi.de

J. von Thienen
e-mail: julia.vonthienen@hpi.de

K. von Schmieden
e-mail: Karen.Schmieden@hpi.de

L. Mayer
e-mail: Lena.Mayer@hpi.de

S. Osman
e-mail: Sherif.Osman@hpi.de

C. Meinel
e-mail: meinel@hpi.de

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_4

1 Introduction

Design Thinking as an approach for innovation and creativity has been widely implemented in different fields and used in many aspects of life (Plattner et al., 2009; Meinel et al., 2015). The application of Design Thinking, however, may differ from one field to another. To understand how Design Thinking is understood and applied in the corporate context, researchers from the Hasso Plattner Institute (HPI) conducted a study in 2015, which was the first large-sample survey of Design Thinking adoption in practice, titled "The Current State of Design Thinking Practice in Organizations" (Schmiedgen et al., 2016). The study provided first evidence on how Design Thinking is perceived and applied by the business community worldwide. However, there was a significant lack of survey respondents from Arabic-speaking countries, making it difficult to draw conclusions on the spread of Design Thinking there, or how to further report on the ways in which the methodology is applied in this region. To fill this gap, we conducted a large-scale study to investigate the adoption of Design Thinking among individuals as well as organizations in Arabic-speaking countries. This overall study uses multiple methodological approaches to explore and understand the development of Design Thinking in the region, including online surveys, semi-structured interviews and social media analysis. In this paper, we provide the first report of the research outcomes based on a systematic literature review following the guidelines of the PRISMA framework. Our review covers the entire year of 2010 (the first time the Arabic translation of Design Thinking "التفكير التصميمي" was coined) and extends up to, and includes May 2019. The questions guiding our study were as follows:

1. What is the most widely adopted Arabic term of Design Thinking, when did it first appear and who coined it?
2. Who are the organizations/people most active in promoting Design Thinking in the Arab world (for-profit or non-profit, government, others?)
3. Which industries/sectors in the Arab world utilize the concept of Design Thinking the most?

2 Research Design

In order to provide a comprehensive assessment of the adoption of Design Thinking in the Arab world, a systematic literature review was conducted following the guidelines of the PRISMA framework (Liberati et al., 2009) to ensure the quality and transparency of the review process. The review consisted of a four stage process: (1) Identification of keywords; (2) Developing the search strategy; (3) Relevance-based filtration and selection; and (4) Reporting on findings. Our review covers the entire year of 2010 (the first time the Arabic translation of Design Thinking "التفكير التصميمي" was coined) and extends up to, and includes, May 2019.

1. **Identification of keywords**

To establish the widest possible search of Design Thinking in the Arab World, we started by scouting for potential keywords that could be associated with Design Thinking in the Arabic language. The initial mapping of all terms that are associated with Design Thinking showed that there is no widely accepted translation of the term "Design Thinking" in Arabic, and that there were in fact several terms used to refer to the subject. These terms include التفكير التصميمي, الفكر التصميمي, تصميم التفكير and التصميم المتمحور حول الإنسان. A closer inspection of the different translations used among the Design Thinking community in the Arab world led us to add several other English keywords to capture the related English content published by Arabic-speaking authors. Table 1 presents the range of keywords that were identified as referring to Design Thinking in Arabic and English.

2. **Developing the search strategy**

Our search process consists of three steps. First, we searched a number of research databases (Academia, Research Gate, Institute for Education Sciences, Science Direct and Google Scholar) using the broad search keywords that were previously identified in Arabic and English. We searched for articles published in English and Arabic from January 2000 up to and including May 2019, which we assumed to be an appropriate time period that captures the global emergence and growth of Design Thinking. We found 4 results in Arabic and 16 in English. In the second step, we used the terms "Design" and "Arab" to include as many different approaches as possible. It is possible that scholars may apply design without labeling it as such and, therefore, some relevant results may have not been retrieved in our searches. However, adding more search terms could increase risks of contamination of the findings. Based on the limited number of peer-reviewed search results, we concluded that the field of Design Thinking in the Arab world was probably mostly confined to the practical realm and has not yet been properly studied by the academic research community. Therefore, as a third step, we expanded our discovery by conducting a web search for attempts, initiatives, workshops, trainings, articles, blog posts and videos that mentioned Design Thinking. This expanded search yielded 150 results, which supports our initial assumptions about the wider adoption of Design Thinking outside the research community.

Table 1 Keywords used for the systematic literature review

Keywords	
Arabic keywords	English keywords
التفكير التصميمي	Design Thinking Arabic
الفكر التصميمي	Design Thinking Arab
التصميم المتمحور حول الانسان	Design Arab
التفكير تصميم	Human Centered Design Arab

3. **Filtration and selection**

We decided to adopt the widely accepted conceptualization of Design Thinking as a process, a mindset, and a human-centered approach to creativity, collaboration and innovation. We used a two-step process for content selection, both focused on the content's relevance to this conceptualization of Design Thinking. First, we assessed titles and abstracts only. For the most part, we excluded studies that were not relevant to Design Thinking–for example, content that did not address Design Thinking as a concept, a process or a mindset. Second, we conducted a full-text analysis of all content in order to ensure relevance. In this step, many articles were excluded because of their misinterpretation of Design Thinking. This process resulted in the inclusion of 11 academic articles and 34 web links (Fig. 1).

4. **Report on main findings**

The results section is divided into five sub-sections, each reflecting the findings of a specific sector with notable Design Thinking activity: education, development, private sector, entrepreneurship, and public sector.

3 Coining the Arabic Term for Design Thinking

In order to identify the first instance of the Arabic term of Design Thinking, an online search using different Arabic translations of "Design Thinking" was carried out. Using the search engine Google, each of the Arabic terms was searched with date filters applied until the first Arabic term "التفكير التصميمي" appeared in 2010. The first introduction of the Arabic term of Design Thinking "التفكير التصميمي" appeared in a translation of Tim Brown's Ted Talk: *Designers—think big*.[1] This translation for Design Thinking was coined by Chafic Jaber, a research engineer at Telecom ParisTech, who—based on his bio—does not seem to have significant involvement with Design Thinking as a concept. Translated into 22 languages, *Designers—think big!* covers a snippet of the evolution of Design Thinking: from the pre-design era in the nineteenth century, where "systems thinkers were reinventing the world" to the twenty-first-century commercialized design era led by a few, whom he refers to as the "priesthood of folks in black turtlenecks and design glasses." Brown argues for an urgent need to reclaim design not as a tool to produce high street products but as a way to redefine global problems. In doing so, Brown urges people practicing Design Thinking to "Think Big" and to "start asking the right questions" (Brown, 2009).

[1]Tim Brown is the Chair of IDEO, a renowned international design and innovation company which roots go back to 1978. Brown has been playing a vital role as an ambassador of Design Thinking all over the world.

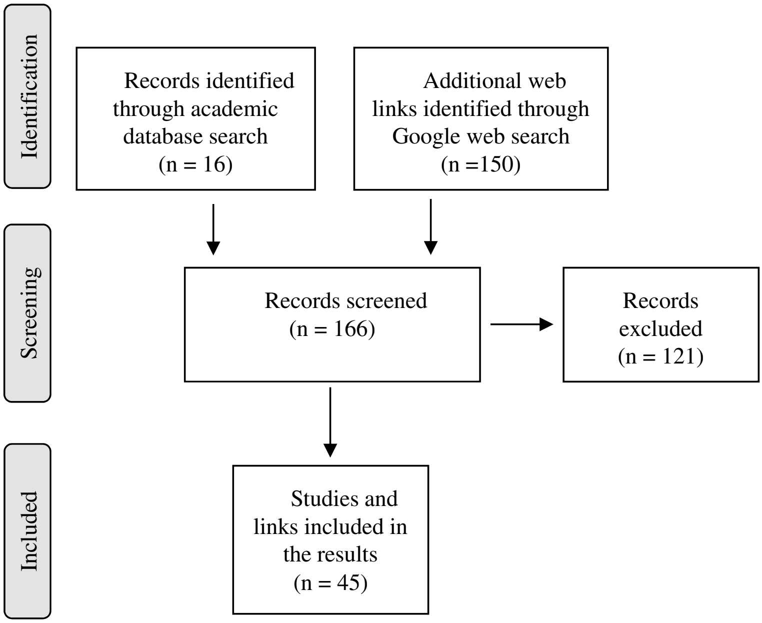


Fig. 1 Filtration of identified articles and selection process

Although the translation of Design Thinking into التفكير التصميمي was coined in 2010, there appeared to be an ongoing discussion on the appropriate way to translate the concept to Arabic. This was picked up in 2015 by a blogger on oktob.io, who claimed the term التفكير التصميمي is the correct translation of the concept "Design Thinking." She rejected the literal translation تصميم التفكير that some authors were using, and instead advocated for the adoption of التفكير التصميمي as the official translation. She made the case that the former term refers to the idea of visualizing one's thoughts, which is only one aspect of Design Thinking, whereas the latter term comprehensively captures all facets of the concept (Aldakheel, 2015). Another blogger, Salah Taha (2019), addressed the confusion over the translation of Design Thinking in Arabic, when he emphasized that the fitting translation of the term Design Thinking is التفكير التصميمي and not تصميم التفكير because the intended meaning is the way of thinking and not the design of one's thinking (Taha, 2019). In 2019, The Shoman Cultural Forum, in cooperation with the Ideas Lab, hosted a lecture in Amman, Jordan, on "Design Thinking" by Tania Amaysi from Stanford University. The host introduced the lecture by discussing whether to use a different term than التفكير التصميمي so that audiences may understand it better. However, she concluded by claiming that التفكير التصميمي is the Arabic term of "Design Thinking" and as of today that is the term most widely used in the region (Shouman Foundation, 2019).

4 Early Attempts to Understand Design Thinking in Arabic

Recognizing the lack of Arabic content on Design Thinking, an unknown Arabic-speaking blogger–owner of a blog called Arabstarts-decided to address the situation. She attempted to fill the evident knowledge gap by beginning to address the topic in Arabic through a series of blog posts. In her first article, she explained the concept of Design Thinking and its function as a management methodology. She highlighted applications in the private sector by companies such as Apple, Yahoo and General Electric, all of which were in the area of management and strategic planning. This was followed by a more detailed investigation, where she attributed Apple's unparalleled success to CEO, Steve Job's adoption of a Design Thinking management style. She further explained that Design Thinking is essentially a methodology used by designers to solve "design" problems. However, over the years, evidence has shown that the way designers analyze problems and generate solutions is potentially transferable to "non-design" fields. This transferability hinges on four main factors: (1) Creating an environment that is conducive to creativity and innovation; (2) Focusing on the user experience; (3) Continuous testing and learning from trial and error approaches; and (4) Incorporating failure as an integral element of the testing process (Unknown, Arabstarts, 2012). It is worth mentioning that the blogger did not write any further articles about the topic after this contribution.

Other voices from the Arab World included an online learning platform (Xschool) with presence on a YouTube Channel (64,000 subscribers), a Facebook page and a Twitter account. The platform's main goal is to engage the audience through short videos which dwell on a number of concepts dealing with problems that concern the general public. Despite a long list of produced videos, only one video introduced Design Thinking, or what they refer to as[2] تصميم التفكير as a problem-solving approach that relies on perpetual testing to generate the most suitable solutions (Aldawood, 2012).

In Al-Dousary and Al-Robayaa (2013) two researchers from Saudi Arabia, published a SlideShare presentation on LinkedIn about Design Thinking. The presentation introduces Design Thinking in reference to education. They explained that Design Thinking is a problem-solving approach which relies on exposing students to real problems and allowing them to brainstorm solutions, test them in real life, and learn from their experiences. The researchers refer to Design Thinking as تصميم التفكير or التفكير التصميمي. They highlight possible applications in the private sector with a specific focus on technology companies like Apple. Converging with earlier contributors, they emphasized that Design Thinking is a methodology is derived from the way designers think. However, over time, the methodology started to spread into other fields as more people recognized its merits in generating

[2]This Arabic word is one of the translations that other bloggers objected to due to the literal translation of the term, which they believe does not reflect the true meaning of the concept.

innovative solutions for complex problems. They also explained how the application of the methodology in "gamification" can facilitate learning for students in preschool to primary school levels (Al-Dosary & Al-Robayaa, 2013).

5 Mapping Design Thinking in the Arab World

5.1 Design Thinking and Education

It seems that in the Arab World, Design Thinking was widely experimented with by young student enthusiasts in various levels of the education system. In 2011, a second year female graphic design student in Saudi Arabia came across the book "Design Thinking Basics 08" by Gavin Ambrose and Paul Harris. She gave a presentation in Arabic about the topic and explained the concept and the phases of the design process. An architect who attended the presentation was fascinated by what she heard and reported on the presentation in a blog post in her series of design posts "Design…Ask" (Al Shaddy, 2011). This blog post is the first article that was identified on the web that mentioned Design Thinking in Arabic after the term was coined in 2010 by Jaber.

More thorough attempts to apply Design Thinking to the education sector came out of the Khalifa University of Science, Technology and Research (KU) in the United Arab Emirates. KU's leadership vision was to internationalize and vertically integrate inquiry-based design pedagogy into all of its undergraduate degree programs in the College of Engineering. This was achieved through strategic partnerships with institutions like Georgia Institute of Technology in the US and KAIST in South Korea. The first step in the integration was the creation of a college-wide, "cornerstone" freshman engineering design experience that meets the needs and context of KU's diverse international student body, and uses effective pedagogical alternatives to lecture-based instruction. A paper titled "Cultivating design-thinking in freshmen: The evolution of the KU freshman design course," published in 2013 by Shadi Balawi et al., describes the research-driven evolution of the freshman design experience at KU from a discipline-specific offering within the departments of aerospace and mechanical engineering, to an interdisciplinary, college-wide Freshman Engineering Design Course (FEDC) (Balawi et al., 2013)

By 2018, education became an area where Design Thinking flourished in the Arab World. Ibrahim Wa'ed (2018) attempted to analyze the presence and possibilities of social design in Arab undergraduate design education and its related courses at a renowned university in the Middle East (the university name is kept confidential). His argument was centered on the fact that social design has the potential to deliver sustainable solutions for numerous social, economic, and environmental challenges, particularly in the Arab World. The main findings suggest a basic current understanding about social design at the university. The research identified a number of existing courses, studio projects, and other assignments that briefly mention social design principles. However, the results highlighted that the subject requires further integration as many design students were unfamiliar with the current global challenges, and also exhibited a lack of critical thinking. This deficit showed itself in an inadequate design process and a limited participation in interdisciplinary collaborative work. Several recommendations are made to facilitate a stronger social design presence in the curriculum such as introducing the concepts of participatory design/co-design, design thinking, and critical thinking (Wa'ed, 2018).

Ahmed Hammam, an assistant lecturer at Helwan University, Egypt, whose research is focused on applying Design Thinking methodologies to the study of sciences among primary school students, is a firm believer that using Design Thinking may revolutionize education in the region. Hammam argues for the potential of Design Thinking to attract young students to pursue the study of natural sciences, which remains limited in Egypt (Hammam, 2018).

In Saudi Arabia, King Fahd University of Petroleum and Minerals introduced a series of workshops which were designed to introduce students to Design Thinking. The goal of the workshops was to help guide learning and to develop the students' creative thinking skills, in addition to bridging the gap between theory and practical applications. The workshops also targeted university staff, with the goal of enhancing their professional and soft skills (Ministry of Education News, 2019).

Other initiatives in the education sector were conducted by the National Center for Assessment in Saudi Arabia (Ministry of Education News, 2019) where a series of workshops targeting female school superintendents, supervisors and teachers were delivered. The workshops primarily aimed at stimulating the application of Design Thinking in education. The female educators were introduced to Design Thinking as a concept and were then guided in generating ideas as to how to incorporate it in school curricula.

At Zayed University in the UAE, a group of researchers looked into how to co-design an immersive, transformative, and sustainable education system, whereby students are placed on a change maker pathway for social innovation (Chung-Shin et al., 2018). The research group leveraged community initiatives and institutional support from Zayed University to build a platform for social innovation named INNOCO (Innovation and Co-design). INNOCO was designed to actively support young people interested in expanding their capacity as change makers. This model strives to challenge the existing linear educational approaches and builds on the need for a paradigm shift in education in the region. Their research design utilized a human-centered and evidence-informed approach called "ME = WE" that resonates with the Panarchy Theory in understanding the systemic and symbiotic relationships between self (ME) and society (WE). This framework concentrates on "action and reflection" contributing to social change that one can affect within an individual, community and systemic level. In a three-year implementation period (2015–2018) the team was able to develop a research framework that was successfully implemented and had already been modeled for a partner project at Zayed University. The resulting project proved useful in facilitating a youth engagement program with participants in the UAE and Nepal; and has chronicled the change maker journeys of program participants through quantitative and qualitative narratives. Through these iterative processes, the youth explored ways in which they could connect, collaborate and contribute to their larger communities.

Further efforts to experiment with Design Thinking and education in the Arab world included a paper that was presented at the CEID conference in London in 2018. The paper examined the current state of higher education in the Arab world, and explored the challenges and obstacles that both students and higher education institutions face. It offered a thorough exploration of the potential of digital learning in overcoming some of these challenges, and proposed a Design Thinking framework as a potential pathway to be considered when redesigning the current learning experiences in Arab universities. The paper concludes by providing recommendations for using a Design Thinking framework to support and facilitate the transformation of higher education in the Arab world. The recommendations are: (1) Promote understanding of the Design Thinking methodology and its implications on education among students and educators (through lectures and workshops); (2) Engage all stakeholders to understand their needs and involve them as co-creators in the process; (3) Design pilot interventions at Arab higher education institutions to test the framework and iteratively develop the learning experiences; and (4) Share pilot outcomes and lessons learned to inform further experimentations (Traifeh & Meinel, 2018).

Several Arabic educational MOOCs (Massive Open Online Courses) on Design Thinking also started to appear in the past few years. The first free MOOC was released in 2017 by King Khaled University in Saudi Arabia, and hosted on the university's MOOC platform: kkux.org. The course was delivered by two faculty members, Fahad Alahmari and Abdullah Alwalidi, who attended Design Thinking training at Stanford University and decided to introduce the methodology to Arab students through a MOOC that includes a series of recorded videos, followed by quizzes and discussions on the course forum. The course was delivered over nine weeks and covered the following topics: (1) What is Design Thinking? (2) Exploring the "needs" space, (3) Exploring the "solution" space, and (4) Making difference in your projects. The same course was offered again in 2018 (KKUx, 2018).

Another MOOC provider called Rwaq,[3] in partnership with Monsha'at,[4] offered a free Arabic course on Design Thinking in 2019 which was delivered by Nisreen Alshami. This course follows the self-paced study approach which makes the course open and available to anyone who is interested in joining at any time. It includes four modules that present pre-recorded videos followed by theoretical articles and a quiz after each module. The topics covered are: (1) Concept and

[3]Rwaq is one of the earliest Arabic language educational MOOC platforms that was established in the region. It was launched in 2013 by two Saudi businessmen, Fouad Al-Farhan and Sami El Hassine.

[4]Monsha'at is the Saudi's General Authority for Small and Medium Enterprises, and was established in 2016.

importance of Design Thinking, (2) Design Thinking phases-part 1: "Understanding," which consists of "Empathy" and "Define," (3) Design Thinking phases-part 2: "Discovery," which consists of "Ideation" and "prototyping," (4) Design Thinking phases-part 3: "Embodiment," which consists of "Testing" and "Implementation." The course discussion forum is not very active, and the majority of the questions or requests posted by the MOOC participants are not answered. Despite the low number of posts ($n = 20$), it is worth mentioning that the majority of participants were asked about how to obtain a certificate of attendance ($n = 16$). Four participants complained about the lack of case studies and resources, the way topics were explained, the disconnection between the topics presented in videos and the associated articles, and finally the content of the MOOC: "This content is almost a literal translation from English books. It is not designed for Arabs, and there is a mix of two translated terms of Design Thinking which are used alternatively, while their meanings are different and may completely change the context" (Rwaq, Design Thinking MOOC discussions, 2019). However, there is another space that was dedicated for discussions and questions called "the wall," where 16 notes and questions were posted. Seven participants expressed their satisfaction about the content, and the majority of other participants asked about receiving a certificate of attendance (Rwaq, Design Thinking MOOC wall posts, 2019/2020).[5]

5.2 Design Thinking in the Development Sector

One of the region's few institutions that focuses on design as a multidisciplinary tool for social development and research is the MENA Design Research Center, a non-profit organization founded in 2011 and based in Lebanon. The core of the center's early publications and events was to highlight the inherent ideological connection between the Arab Spring's grass-root movements calling for social, economic and political change in the region, and Design Thinking's participatory and bottom-up problem-solving strategies. In 2012, the center initiated and organized the first Beirut Design Week, which marked the beginning of design weeks in the Middle East & North Africa. It later extended its work to the Desmeem initiative, named for a combination of "design" and "tasmeem" ("design" in Arabic). The initiative creates 10 teams of designers, each team is composed of 3 Lebanese designers and 1 European designer, partner with local NGOs for three months to think up creative solutions to issues that organizations face (MENA Design Research Center 2019).

Recognizing the potential of applying Design Thinking to development, researchers Saad and Shoushanian (2013) studied the case of Egypt as a successful example of applying the model of Design Thinking to conventional design practices, with the intention of improving the production of material objects. The study conducted a thorough tracing of the case of Egypt and developed a conceptual

[5]We could not verify the total number of participants of both MOOCS offered by KKU and Rwaq.

framework linking the factors involved in addressing the role of design in socio-cultural practice (Saad & Shoushanian, 2013).

Major Design Thinking applications in the international development sector were adopted by the United Nations Development Program (UNDP) as one of its main focus areas in the field of innovation for development among Arab states. UNDP identified human-centered design as "a creative approach to problem-solving that starts with the needs of the user, emphasizes the importance of diverse perspectives and encourages solution-seeking among multiple actors" (UNDP, 2018). It applied the approach as a way to respond to the growing need for designing sustainable products, services or experiences in the Arab states. Focusing on applying Design Thinking among youth between 15 and 29 years old, the UNDP launched its regional Youth Leadership Program (YLP) in 2015. YLP aims to build a generation of young leaders, innovators, and change makers in the Arab region. The program applies innovation methodologies, most notably Design Thinking, which is designed to help youths in developing effective and sustainable solutions to address development challenges.

In its first year, YLP was held in Amman, Jordan, and brought together 40 youths from 18 Arab countries to foster their creativity and advance their leadership skills to help them improve their communities. In its second year, under the theme *Innovation for Sustainable Development*, YLP2 supported more than 700 young people through national activities.

Accelerating Innovation for Sustainable Development was the theme of YLP3, which was held in Egypt and supported more than 2000 youths through national activities. With the theme *Innovating for Sustainable Impact*, YPL4 targeted 14 countries across the Arab region[6] and engaged 5000 youths in national activities across the region, and partnered with national organizations that work toward youth empowerment and fulfillment of the Sustainable Development Goals (SDGs). YLP5, with the theme *Explore, Experiment, Expand*, is designed to build on lessons learned during the previous four years and increase its reach by exploring innovative approaches to addressing sustainable development challenges, experimenting with potential solutions, and expanding youth knowledge and networks (UNDP Arab States, 2019).

Another initiative taken by the UNDP was done in partnership with the European Union Commission, promoting Design Thinking and lean startup methodologies among Iraqi Youth in preparation for their participation in the Innovation for Development Forum. Design Thinking was introduced as a methodology to help young people design their prototypes and pilot projects to present in the Innovation for Development Forum (Hilal, 2016).

Further approaches to applying Design Thinking to the field of development included the concept of Social iDesign proposed by El Aidi (2017). Social iDesign is a concept that aims at improving the quality of people's life in accordance with their cultural values by stimulating their social intelligence. The study's main

[6]The partner countries include but are not limited to: Egypt, Syria. Iraq, Bahrain, Yemen, Sudan, Jordan, Palestine, Algeria, Morocco, Tunisia, Lebanon and Saudi Arabia.

premise is that designers seek to generate meaningful and sustainable solutions by co-designing conditions and situations for more empathetic interaction between people and their society. Having this as its main thesis, the research on Social iDesign is "built upon integrating three forms of the cognitive process of design: learning, doing, and reflecting" (El Aidi, 2017). Therefore, this model offers a holistic approach to improving the quality of life in Egypt by thinking in a systematic manner and developing the three pillars of design simultaneously, rather than only focusing on one design component at a time. These pillars are: (1) Design education, (2) Design practice, and (3) Design research.

Similarly, research into applying Design Thinking principles to local development in Algeria showed that there is a lot of merit in focusing on citizens as the main users of local development policies. The research also highlighted the importance of involving citizens in the implementation of local development programs as a means of enhancing ownership and enabling sustainable policies (Haroush & Maaroufi, 2017).

In 2018, an academic scholar conducted a research study to identify the impact of the Design Management Processes (DMP) on decision making (DM) through Design Thinking (Ouda et al., 2019). The study focused on decision makers from 78 local NGOs in the Gaza strip. The research findings showed no direct relationship between the design management processes and decision making. Nevertheless, in-depth interviews with senior leadership revealed that Design Thinking has the potential to mediate the relationship between the design management processes and decision making.

In the area of social science studies, researchers Osman and Dahlan (2019) adapted the Design Thinking and System Thinking approach to solving the Eritrean refugee's problems in Sudan. Understanding the needs of the people, business modeling tools such as Business Model Canvas (BMC) and Value Proposition Canvas (VPC) were built and developed. The proposed solutions include educational skills such as technical skills programs, study materials, literacy programs, and mentoring programs. The researchers claim that the main contribution of their work is empowering the refugees with knowledge, values and skills to live productive and independent life (Osman & Dahlan, 2019).

In 2019, Sabr, an Arabic business design company based in Turkey, launched a book titled "Design Thinking for Social Innovation" in which Ghaiath Howari and Kinda AlMemar introduce Design Thinking as an innovation methodology. The authors present steps on applying Design Thinking to solve social problems, and they discuss case studies (Howari & AlMemar, 2019). This book is believed to be the first published book in Arabic that focuses on Design Thinking.

5.3 Design Thinking in the Private Sector

During the past few years, many of the private design firms established in the region changed their business strategy and shifted from offering only visual design services to focusing more on innovation methodologies and adopting Design Thinking and human-centered approaches in their work. HUED is, for example, one design and innovation consultancy that was established in 2013 in Saudi Arabia by a small service design team. HUED expanded and developed their work methods based on principles of Design Thinking such as empathy, building user insights, dealing with ambiguity and uncertainty, and designing and executing solutions based on a human-centered approach. HUED is now considered Saudi's first and largest innovation & design consultancy (HUED, 2019). They offer services to public as well as private companies and organizations in the country, and they also work on promoting the culture of Design Thinking by organizing and running different trainings and annual service design and customer experience events.

Another example is ERGO, one of the region's first human-centered innovation and strategy design firms. Established in 2016 in Cairo, Egypt, the firm is committed to advancing human experience through design and innovation. With a deep belief that design must be rooted in people, culture and human values, ERGO believes in the power of Design Thinking in bringing innovation to one of the most culturally sensitive regions in the world. By restoring focus to human-centricity, and capturing key market and behavioral insights, ERGO conceptualizes and delivers breakthrough innovations to help the Arab market create positive deviance, and achieve sustainable growth. The company offers services that range from delivering insight-driven products, services and experiences, to designing new impact-driven ventures, to building innovation capacity in corporations, NGOs, enterprises and governments to unleash their creative potential and become better innovators (ERGO, 2016).

Beyond private design firms, Design Thinking and user-centered approaches have been applied in other types of private entities such as private medical clinics. In Egypt, for example, recognizing the issues associated with physician-oriented systems such as waiting times and appointment, researchers applied design features and functions to improve the quality of communication between patients and clinic staff (Heshmat et al., 2017). The researchers aimed at decreasing waiting times and developing schedules for patients and staff. Using an Egyptian private clinic as a case study, the team designed a patient-oriented mobile application called "Your Clinic" in order to support online booking and consultation. The application combines 6 main features: (1) Creating accounts; (2) Clinic search; (3) Patient booking; (4) Nurse scheduling; (5) Doctor scheduling; and (6) Online consultation.

In 2017, the Innovation Academy at King Abdullah University for Science and Technology (KAUST) launched the REVelate Corporate Innovation Program (CIP, 2017): an intense three-day, team-based program designed to support new innovation by organizations in the Kingdom of Saudi Arabia. The program is delivered twice a year at KAUST's premises and targets a variety of organizations medium-to-large companies, non-profits and government agencies. As part of REVelate, CIP is based on Design Thinking, change management and entrepreneurship. Additionally, CIP combines elements of innovation, new internal venture creation, team development, business model development and executive education.

Private sector applications of Design Thinking continued to thrive. With a focus on the human resources departments and employees' experience, Ahmad Al-Ghamdi, VP of HR in a big private company (name of company kept confidential), argues that this approach can boost morale and preemptively deal with human resource dissatisfactions and deal with potential sources of demotivation. He claims that "Design Thinking can be used in all human resource processes, starting from attracting talent, identifying potential employees, employing them in specific roles, creating opportunities for their development and transferring them to new roles, and of course, retaining key talent in the facility, developing and turning them into large and important assets for the organization." (Al-Ghamdi, 2019).

Other practical applications of Design Thinking included the creation and implementation of an identity re-brand and successive marketing campaign for the British Football Academy, in Kuwait. Design Thinking was applied in the course of a "Value Management Branding Workshop," which involved key stakeholders in the co-design of a new brand for the British Football Academy. The workshop had two main outcomes: (1) A clear vision of British Football Academy's brand position toward customer experience, (2) A direct indication of how the brand needed to develop to fulfill the objective of increasing its market share. "This cyclic approach (of Design Thinking) utilized visualization as a pragmatic tool, aiding the development of marketing strategies and resulting in innovative solutions, with the launch of the new brand identity and marketing campaign receiving an overwhelmingly positive response from all parties involved" (Winstanley, 2019).

5.4 *Design Thinking and Entrepreneurship*

Between 2014 and 2015, Design Thinking became recognized as a viable tool to aid businesses in the Arab World. Mainly focused on small to medium-sized enterprises (SMEs), Design Thinking started to create a buzz among hackathon arenas, business incubators and entrepreneurship hubs. One early adopter of Design Thinking as a method to further small businesses in North Africa was the Egyptian non-governmental organization, Nahdet El Mahrousa. Through its social incubator, Nahdet El Mahrousa was among the first in the region to establish systematic Design Thinking workshops for social innovation. The workshops were conducted all over Egypt and aimed at preparing young entrepreneurs to start their own businesses. Similar initiatives that utilized Design Thinking in social innovation include BINA, a non-profit social development and capacity building organization established in 2014. Based in Turkey but directed at Syrian refugees, Bina's work focused on guiding Syrians to develop their own community initiatives to address their most pressing issues (BINA, 2017).

Another example of the adaptation of Design Thinking in entrepreneurship is Wamda, an entrepreneurship empowering platform for entrepreneurs in the Middle East and North Africa. Wamda's team, based in Dubai, aim to offer support that entrepreneurs need via three primary platforms: (1) a media site that is the leading source for startup and entrepreneurship news in the region, (2) a fund, which is the largest early stage investment fund in the region, and (3) a programs arm, that includes "Mix N Mentor" event series (wamda.com). Among the articles they feature and the workshops they host, the awareness building of Design Thinking has been a hot topic: from employing Design Thinking in workspaces, applying it in designing a great UX experience, using the methodology to guide students into solving problems concerning education and transportation (Menon, 2016; Wamda events, 2017; Rahal, 2017).

Similar steps were taken in Saudi Arabia in 2017, whereby the Taqadam Startup Accelerator collaborated with King Abdullah University for Science and Technology (KAUST) to produce a series of articles outlining the Design Thinking process to foster agile management for companies (Taqadam Startup Accelerator, 2017). The series also draws attention to the corresponding training program offered by Taqadam and related courses offered by KAUST. The series consists of five articles; each article depicts a phase of the Design Thinking process. The first, titled "Why Your Company Needs Design Thinking?" introduces Design Thinking, its attributes and potential benefits for companies. The second, titled "Emphasize: How to Know What Your Customers Really Want?" addresses the question of empathy and how to attain it in the context of product development and customer relations. The article briefly highlights the importance of embedding products in user needs and provides a snippet of the main methods used to gain empathy, namely: observation and body-storming (roleplaying). The third article: "Can You Define Your Startup in One Sentence?" tackles the "Define" phase by walking the reader through defining the right problem to solve, identifying the point of view (POV)—in accordance with the Stanford d.school template—and developing a problem statement. The fourth article: "What Makes a Good Startup Idea?" covers the "Ideation" phase, which the authors define as "the transition between identifying the problem and creating the solution." In doing so, the article references the Stanford d.school's Design Thinking Process Guide by outlining three main aspects. It starts with the "How Might We?" questions, designed to kick-start different brainstorming ideas. It follows by guiding the reader through the brainstorming process, where the importance of five main rules are highlighted: (1) Stay focused; (2) Defer judgment and criticism; (3) Encourage free thinking; (4) Quantity over quality and 5) Be visual. The article ends by helping the reader to go through the idea selection process. Here, they put forth a sequence of grouping, clustering and voting as the way to go. The fifth and final article: "Why and How to Use Prototypes?" explains that prototyping is an essential step in bringing an idea to life as it can prove viability and test the real-world impact of a product or idea. It also highlights that prototypes help designers see the major errors of their product before the going-to-market phase. The authors then follow by drawing special attention to the distinction between low-fidelity prototypes and high-fidelity prototypes as well as the benefits and approaches of each.

Other voices on the Internet and social media outlets have continued to emerge. Prominent examples include blog posts and magazine articles addressing the relationship between Design Thinking, business success and entrepreneurship.

These include the "Entrepreneur AlArabiya" online magazine which was established in 2014 to become one of the trustworthy sources of information for Arab entrepreneurs. Among the different topics covered by the magazine, some articles talked about Design Thinking as an innovation tool that entrepreneurs should use to help them define and solve the challenges around them. In one of the articles, Hanan Sulaiman reports on a workshop conducted by a Design Thinking expert, Eman Abo El Atta. The workshop was part of a series titled "Maker Faire," which is an annual event that takes place in Cairo and showcases invention, creativity and resourcefulness from all over the Arab region. It was modeled on the US version which currently spans over 40 major cities in the world. The workshop report covers the five-step Stanford Design Thinking process: "Empathize," "Define," "Ideate," "Prototype," and "Test" (Sulaiman, 2016).

5.5 Design Thinking in the Public Sector

Over time, there has been a noticeable increase in the uptake of Design Thinking by the public sector in the Arab World. Illustrating the incorporation of Design Thinking in the public sector is the UAE Government Leadership Program, which considered Design Thinking methods as integral to the development of UAE's future leaders (UAE Youth Leadership Program, 2017). The UAE Youth Program was one of the programs launched by the UAE Government Leaders Program in June of 2017 after the development of the Innovative Leaders Program, which was launched in 2014. The program is implemented in collaboration with the Youth Office and it targets UAE youth from federal and local government, as well as the private sector. The eight-month program is designed to train ambitious and driven young adults between the ages of 21–35. Consisting of five modules, the Design Thinking methodology underpins every facet of the program in order to promote innovation, creative thinking, and continuous development and learning. The program ultimately aims at realigning and redesigning work practices in the UAE.

In Egypt, the Ministry of Youth and Sports, together with USAID'S Youth Leadership Program (YLP) developed annual youth training camps to support youth initiatives across the country (Bikir, 2019). The program is based on the Design Thinking methodology. The young people, first trained in Design Thinking phases, are expected to translate their acquired knowledge into youth initiatives to be implemented across the country. In this sense, the methodology is perceived as a gateway to promote youth social activism.

Another major stride toward the integration of Design Thinking methods in the public sector was the establishment of the Mohamed Bin Rashid Center for Government Innovation in the UAE. The center presents Design Thinking as part of their implementation toolbox, which highlights its merits in generating creative solutions for "wicked" problems in the public sector. Moreover, the UAE government launched an annual event titled the "Dubai International Project Management Forum" (2019), which offers a series of master classes on diverse project management topics. At the forefront is the utilization of Design Thinking in fostering flexible project management. An example of this is the UAE Ministry of Finance which provides its employees with an intensive training in Design Thinking methodologies with the aim of improving their problem-solving and analytical skills.

Public sector efforts at advancing Design Thinking in the health care sector manifested itself in Saudi Arabia, where the Ministry of Health collaborated with King Abdullah Medical City to launch a series of Design Thinking workshops aiming at training medical professionals in applying Design Thinking methodologies to improve the patients' experience and enhance safety procedures in health services (Saudi Press Agency, 2015).

6 Discussion and Conclusions

In this chapter, we reported on a preliminary exploration of the current state of Design Thinking in the Arab world through a systematic literature review based on the guidelines of the PRISMA framework. Our review covers the entire year of 2010 (the first time the Arabic translation of Design Thinking "التفكير التصميمي" was coined) and extends up to, and includes, May 2019. We acknowledge the following limitations in this study:

- The English keywords identified to conduct the online search focused on the Arabic-speaking countries were limited to two keywords "Arab" and "Arabic." The Arab world however consists of 22 countries, and hence, a country by country filtration may yield more articles and expand our understanding of the state of Design Thinking in the region.
- A few Arab countries including Morocco, Tunisia and Algeria mainly communicate in French. This study did not cover articles and events conducted in French in the Arab region.
- Most of the data collected are news reports, blogs and online articles for which the web links may change over time. Therefore, we may have missed some of the events that happened earlier when Design Thinking started to emerge in the region.

This study covers one part of a larger-scale study. We aim to compare its findings with the results of the other parts to further enhance our understanding of the history of Design Thinking in the Arab world and its application across different sectors. Nevertheless, one of the major findings of this study is tracking down the first translation of the term "Design Thinking" into Arabic as "التفكير التصميمي." The term appeared in a translation of Tim Brown's Ted Talk: Designers—think big, in 2010. This translation for Design Thinking was coined by Chafic Jaber, a research engineer at Telecom ParisTech, who voluntarily translated the video. The study also shows that there is an ongoing debate about the most accurate Arabic translation of

Design Thinking. In addition to التفكير التصميمي, there are three more terms that researchers, designers or authors in the region are using: الفكر التصميمي, تصميم التفكير and التصميم المتمحور حول الإنسان. However, based on the data collected, we can conclude that "التفكير التصميمي," the term that was coined in 2010, is the most widely adopted Arabic translation of "Design Thinking." The data also shows that Design Thinking is still in its early stages of adoption in the region, but it appears to have started to grow in popularity in the past few years, especially in the education, development and entrepreneurship sectors. The Arab countries that seem to have the highest adoption of the methodology are Saudi Arabia, Egypt and UAE.

References

Al-Dousary, M., & Al-Robayaa, M. (2013). Design thinking. Retrieved from https://www.slideshare.net/mahaalrobayaa/design-thinking-2946209.

Al-Ghamdi, A. (2019, August 19). التفكير التصميمي والموارد البشرية. Retrieved from صحيفة مال الاقتصادية website https://maaal.com/archives/20190819/127116.

Al Shaddy, W. (2011). Ma'reb Okhra Blog, hashtag: Design Ask. Retrieved from http://www.w-dd.net/wp/?p=3188#more-3188.

Aldakheel, B. (2015, Sptermber 29). Oktob blog. Retrieved https://oktob.io/posts/1642.

Aldawood, E. (2012, Septmeber, 4). Xschool Youtube Channel. دقائق مع مفهوم (7): هل الأفكار الناجحة عشوائية؟ https://www.youtube.com/watch?v=zAv3UYBrA98.

BINA. (2017). *About us*. Retrieved from https://www.bina-org.com/en/about-us/.

Balawi, S., Hitt, G. W., Khalaf, K., Hirsch, B., & Powell, L. (2013, August). Cultivating design-thinking in freshmen: The evolution of the KU Freshman design course. In: *Proceedings of 2013 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)* (pp. 350–353). IEEE.

Bikir, R. 2019. The ministry of youth and sports completes the activities of the third camp to support 1100 in Fayoum, Sada Elbalad, August 22. https://www.elbalad.news/3951228).

Brown, T. (2009). Designers Think Big, TEDGloobal 2009. Retrieved from https://www.ted.com/talks/tim_brown_designers_think_big/transcript?language=ar.

Chung-Shin, Y., Renaux, J., Chikermane, V., & Rajani, J. J. (2018). Co-designing a social innovation model for changemakers. In: *Proceedings of RSD7, Relating Systems Thinking and Design* (Vol. 7). October 23–26, 2018, Turin, Italy.

Design Thinking (التفكير التصميمي) MOOC. 2018. KKUx, Saudi Arabia. https://kkux.org/courses/course-v1:KKU+DesignThinking001+2018_sem2/info.

ERGO Homepage. (2016). Retrieved from https://www.ergo-x.design/.

El Aidi, D. (2017, November). The Potential of Design Thinking in Sustaining Development Efforts in Egypt. In the 1st International Conference on Towards a Better Quality of Life.

HUED. (2019). Design and Innovation. https://www.livehued.com/about-hued/.

Hammam, A. (2018). فاعلية وحدة STEM لتنمية التفكير التصميمي في مادة العلوم لدي تلاميذ المدارس الرسمية للغات مقترحة في ضوء مدخل (Master's Dissertation). Retrieved from https://www.academia.edu/38781879/.

Haroush, N., & Maaroufi, A. (2017). Academic interest in the process of local development in Algeria, an analytical study of the process through some university theses using the method of design thinking. *Journal of Ijtihad for Legal and Economic Studies/University Center of Tamingst-Algeria, 12,* 117–138.

Heshmat, M., Mostafa, N., & Park, J. (2017). Towards patient-oriented design: A case of the Egyptian private outpatient clinics. In: *Proceedings of the 50th Hawaii International Conference on System Sciences* (pp. 3839–3849). https://doi.org/10.24251/hicss.2017.464.

Hilal, N. (2016, October 9). EU/UNDP Innovation for Development Initiative positions Iraq on the Social Good Summit global map with 6 youth gatherings. Retrieved from UNDP in Iraq website: https://www.iq.undp.org/content/iraq/en/home/presscenter/pressreleases/2016/10/09/eu-undp-innovation-for-development-initiative-positions-iraq-on-the-social-good-summit-global-map-with-6-youth-gatherings.html.

Howari, G., & AlMemar, K., (2019). Design thinking for social innovation, book. https://sabr-sp.com/ar/home/details/design-thinking-in-social-innovation.

Ibrahim, W., (2018). Social design in the Arab world: How undergraduate design education can lead the way to a better future, Masters thesis, Aalto University. http://urn.fi/URN:NBN:fi:aalto-201812126400.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C et al. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies That evaluate health care interventions: explanation and elaboration. PLOS Medicine 6(7): e1000100. https://doi.org/10.1371/journal.pmed.1000100

MENA Design Research Center. (2019). https://www.menadrc.org/.

Meinel, C., Weinberg, U., & Krohn, T. (Eds.). (2015). Design thinking live: Wie man Ideen entwickelt und Probleme löst. Murmann Publishers GmbH.

Menon, A. (2016, February 1). What I know about design thinking: Johannes Comeau Milke. Retrieved from Wamda.com website: https://www.wamda.com/2016/02/what-i-know-about-design-thinking.

Ministry of education, Saudi Arabia. (2019, May 17). The implementation of design thinking in education training. Retrieved November 2019, from edu.moe.gov.sa website https://edu.moe.gov.sa/Sharqia/MediaCenter/News/Pages/NMA1000143.aspx.

Ministry of Education News, National Center for Assessment, Saudi Arabia. (2019, March 17). Applying design thinking in education—A workshop for female teachers in The Education Sector in Riyadh. Retrieved from spa.gov.sa website https://www.spa.gov.sa/1899834.

Osman, G. M., & Dahlan, A. R. B. A. (2019). Empowering Eritrean refugees in Sudan through quality education for sustainable development. *International Journal of Social Science and Humanities Research, 7*(2), 476–483.

Ouda, R. O., Ouda, M., Abed, M. N., Abdalmenem, S. A., Abu-Naser, S. S., & Al Shobaki, M. J. (2019). Design thinking and its impact on decision making in NGOs," PhD Dissertation, Al Azhar University, Department of Information Technology. http://dstore.alazhar.edu.ps/xmlui/handle/123456789/472

Plattner, H., Meinel, C., & Weinberg, U. (2009). Design thinking. Innovation lernen. Ideenwelten öffnen. München: Mi-Wirtschaftsbuch.

REVelate Corporate Innovation | KAUST Innovation & Economic Development. Kaust.Edu.Sa, November 2017. innovation.kaust.edu.sa/industry/explore-corporate-innovation-training/revelate/. Accessed November 19, 2020.

Rahal, M. (2017). Wamda hosts student workshop around education and mobility in UAE. Retrieved from https://www.wamda.com/2017/02/wamda-hosts-students-workshop-education-transport-uae.

Rwaq. (2019). Design thinking MOOC discussions. Retrieved from https://www.rwaq.org/courses/designth/discussions.

Rwaq (2019/2020). Design thinking MOOC wall posts. Retrieved from https://www.rwaq.org/courses/designth/wallposts.

Saad, Q., & Shoushanian, C. (2013). Exploring Models of Design Thinking in Egypt. *Making Futures: Interface Between Craft Knowledge and Design: New Opportunities for Social Innovation and Sustainable Practice, 3,* 530–540.

Saudi Press Agency, Saudi Arabia, The conclusion of the summer research program at king abdullah medical city. (2015, Feb. 21). Retrieved October 2020, from spa.gov.sa website: https://www.spa.gov.sa/1377150

Schmiedgen, J., Rhinow, H., & Köppen, E. (2016). *Parts without a whole? The current state of Design Thinking practice in organizations* (Vol. 97). Universitätsverlag Potsdam.

Shouman Foundation. (2019). محاضرة "التفكير التصميمي" لخبيرة التفكير التصميمي تانيا عمايسي "Design Thinking Lecture by DT Expert Tania Amaisi" [YouTube Video]. Retrieved from https://www.youtube.com/watch?v=DRZFH4sNc4A.

Sulaiman, H. (2016). What is design thinking and what do you know about it? Entrepreneur AlArabiya Magazine. Retrieved August 2019, from https://entrepreneuralarabiya.com/2016/04/25/2784/ ماذا-تعرف-عن-التفكير-التصميمي؟/#.X7TlFGhKhaQ.

TAQADAM Startup Accelerator. (2017). Why your company needs design thinking? Retrieved from Kaust.Edu.Sa, 2017, innovation.kaust.edu.sa/why-your-company-needs-design-thinking/.

Taha, S. (2019, March 27). Design Thinking—Salah Taha. Salahtaha.Com. Retrieved from https://salahtaha.com/.

Traifeh, H., & Meinel, C. (2018). Design thinking: a proposed framework for transforming higher education in the Arab World. In: *CEID Annual Conference*.

UAE Youth Leadership Programme. Retrieved from Www.Uaeglp.Gov.Ae, 2017. www.uaeglp.gov.ae/en/areas-of-focus/government-leaders-programme/program?id=2eb44b51-f05b-46df-b450-a72fb7ab641d. Accessed November 19, 2020

UNDP, Innovation for Arab States. (2018). Human centered design. http://www.inno4arabstates.com/Methodology/view/10.

UNDP. (2018). Human-Centered Design | UNDP in the Arab States. Retrieved from www.arabstates.undp.org/content/rbas/en/home/Innovation/human-centered-design-.html.

Unknown, Arabstarts Blog. (2012, June 2). التفكير التصميمي: تعريفه وعلاقته بمجالات الاعمال -١ "Design thinking: Definition and relation to business. Retrieved August 2019, from http://arabstarts.blogspot.com/2012/06/1.html.

Wamda events. (2017). UX Workshop by Willy Lai. Retrieved from https://www.wamda.com/2017/09/ux-workshop-willy-lai.

Winstanley, L. (2019). The British Football Academy: A design thinking approach to branding strategy. *SAGE Research Methods Cases in Business and Management*. https://doi.org/10.4135/9781526465153.

Decoding Nonverbal Online Actions: How They Are Used and Interpreted

So Yeon Park, Mark E. Whiting, and Michael Shanks

Abstract Nonverbal actions form an integral part of the interactive experience between users on online platforms. While convenient, these nonverbal online actions, which often take the form of clicking a button to like or share content with others, sacrifice contextual information compared with their in-person and verbal counterparts as they are simplified into an icon or content featured in one's feed. We show that the simplicity of online nonverbal actions and their indicators may come at a cost to users. Through two surveys, we found that there was no single dominant reason for each nonverbal action; rather, users were varied in what they felt were the reasons for taking particular online actions. Users also felt more positive about some reasons than others, and how one felt about a nonverbal action was found to correlate with their affect for the reason thought to be behind that action. Finally, we found that not being aware of why an action was taken could lead to negative effect for an action. Unexplained nonverbal actions in online platforms evoke more negative affect than those same actions when their reason is known. Our findings suggest that surfacing rationales behind nonverbal actions could be one way to mitigate the unintended uncertainty and misunderstandings inherent in how indicators of nonverbal online actions are currently designed. We present this implication amongst others to consider in designing social platforms for user-user interactions online.

S. Y. Park (✉) · M. Shanks
Stanford Center for Design Research, Stanford University, 424 Panama Mall, Stanford 94305-2232, CA, USA
e-mail: syjpark@stanford.edu

M. Shanks
e-mail: mshanks@stanford.edu

M. E. Whiting
University of Pennsylvania, Pennsylvania, USA
e-mail: markew@seas.upenn.edu

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_5

1 Introduction

Communication is mediated by technology more than ever before, and critically, subtle signals valuable to in-person interactions may go unspoken and not be noticed. In social media platforms, much of our communication and interactions are arguably nonverbal, taking the form of a button-click (e.g., Like, Retweet), inaction (e.g., liking someone else's post but not mine), and even when actions are taken (i.e., chronemics). Unlike actions taken offline, nonverbal online actions are made visible through the platform's design. Platform architecture and feature designs dictate how these nonverbal actions are translated and depicted to the users and their interactants. Hence, nonverbal online actions on social media platforms may or may not result in a change that is visible to the actor, or the observer, or both the actor and observer. Misunderstandings and interpersonal complications arise as subtle cues around these nonverbal signals are increasingly missed online and the context in which these actions are taken are not transparent to those receiving them. Many of these actions also adopt meanings beyond their original design—e.g., in different situations, the Like button is used to indicate appreciation, receipt, or empathy (Hayes et al., 2016; Lee et al., 2016; Sumner, 2018)—further obfuscating the meaning or intention behind the nonverbal action.

Some academics posit that misunderstandings that occur online are not as serious as those that occur face-to-face (Edwards et al., 2017). However, misunderstandings online can result in life-altering consequences too—in several cases, people have been imprisoned over misunderstood online actions.[1] Nonverbal online actions may also signal partiality. BBC has recently gone so far as to ban all their staff from "virtue signaling", which can take the form of liking or following certain accounts, as these may signal to their readers sharing a personal opinion on an issue and thereby "undercut an otherwise impartial post".[2] Virtue signaling can also occur amongst non-journalist users by simply retweeting or sharing a post. While the action of frequently posting about a subject may also connote interest and perhaps knowledgeability in the matter, one nationwide survey found that "individuals most active in contributing to social media were actually propagating inaccurate information" (Groshek & Bronda, 2016). To mitigate fake news propagation, Facebook and Twitter had made changes prior to the 2020 US election—Facebook forbade new political advertisements the week before the election, and Twitter labeled posts that were deemed to be misleading, disputed, and unverified (Elkind, 2020; Pennycook & Rand, 2020; Roth & Pickles, 2020).

---

[1] https://evangelicalfocus.com/world/2559/christian-in-pakistan-sentenced-to-life-imprisonment-for-blasphemy, https://www.smh.com.au/national/sydneyman-unlawfully-imprisoned-in-egyptian-prison-since-january-20200920-p55xff.html, Accessed November 16, 2020.

[2] https://www.pressgazette.co.uk/bbc-social-media-guidelines-ban-virtue-signalling-criticism-of-colleagues-and-breaking-stories-on-your-ownaccounts/, Accessed November 5, 2020.

Both approaches taken by Facebook and Twitter are intended to alter social dynamics in the online world through the design of features with nonverbal components. This highlights that the lack of such approaches prior was also a design decision. In other words, for every nonverbal online action, their indicators are designed by the platforms. These indicators affect the users' experience directly, including how users perceive these nonverbal actions. In our investigation, we will refer to these as *online action indicators*, specifically, *indicators that evidence and/or depict a nonverbal action taken by a user with or without their intention to carry it out on, and as designed by, an online platform.* These online action indicators can manifest themselves in a multitude of forms, according to the philosophy and design of the platforms, and include the use of Like buttons, the inclusion (or lack) of comments in retweets, and social presence indicators in collaboration platforms, but, notably, not the content itself. Comment text, messages, and emoji fall outside of this interpretation of online action indicators.

We believe that systematically understanding the impact of these nonverbal online actions and the impact of their indicators (i.e., online action indicators) on users is critical to understanding how platforms ought to design for their users. We investigate the impact of nonverbal online actions through two research questions:

Research Question 1 (RQ1). *How do users interpret nonverbal online action indicators?*

Research Question 2 (RQ2). *What is the impact of not knowing the reasons for nonverbal online actions?*

We answer these two questions through two studies respectively, and throughout this chapter, we show *why* these questions are relevant and critical. We study three instances of nonverbal online actions in widely-used online social platforms to shed light on what reasons lie behind them and how these reasons can influence the recipients' affect. Through surveys, we found that rationales—the logical basis and reasoning for taking the nonverbal action—matter in how people perceive these nonverbal online actions and that knowing why someone has performed an action can help recipients feel better about it. These perceptions of nonverbal actions suggest that incorporating indicators of rationale could help improve affect when using social platforms. In doing so, we show the need for greater clarity of users' intentions in social platforms. Additionally, other platforms that mediate user-user interactions can also benefit from this insight. Our work highlights the role of design—and the need for design thinking—in our nonverbal communication online. We show how designs impact affect, and we provide implications for how nonverbal actions and their indicators ought to be designed.

2 Related Works

Nonverbal online actions, as considered in our work, are: *actions users (can) take online* (e.g.*, through clicking the Like, Retweet, or turn video off buttons) without any explicit textual communication (*e.g.*, words) that result in a change.* This change may or may not be visible to the actor or both the actor and observer (e.g., count of Likes increasing, video-conferencing person no longer being visible). Here, we provide a contextual backdrop for these *nonverbal online actions* and their *indicators*, the subjects of our investigation, and expected results with regards to our hypotheses.

2.1 Nonverbal Communication

Nonverbal communication has been a topic of extensive research since its conceptual formulation by Ruesch (Ruesch, 1953, 1955), and is defined as “all of the ways in which communication is effected between persons when in each other’s presence, by means other than words” (Kendon, 1986; Kendon et al., 2010). This includes potentially informative behaviors that are not purely linguistic in content (Knapp et al., 2013). Matsumoto et al. extends this to define nonverbal communication as “the transfer and exchange of messages in any and all modalities that do not involve words” (Matsumoto et al., 2012). Nonverbal communication can be approached in several ways: “proxemics (communicative use of space), kinesics (communicative use of the body movements), haptics (communicative role of touch), physical appearance (communicative use of body endowments, e.g., body shape and size, skin and eye color, and body adornments, e.g., clothing, makeup, tattoos), oculesics (communicative eye behavior), chronemics (communicative role of time), objectics (communicative use of artifacts), and vocalics (communicative aspects of voice, such as tone, accent, and loudness)” (Antonijević, 2013).

Since the advent of computer-mediated communication (CMC), research in nonverbal communication has expanded to encompass not only the same corollaries in video-conferencing (Nguyen & Canny, 2005; Okada et al., 1994) but also similar communication in virtual environments (Becker & Mark, 1998; Guye-Vuillème et al., 1999; Yee et al., 2007). With avatar-based communication that can also occur in virtual reality, nonverbal communication can take more embodied forms with multimodal cues (Benford et al., 1995; Moore et al., 2007). In a gaming context, this includes “pings”, which are “alerts that are easy to activate and provide auditory and visual cues for teammates” (Leavitt et al., 2016).

The nuances considered within nonverbal communication have differed within research, with some considering certain types of behavior that is sent and received (see Table 1) as part of nonverbal communication, while others do not. For example, only

Table 1 Matrix of behavior sent with behavior received reproduced from Guerrero et al.’s work (Guerrero et al. 1999)

	Not received	Received inaccurately	Received accurately
Sent with intent	Attempted communication	Miscommunication	Successful communication
Sent without intent	Unattended behavior	Misinterpretation	Accidental communication

behavior sent with intent (e.g., (Motley, 1990)) or behavior that is received by another (e.g., (Andersen, 1991)) were considered as nonverbal communication (Guerrero et al., 1999). Some, including Guerrero et al., argue that nonverbal communication encompasses all—that they are "behaviors (other than words) that are typically sent with intent and/or interpreted as meaningful by receivers" (Guerrero et al., 1999). In our work, while we adopt the expansive concept of nonverbal communication, we focus on how users interpret the indicators that they receive of the action (or inaction).

2.2 Nonverbal Actions

Tied intricately with nonverbal communication are *nonverbal actions*, upon which we center our work. While "pings" and emojis themselves are a form of nonverbal communication (Leavitt et al., 2016; Pohl et al., 2017; Rodrigues et al., 2017), choosing to send them (i.e., taking these actions over others) or performing any other equivalent action—as well as inaction (or absence of action)—are the focus of our investigation. These nonverbal actions possess communicative value and therefore can be seen as communicative acts, "through which [people] get others to understand what they mean" (Clark & Brennan, 1991). Nonverbal online actions include conceptualizations such as paralinguistic digital affordances. Hayes et al. describe paralinguistic digital affordances as being "representative of phatic communication in their design but possibly not in their usage" (Hayes et al., 2016), as indicators online present themselves as an affordance to their users. Yet, our notion of nonverbal online actions extends beyond paralinguistic digital affordances in that they also include the indication of a user having taken the action of retweeting (lexical) in addition to clicking emojis (paralinguistic).

Ekman and Friesen distinguished five types of nonverbal behaviors; one type is "nonverbal actions". These are regulators that "have as their sole function the management of the conversational flow or exchange" (Ekman & Friesen, 1969). As described by Goffman et al., nonverbal actions "may inadvertently distract from the performance" (Ekman & Friesen, 1969) and their "full significance is not appreciated by the individual who contributes them to the interaction" (Goffman, 1978). As with nonverbal communication, nonverbal actions can be taken with intent and/or interpreted as meaningful. As shown in Table 1, only one of the six possibilities of nonverbal communication is successful communication; others show possible misalignments between the sender and the recipient. An example of misalignment between intent and interpretation in nonverbal online actions is when Facebook Likes are over-interpreted as indicating a nuanced social preference. BBC's decision to ban their staff from "virtue signaling" evidences the grave consequences of misalignment in such nonverbal online actions at the corporate level.[2]

Virtue signaling has been found to be one of four motivations for using Likes, Favorites, and Upvotes in particular: literal interpretation, acknowledgement of viewing, social support/grooming, and utilitarian purposes (Hayes et al., 2016).

Literal interpretation means "participants liked Facebook content they actually liked"; acknowledgement of viewing is using Likes to "acknowledge they had seen a post"; social support/grooming is when paralinguistic digital affordances were used as a "form of social support" with an additional aspect of social grooming, such as taking action as part of "relationship maintenance"; and utilitarian purposes included using paralinguistic digital affordances for "keeping a record of particular content" (Hayes et al., 2016). Another study found that out of "365 Likes, 52% were meant to convey content-related thoughts, 23% were meant to communicate relational-based sentiments, and 24% were said to contain both types of meaning" (Sumner, 2018), similar to findings in earlier work (Lee et al., 2016). The applicability of the four usages of paralinguistic digital affordances found from Likes and Upvotes to other nonverbal online actions across platforms have yet to be validated. Furthermore, the extent to which nonverbal online actions are used for various reasons is also open to question. This leads us to ask RQ1—how do users interpret nonverbal online action indicators?

We seek to answer how the indication of these nonverbal online actions affect online platform users. In a similar manner to Likes or Upvotes, we expect that nonverbal online actions at large will be taken for different purposes. Given the false consensus bias that embodies the tendency for layman to "see their own behavioral choices and judgments as relatively common and appropriate to existing circumstances while viewing alternative responses as uncommon, deviant, or inappropriate" (Ross et al., 1977), we expect to observe heterogeneous, or diverse, interpretations from users who take these same actions for heterogeneous reasons.

Hypothesis 1 (H1) *People are heterogeneous in their interpretation of nonverbal actions.*

Hayes et al. also found that "faithful" and "ironic" use cases of paralinguistic digital affordances were "encoded and decoded by senders and receivers" (Hayes et al., 2016). Extending Adaptive Structuration Theory, which acknowledges that users can appropriate the technology faithfully (i.e., as intended by the developers) or unfaithfully (i.e., as not intended by the developers and "out of line with the spirit of the technology") (DeSanctis & Poole, 1994), faithful use cases of paralinguistic digital affordances involved using and understanding the PDA for its face value, whereas ironic (i.e., unfaithful) use cases involved social motives (Hayes et al., 2016). Such ironic appropriations of paralinguistic digital affordances were found to be used frequently with senders' social motives directed "toward the message poster, rather than the content itself" (Hayes et al., 2016).

For our purposes, we will consider these "ironic use cases" as having an *underlying rationale* for taking the nonverbal actions; we consider many of the cases here exemplifying underlying, indirect communication rather than embodying irony. Similarly, we will call nonverbal actions with "faithful use cases" as being taken for their *face value rationale*. For example, liking content, by face value, expresses fondness for the content, whereas underlying rationales for liking could include those for the purpose of a Like being returned (i.e., reciprocation of a Like). Similarly, one could unsubscribe from a newsletter because they are simply no longer interested in it (i.e., face value rationale), whereas an underlying rationale could be to convey to the author that they are displeased with the content. By using the terms "face value" and "underlying", we take the perspective of *users* taking (and receiving) these nonverbal online actions, rather than that of platform designers—the words "faithful" and "unfaithful" reflect the perspective of platform designers.

Attributing the correct intent to nonverbal online actions is tenuous, especially when the actions are taken for underlying reasons, as the indicators do not take different forms to cue users in on senders' intentions. Furthermore, researchers have concluded that "ironic appropriations of paralinguistic digital affordances may be more widespread and salient to social media users than the faithful appropriations of [paralinguistic digital affordances]" (Hayes et al., 2016). In other words, online platform users may use nonverbal actions to indicate their underlying rationale more often than they do for their face value rationale. Consistent with the false consensus bias (Ross et al., 1977), we expect that online platform users would more frequently interpret others' actions as indicating underlying rationales just as they do themselves.

Hypothesis 2 (H2) *Most actions taken online are not interpreted as indicating their face value.*

The act of liking is not always interpreted positively. Some also used Likes to avoid others questioning their lack of liking (i.e., inaction) offline (Hayes et al., 2016). Given these varying motivations for liking, and the likelihood of users knowing these varied reasons for their nonverbal actions, we argue that people will associate different affect to each of these reasons. In particular, given the availability heuristic (Schwarz et al., 1991), we expect that the affect associated with the most readily available reason in the users' minds will be closely correlated with how they feel about the nonverbal action online.

Hypothesis 3 (H3) *The affect associated with a nonverbal action depends on the person observing the action and the reason they think the action occurred.*

Hypothesis 4 (H4) *How people feel about a nonverbal online action is predicted by the reason they most readily associate with it.*

The issues of nonverbal actions pervade other online experiences too. For example, in collaborative settings, such misinterpretations can be especially detrimental. They can lead collaborators to misunderstandings and misattributions, which often undermine team success and can even cause teams to fall apart (Whiting et al., 2019). This particularly affects creative teams, which rely heavily on these online platforms to exchange ideas, make decisions, and co-create. Inadequate scaffolding of nonverbal actions leads team members to continually question their remote colleagues' behavior, often attributing platform inadequacies to each other rather than to the platform: Why do they not actively contribute to the shared collaborative product even when they seem to be present? Why are they only deleting my content?

Not only those working in collaborative settings, but people, in general, become biased in assuming a negative intent in others' neutral actions (Nasby et al., 1980)—a phenomenon known as hostile attribution bias. This often causes people to attribute someone's negative behavior to their personality, rather than their context or situation (Ross, 1977). These biases, studied extensively in psychology, are likely more pronounced in technology-mediated interactions because there are fewer, and less effective nonverbal cues on which to base the evaluation of a stimulus (Dennis & Kinney, 1998). Our understanding of nonverbal actions online, analogous to those offline (Kraut et al., 2002), pales in comparison. This leads us to ask RQ2—what is the impact of not knowing the reason behind nonverbal online actions?

Given hostile attribution bias, we expect online platform users to assume negative intent when they do not know the reason behind a nonverbal online action. As a consequence of interpreting negative intent, we hypothesize that users will feel more negatively about the nonverbal online action.

Hypothesis 5 (H5) *People will receive nonverbal online actions more negatively when they do not know the reasons behind them—consistent with hostile attribution bias.*

3 Methods

The relative lack of understanding regarding the impact of how nonverbal actions are received and interpreted, despite the ubiquity of these actions, motivated our investigation. With nonverbal actions and rationales derived from platform users, we conducted studies that measured the impact of *knowing* (in contrast with *not knowing*) the reasons for why someone has taken a nonverbal action, and evaluated the affect of the rationales associated with these actions.

For this investigation, we compiled *nonverbal online actions* and *rationales* for these nonverbal actions. To garner a comprehensive list of nonverbal actions, we used a crowdsourcing approach and conducted a survey with open-ended questions asking participants to enlist nonverbal actions on platforms they were each familiar with. We inquired about actions across a variety of online platforms so that we would be able to consider the actions themselves and account for the design subtleties across different platforms, and also identify the three online platforms that are used in tandem amongst the average social media platform user. From this, we resulted in three nonverbal online actions that we would consider for our two main studies of this investigation. To identify rationales for online nonverbal actions, we conducted dyadic brainstorming with platform users who were well-versed in the selected platforms. We used group interviews instead of online crowdsourcing elicitation to induce interdependent discussion to cultivate subtleties in the generated reasons. In pilot testing, traditional crowdsourcing approaches did not achieve this goal. Taking these steps to identify actions and rationales enabled us to make decisions based on results, rather than relying on research instincts or arbitrary decision-making.

Using these actions and rationales, we conducted survey-based studies investigating the affect and frequency of rationales attributed to nonverbal online actions, as well as the differences in affect on recipients of these actions when they know the reasons for them, in the following two studies:

(1) Study 1. Rationales for nonverbal online actions, and
(2) Study 2. Impact of knowing why nonverbal online actions occur.

3.1 Selecting Nonverbal Online Actions

We first identified platforms that are most commonly used among participants to narrow down the platforms considered in our within-subjects study for nonverbal action evaluation (i.e., for all participants to be able to answer about the various actions on all the platforms considered). We surveyed usage (e.g., years and frequency of use) of 19 platforms known to have collaborative elements for enumerating nonverbal actions: Discord, Dropbox, Facebook, Figma, Github, Google Docs, Google Drive, Google Sheets, Google Slides, Instagram, Pinterest, Reddit, SketchUp, Slack, Snapchat, Spotify, Twitter, WhatsApp, and YouTube.

In a survey, we asked $N = 82$ participants recruited from Amazon Mechanical Turk to describe at least five possible actions on each platform from our selection that they use to build a representative corpus of actions through task analysis (Hackos & Redish, 1998). Participants in this survey were offered an incentive for more action responses, resulting in a total of 3643 platform actions across 17 of the platforms. Table 2 shows sample responses.

We lemmatized and extracted the main verb in each collected action with tidytext and textstem for R (Rinker, 2018; Robinson, 2016) and manual evaluation (Grimmer & Stewart, 2013). We categorized the actions with thematic analysis (Thomas, 2006) to be able to group all similar nonverbal online actions together. For this, we chose a representative action verb for synonyms, e.g., "edit", "change", and "modify" were grouped as "modify", and similarly all the nonverbal online actions described in Table 2 were grouped under the representative action verb of "comment". Using

Table 2 Sample of nonverbal online action responses and the platforms upon which the action was said to be possible

Platform	Described nonverbal online action
Facebook	*"Commenting on statements from friends"*
Google Drive	*"Real time comments to guide action and work on a project"*
Google Sheets	*"Ability to comment in real-time about priorities working on the sheets"*
Instagram	*"Comment on photos"*
Twitter	*"Comment on someones tweet"*
YouTube	*"Writing comments in livestreams"*

Most responses reflected a common use case in a scenario format. The sample provided here shows various ways in which participants referred to "commenting", which was one of the three nonverbal online actions that we have investigated further these groupings, we were able to identify which were the top actions taken for each platform. Figure 1 depicts occurrences of representative action verbs on each platform.

Facebook, Twitter, and YouTube were used in tandem by a majority of participants, so we restricted our remaining analysis and experimentation to just these platforms. From the most common ten representative actions across all three of these platforms, we selected three actions upon which to focus (Fig. 2).

To balance the inherent valence of the actions, we selected focal actions that were qualitatively distinct in function (e.g., modifying, adding, and commenting were seen as having a similar nature). More specifically, we selected one positive, one negative, and one neutral nonverbal online action: liking, commenting, and unfollowing (opposite of "following"), respectively (Table 3).

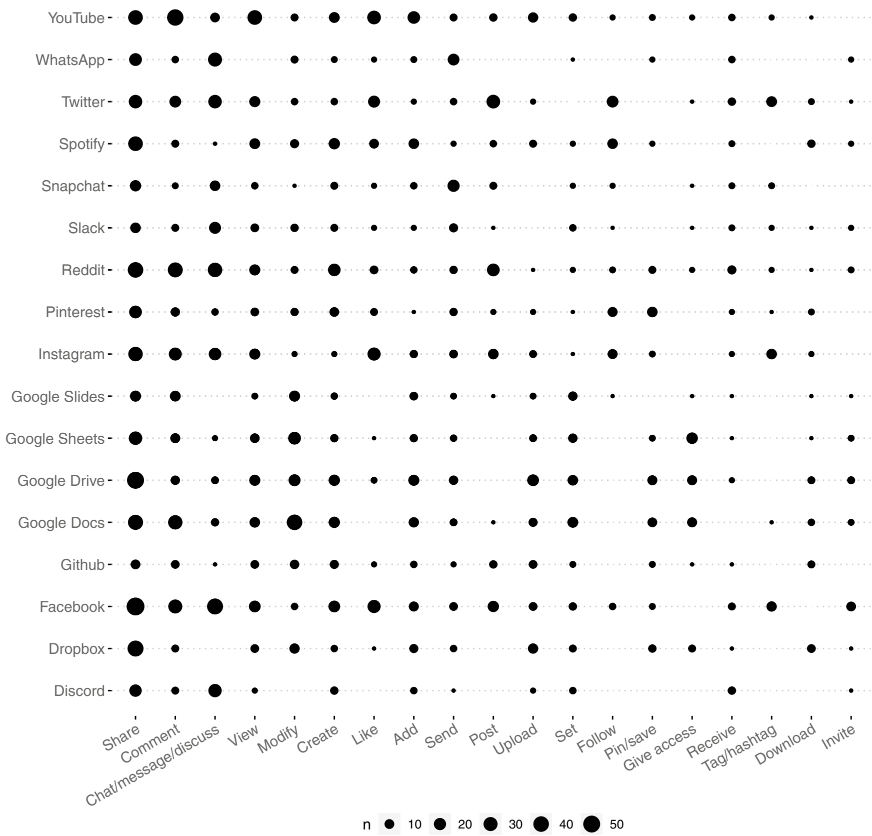


Fig. 1 Counts of representative actions identified by participants, separated by platform (ordered in descending total counts). This shows only actions that were mentioned at least 30 times by participants for ease of readability. As shown by the distribution of the counts, some nonverbal actions were more salient than others and were available across multiple platforms

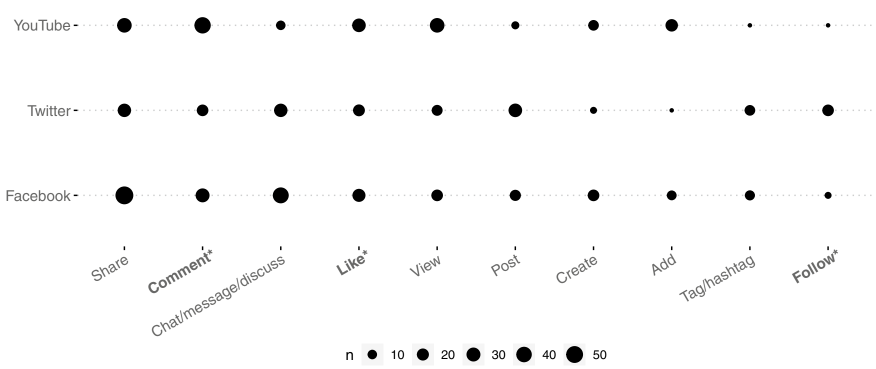


Fig. 2 The 10 most common representative actions reported by participants are present in all three focal platforms (ordered in descending total counts). The three selected nonverbal online actions for our investigation are bolded and indicated with an asterisk. For the three choices to be qualitatively distinct in function, we chose "unfollowing" instead of "following" presented here

Table 3 List of the nine actions considered for the rest of our investigation; nine total from the three most frequently mentioned nonverbal online actions and three platforms most commonly used in tandem by our survey respondents

	Liking	Commenting	Unfollowing
Facebook	Liking on Facebook	Commenting on Facebook	Unfollowing on Facebook
Twitter	Liking on Twitter	Commenting on Twitter	Unfollowing on Twitter
YouTube	Liking on YouTube	Commenting on YouTube	Unfollowing on YouTube

3.2 Identifying Rationales for Nonverbal Online Actions

Through pilots, we found that the functionality of each action had much greater influence on perceived affect than feature design between platforms—unfollowing brought about more negative affect than did liking across the board. Yet we also found differences in participants' affect ratings for each action. We attributed this to the most prominent rationale people thought of for the given action.

To better understand the effect of the rationale, we conducted four rounds of dyadic brainstorming (Paulus & Dzindolet, 1993) with participants cognizant of the functionalities of all three platforms (and well-versed in at least two) to derive all the reasons for why someone would conduct any of the nine actions in Table 3. We garnered an average of 16 reasons for why someone would perform each of the actions from each session. Eliminating similar responses led to a total of 42 reasons for all nine actions (i.e., 14 reasons each for liking, commenting, and unfollowing) (Table 4).

We conducted thematic analysis for these reasons and found that the categories that emerged from our analysis were similar to that of the motivations associated

Table 4 Selected reasons extracted from interviews and edited for each nonverbal online action, ordered, and emboldened to indicate analogous implications across rows

Liking	Commenting	Unfollowing
To align with social expectations (e.g., peer pressure)	To align with social expectations (e.g., peer pressure)	To align with social expectations (e.g., peer pressure)
To respond to request from creator/poster	To respond to request from creator/poster	To respond to request from creator/poster
To project a personal image for self-promotion	To project a personal image for self promotion	to project a personal image for self-promotion
To (express) support for creator/poster socially	To (express) support for creator/poster socially	Because they want to unfollow the creator/post
To (express) support for the content itself	To (express) support for the content itself	To reduce the amount of content they see
To promote others or their content to others	To promote/bump others or their content	To take action without others knowing
To reciprocate	To reciprocate	To express unsupported
To affect the algorithm for recommendations for themselves and others	to respond to others (e.g., creator, audience, other commenters)	To prevent receiving undesired content from creator/user (i.e., boycotting)
To influence the content creator/poster	To be malicious/out of spite	To be malicious (e.g., out of spite)
By accident, out of habit, or no reason	Because they want to comment	By accident, out of habit, or no reason
Because they like the content	To signal power	To not be seen or harassed
To confirm receipt or to acknowledge	To express an opinion	To respond to a change of mind
To make it easy to find the content again	To provide information/more context for another action	To prevent receiving undesired content
To express support at a level associated with this action but not more or less as might be associated with other	To express themselves adequately (not more or less) at the level associated with this action	To express themselves adequately (not more or less) at the level associated with this action

with paralinguistic digital affordance (Hayes et al., 2016) except for one additional rationale "mindless" (e.g., accidental) that was not included in the work by Hayes et al. (Table 5). For our investigation, we also separated "social support/grooming" into two motivation categories, as we expected different affects to be associated with these two categories as the beneficiary of the action is different—social support is for the recipient of the action whereas social grooming is to benefit the one who is taking the action.

Table 5 Six categories of reasons for performing online actions that extend prior work by Hayes et al. (Hayes et al. 2016) by differentiating between Social support and Social grooming, as well as adding Mindless

Type of rationale	Rationale categories
Face value	Literal interpretation
Underlying	Acknowledgement of viewing, Social support, Social grooming, Utilitarian purposes, Mindless

3.3 Study 1: Rationales for Nonverbal Online Actions

The first study ("Study 1") was a survey designed to examine how people associate certain reasons for nonverbal online actions on social platforms, and at which frequencies these reasons are attributed to the actions. For this, the survey was structured with two parts that were randomly presented to prevent order bias. One part asked about the participants' affect associated with the nine nonverbal online actions (Table 3). For each of the nonverbal actions, participants also selected the first reason out of the six categories of reasons (Table 5) attributed to the action and the frequency with which these six reasons for the nonverbal online action could be attributed on a slider scale from Never (0; no time ever) to Always (100; every single time). The other part asked the same affect questions but for the six reasons from Table 5. To evaluate affect, we used the Positive and Negative Affect Schedule (PANAS) (Watson et al., 1988). Of PANAS baseline measures, the particular PANAS metric we used was an international version of the shortened PANAS (i.e., I-PANAS-SF) with the 5-point Likert scale (Thompson, 2007). I-PANAS-SF words consisted of five positive words (i.e., active, alert, attentive, determined, actions inspired) and five negative words (i.e., afraid, ashamed, hostile, nervous, upset); this version was used in favor of the full PANAS scale to prevent participant fatigue as participants would be asked to evaluate their affect multiple times, either for actions or for categories of reasons. In both parts, the actions and categories were randomized and evenly presented to remove any ordering bias.

Our main hypotheses were that people feel differently about reasons behind nonverbal online actions (H3) and that how people feel about an action would be associated most strongly with the first reason attributed to the action (H4). We also hypothesized that people would attribute different reasons with an online action (H1), of which most are not the reason taken at face value (H2). These hypotheses were confirmed with χ^2 test, ANOVA, paired-samples t-tests, and linear regressions.

We administered our pre-registered survey,[3] to 70 Amazon Mechanical Turk workers, which was reduced to 51 after filtering out participants who had completed the survey in less than three minutes and had used any of the three focal platforms for "a few weeks" or less. The survey took approximately 12 minutes to complete, and the respondents were compensated with $3.00 for their participation, achieving a suggested pay rate of $15 per hour (Whiting et al., 2019).

[3] https://aspredicted.org/blind.php?x=tb8bq9.

3.4 Study 2: Impact of Knowing Why Nonverbal Online Actions Occur

The second study ("Study 2") was a survey designed as a within-subject test focused on participant affect using the PANAS metric with the control condition of not knowing the reason behind a nonverbal action ("unknown"), and with the experimental condition of knowing the reason ("known"). These PANAS questions took similar forms (with only slight modifications to wording for accurate content and grammar) as those in Study 1. The affect questions were posed as: "Thinking about the receiving person and how they would feel when someone is [one of nine actions from Table 3], which the receiving person believes is [one of the 14 reasons from Table 4],[4] to what extent do you think they would feel: [I-PANAS-SF words with the 5-point Likert scale presented as Very slightly (1) or not at all to Extremely (5)]". All questions, as shown, were asked in the Bayesian truth serum (Prelec, 2004) format such that the participants would not be biased by their own motives (e.g., impression management) and could answer more truthfully from the perspective of someone else. The participants were randomly assigned with equal distribution to four of nine total possible actions and were asked about three of 14 total possible "known reasons" to prevent survey fatigue, and the ordering of PANAS words, actions evaluated, and conditions ("known reasons" or "unknown reason") presented were all randomized to prevent any order bias.

Our main hypothesis was that awareness of the specific reason behind an action will be perceived more positively and less negatively than just the action without knowing the reason for it (H5). This hypothesis was tested with paired-sample t-tests.

We administered our pre-registered survey,[5] to 50 Amazon Mechanical Turk workers who used all three platforms and did not participate in Study 1 (data including irrelevant text responses were excluded). We eliminated responses from participants who had completed the survey in less than three minutes, had not used all three platforms for at least "a few weeks", or provided irrelevant text responses, resulting in responses from $N = 47$ participants. The survey took approximately 18 minutes to complete, and the respondents were compensated with $4.50 for their participation, achieving a suggested pay rate of $15 per hour (Whiting et al., 2019).

4 Study 1: Rationales of Nonverbal Online Actions

We present the results from Study 1 in the order of our Hypotheses 1, 2, 3, and 4—with the latter two having been pre-registered. We first present results regarding the first reasons (of the six in Table 5) that users attribute to the three nonverbal online actions of liking, commenting, and unfollowing, followed by the frequencies users

[4] This part of "which the receiving person believes is [one of the 14 reasons from Table 4]," was presented for the "known" condition; this portion was not included in the "unknown" condition.

[5] https://aspredicted.org/blind.php?x=mu8qg9.

believe recipients attribute the nonverbal actions to each of the six reasons. Then we present statistical analyses that confirm that different affects are associated with the reasons behind the nonverbal actions and that the first reason one associates with an action is highly correlated with how one feels about the nonverbal action itself.

All of our participants used all three social media platforms Facebook, Twitter, and YouTube. Most of our participants were long-term users who indicated having used the platforms Facebook (80%), Twitter (63%), and YouTube (86%) for “a few years”. Only 2–4% of the participants identified as having used these platforms for “a few weeks”.

4.1 Are People Heterogeneous in Their Interpretation of Nonverbal Actions?

We first found that everyone had diverging first reasons attributed to each action (Fig. 3), though there were clear dominant reasons for some (e.g., liking was largely seen as an action performed for social support). The most frequently occurring first reasons for each of the actions also had the highest frequency means (e.g., for unfollowing, utilitarian purpose and social grooming were the two most occurring reasons in Fig. 3 and also had the highest means of frequency of reason being attributed to in Fig. 4). However, there were also some differences; most notably, while liking for the purpose of social grooming was not mentioned as the first reason much (Fig. 3), it was the third most frequent reason attributed to the action, predicted to be the reason for more than half of the time (Fig. 4).

Performing a Chi-square test showed that the first reason attributed for the three nonverbal actions was statistically significant ($\chi^2(10, 51) = 118.26, p < 0.001$). Clear differences between each of the categories supported our decision to separate social support from social grooming, as well as to add “mindless”. Further separate Chi-square tests showed that the differences in first reasons attributed to each action were significant for each platform ($p < 0.001$); however, differences between platforms were not significant ($p > 0.1$). These analyses show that people are heterogeneous in their interpretations of nonverbal online actions themselves and not across platforms.

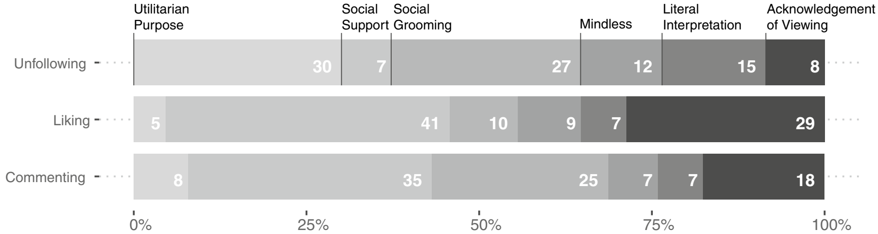


Fig. 3 Proportions of the most prominent reason (out of six possible from Table 5) that participants attributed to each of the three nonverbal actions across all three platforms ($N = 51$)

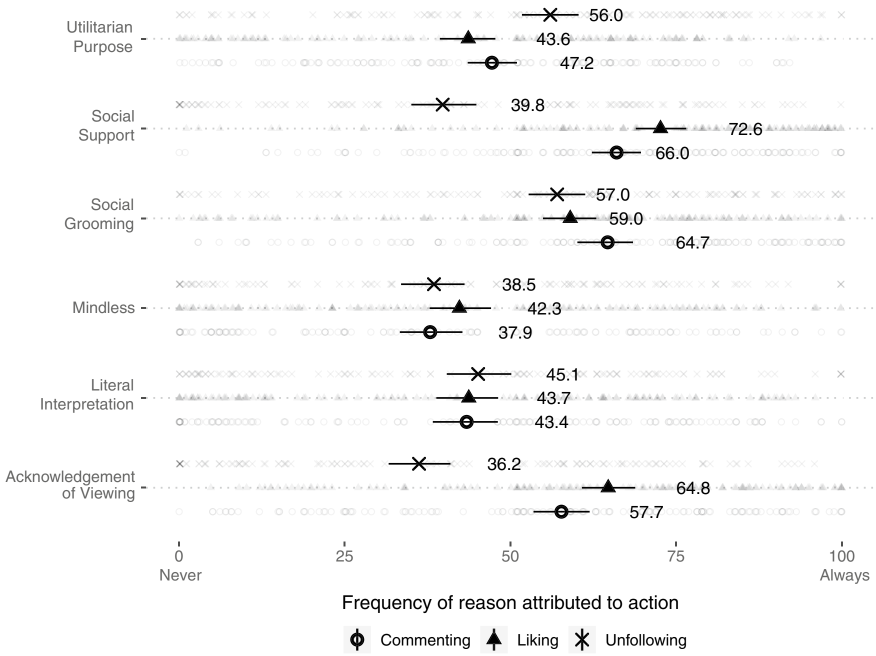


Fig. 4 Frequency ratings of the reasons attributed to each of the three actions on a scale of 0: Never to 100: Always ($N = 51$). Large markers show means and error bars indicate bootstrapped 95% CI

4.2 *Are Most Actions Taken Online not Interpreted as Indicating Their Face Value?*

Using the taxonomy of face value and underlying rationales (Table 5), we can separate literal interpretation from the other reasons. In doing so, we find that only 7–15% of the first reason attributed to the nonverbal actions taken is the face value rationale (Table 6). Figure 4 corroborates this relative infrequent interpretation of nonverbal actions for their face value rationale—literal interpretation is the third highest in

Table 6 Counts and proportions of types of rationales (as shown in Table 6) are associated most readily with each nonverbal online action

Nonverbal action	Face value rationale	Underlying rationale
Liking	10 (7%)	143 (93%)
Commenting	10 (7%)	143 (93%)
Unfollowing	23 (15%)	130 (85%)

We find that online platform users most readily associate underlying rationales (e.g., social grooming, acknowledgement of viewing) with nonverbal online actions more than their face value rationales (i.e., literal interpretation) frequency of reason attributed only for unfollowing, and not one of the top three highest rationales for liking and commenting.

4.3 Do People Have Different Affects Associated with Reasons for Nonverbal Online Actions?

Analyzing the PANAS for each of the six reasons (Table 5), we found that some reasons vary in how much positive and negative affects are associated (Fig. 5). For example, positive affect associated with an action for social support is greater than that for a literal interpretation and social grooming, and the negative affect associated with social support is lower than both. Mindless is a reason with similar aggregated positive and negative affects. Further analyzing the aggregated affects, we found that statistical significance in aggregated positive affect across the reasons (F (5, 300) $= 6.889$, $p < 0.0001$); there was no statistical significance for aggregated negative affect (F (5, 300) $= 0.754$, $p > 0.1$).

To validate our decision to distinguish between social grooming and social support, we conducted a paired t-test that showed the difference between the two reasons to be significant in their positive affect ($t(50) = -2.42$, $p < 0.05$) and negative affect ($t(50) = 2.12$, $p < 0.05$). We also evaluated social support against

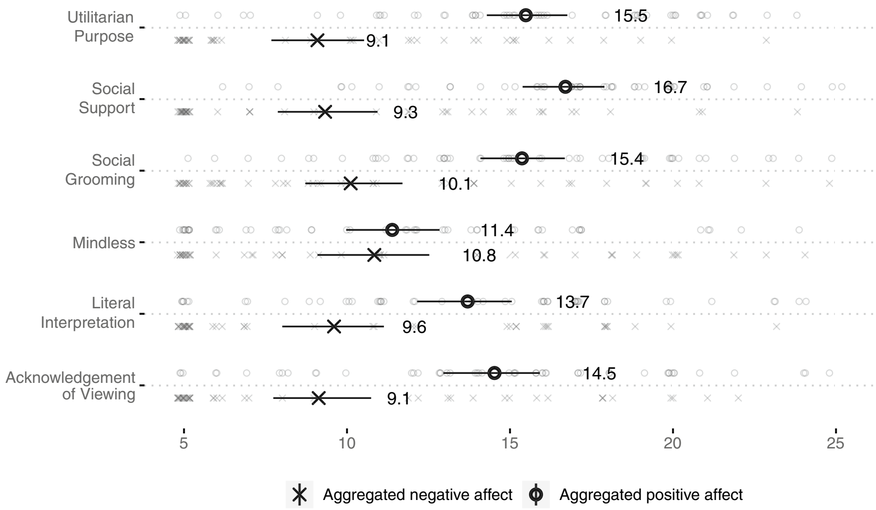


Fig. 5 Aggregated positive and negative affect values from PANAS (ranging from 5 to 25) for each of the six reasons that nonverbal actions are taken ($N = 51$). Large markers show means and error bars indicate bootstrapped 95% CI

acknowledgment of viewing, which showed similar differences between the aggregated affects. Paired t-test results revealed statistical significance in aggregated positive affect ($t(50) = -3.43, p = 0.0012$) between social support and acknowledgment of viewing but no significance for their negative affects ($t(50) = -0.59, p = 0.561$).

4.4 Can How People Feel About a Nonverbal Online Action Be Predicted by the Reason They Most Readily Associate with It?

To test our hypothesis on the expected correlation between affect of an action with the first reason attributed to the action, we conducted linear regressions for the aggregated positive and negative affects of nonverbal online actions and rationales. We found statistically significant results for both negative (F (1, 457) = 482, $p < 0.001$) and positive (F (1, 457) = 185.4, $p < 0.001$) affects, with R^2 of 0.51 and 0.29, respectively, which supports our hypothesis: how people feel about a nonverbal online action is strongly correlated with that of the reason they most readily associate with the action.

5 Study 2: Impact of Knowing Why Nonverbal Online Actions Occur

We present the results from Study 2 that address our pre-registered Hypothesis 5—that users will receive nonverbal online actions more negatively when they do not know the reasons behind them. We show that the aggregated positive affect is slightly higher for nonverbal actions taken when their reason is known compared to when it is not known. Conversely, the aggregated negative affect is slightly lower for knowing than not knowing the reasons behind a nonverbal online action. We also present how platform users believe they would feel about knowing the reason for a nonverbal online action, in particular, whether they would feel more comfortable with it.

As with Study 1, all of our participants for Study 2 used all three social media platforms Facebook, Twitter, and YouTube. Most of our participants were long-term users who indicated having used the platforms Facebook (83%), Twitter (70%), and YouTube (81%) for "a few years". Only 11–13% of the participants identified as having used these platforms for "a few weeks".

5.1 Do People Receive Nonverbal Online Actions More Positively When They Know the Reasons Behind Them?

In line with our hypothesis, we found that knowing the reasons an action occurred resulted in higher positive affect and lower negative affect than the "unknown reason" condition (Table 7).

Figure 6 shows the distribution of the aggregated negative and positive affects. Paired mean difference (Known Unknown) for aggregated positive affect was 0.115 with 95% CI [−0.869, 1.07], and that for aggregated negative affect was −0.165 with 95% CI [−1.13, 0.9].

Paired t-tests show that these differences were not shown to be statistically significant for both the negative ($t(187) = -0.673, p = 0.502$) and positive ($t(187) = 0.531$, $p = 0.596$) affects. Breaking down these aggregated affects by actions (Fig. 7), we are able to see some more variation in aggregated affect, particularly for unfollowing, but paired t-tests show that the difference is not significant ($t(119) = -0.504$, $p = 0.615$). Therefore, while the means suggest that Hypothesis 5 could be true, our statistical analyses show that it is not conclusive.

In line with participant results from the PANAS questions, we also found alignment of results from our final question posed to measure guesses about our main hypothesis via a direct question after completion of other parts of the experiment: "Do you think knowing the reason for an online action will make you more comfortable with it?" To this question, 76% of participants responded "Yes, it will make me more comfortable with the online action", 21% responded "No, it will not change how comfortable I am with the online action" (only one participant selected "No, it will make me less comfortable with the online action"). Therefore, while our affect analysis is not conclusive on whether Hypothesis 5 can be supported, we see that the participants expect to feel more comfortable knowing the reason behind nonverbal actions.

Table 7 Means of aggregated negative and positive affects for when reasons for actions are known and unknown (standard deviation also reported)

Reasons for nonverbal action	Aggregated affect		
	Negative	Positive	Relative
Known	10.15 (SD = 4.91)	15.40 (SD = 4.68)	5.25 (SD = 5.23)
Unknown	10.31 (SD = 5.29)	15.29 (SD = 4.73)	4.97 (SD = 5.91)

Also included is 'Aggregated relative affect' calculated as 'Aggregated positive affect' minus 'Aggregated negative affect'

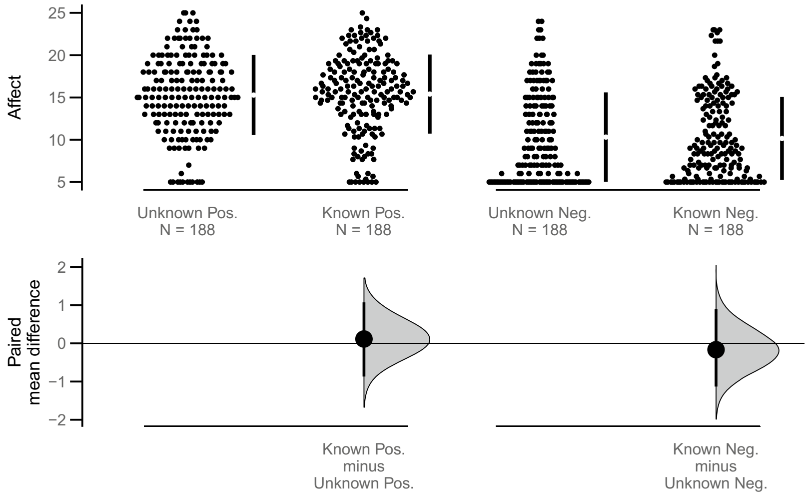


Fig. 6 Paired mean difference between knowing and not knowing the reason for an action across positive and negative affect are shown in a Cumming estimation plot (Ho et al. 2019). PANAS affect scores (broken out by sign, and aggregated within known conditions) are plotted on the upper axes. On the lower axes, each paired mean difference is plotted as a bootstrap sampling distribution. Mean differences are depicted as dots; 95% confidence intervals are indicated by the ends of the vertical error bars. Knowing the motive of an action leads to slightly improved affect—more positive, less negative—however, this difference is not significant

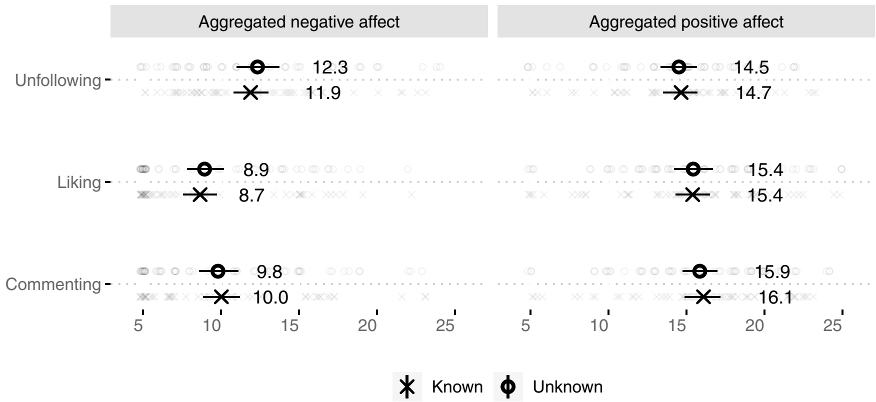


Fig. 7 Aggregated positive and negative affects from PANAS (ranging from 5 to 25) for each online action across 2 conditions—the reason for an action being known and the reason being unknown. Large markers show means and error bars indicate bootstrapped 95% CI

6 Discussion

Nonverbal online actions are ubiquitous and we, as users, interact with other users through these actions frequently. However, there is a relative lack of understanding of the effects of these nonverbal online actions—and the unclarity of intention of these actions—on the users at the receiving end. Through our investigation, we build an understanding of such nonverbal actions online by conducting two studies: one measuring the impact of knowing why actions occur and another evaluating the rationales for nonverbal actions online. For these studies, we selected the nonverbal actions to study by taking crowdsourcing approach that garnered 3643 nonverbal actions on 19 different platforms, and found that the three most commonly used platforms amongst our participants were Facebook, Twitter, and YouTube. We identified the ten most frequently mentioned nonverbal actions amongst these three concomitantly used platforms, and from them chose the three nonverbal interactions—commenting, liking, and unfollowing—for the rest of the investigation. We then identified rationales for nonverbal actions by conducting dyadic brainstorming sessions and resulted in 14 reasons each for commenting, liking, and unfollowing. Comparing our resulting total of 42 reasons with existing taxonomy of reasons from prior work, we found that an existing category of social support/grooming were two distinct reasons as they benefited different people and therefore were felt differently by the recipients, and also found an additional category that we call "mindless" to capture the actions taken without much thought (e.g., habitual reaction) or by mistake (e.g., wrong click). We resulted in 42 specific reasons total for commenting, liking, and unfollowing, and six reason categories of literal interpretation, acknowledgement of viewing, social support, social grooming, utilitarian purposes, and mindless (Table 5).

In Study 1, we conducted a survey in which we asked participants to evaluate expected PANAS for each of the actions, identify the first category of reason they would attribute to each of the three nonverbal actions, how frequently the actions are taken for each of the reasons, and the expected PANAS for each of the reason categories. We found that the first reasons participants attributed to the actions were significantly different, frequencies of reasons attributed to actions depend highly on the action itself, and that the differences in affect between the reasons are statistically significant. We also found statistical significance in linear regressions between the nonverbal action PANAS with the PANAS of the first reason attributed to the action. These divergent views, if left unarticulated and unresolved, can lead to misunderstandings and the accumulation of negative emotion, often unbeknownst to the interactants. Through this study, we confirmed Hypotheses 1, 2, 3, and 4.

In Study 2, we conducted a survey in which we asked participants to rate expected PANAS regarding nonverbal actions taken for unknown reasons and for known reasons. From analyzing their PANAS evaluations, we found a mild trend that knowing the reason for a nonverbal action led recipients to report more positive (and less negative) affect, regardless of the action. This finding suggests that people's in situ perceptions of knowing reasons for online nonverbal actions are less impactful than in an imagined situation (i.e., unknown reason). Separating PANAS out for the actions, we find that the aggregated positive affect is consistently higher for when the reasons for actions are known compared to when reasons are unknown. Similarly, aggregated negative affect is lower when the reasons for actions are known compared to when reasons are unknown, except in the case of liking. Moreover, the majority of our participants expected to feel greater comfort with knowing the reasons for nonverbal online actions. Through this study, we found that our Hypothesis 5 could be supported.

In short, our results show that users' interpretations of nonverbal online actions are varied and that the way they interpret them dictates the individual affect of the action. While the manifesting use cases are not "faithful" to the presumed designer's intent, they are used in these ways due to the affordances of platform design. Therefore, we can apply our findings from this investigation to understand how we might design nonverbal online action indicators such that misunderstandings from nonverbal interactions and behavior evidenced online do not perpetuate, enabling a more positive user experience.

6.1 Users Feel Better When They Know the Reasons for Nonverbal Actions

Our results show that users expect they would feel slightly more positive about a nonverbal online action if they know the reason for why actions are taken. Although this result is not statistically significant, whether they would feel as similarly with or without reasons in situ has yet to be answered. Regardless of whether they would feel a more positive affect, 76% of our participants indicated they would be more comfortable with knowing the reasons for nonverbal actions. This uncertainty is felt by the recipient in a more negative way than if they had known the reason behind the action. This suggests hostile attribution bias in online platforms, and necessitates the need to understand how to better design online tools to ameliorate the perceived affect due to this uncertainty.

6.2 Design of Nonverbal Actions Afford Ambiguity and Interpretability

While asking why someone is acting in a certain way in-person may not be socially acceptable or timely, this may be possible in online platforms where the social norms are dynamic and prone to change. Just as online platforms introduced the Like button to quickly express appreciation by the users and has now become a norm, they can change the landscape of how users communicate and interact with each other. Such nonverbal online actions as the Like button enable users to interact with one another efficiently, yet our results indicate that these are laden with multiple meanings by the senders and interpreted in multiple ways by the recipients. Whether liking is interpreted as acknowledgment of viewing or social support—both of which are likely reasons from Figs. 3 and 4—significantly changes how users feel about (in their positive affect) and receive the likes, as found from our t-tests.

6.3 Design Implications for Nonverbal Actions

Design implication 1: Affordance of options. The multitude of reasons interpreted from an observed online action (as shown in Figs. 3 and 4), and concomitantly the reasons possible for taking a nonverbal action, indicates that these actions perform more than one function. This suggests that there may not be enough nonverbal actions to capture all the nuances. In other words, online communication is more complex than design often encompasses.

As studies by Facebook have shown,[6] the Like button was not enough to capture the range of emotions users wished to communicate, leading Facebook to implement a set of seven Reactions. Still, many platforms are not providing such diverse nuances of expressions between users. Furthermore, direct communicative features such as Likes are not the only actions that require options. Anecdotes of presence indicators often being disabled outright due to their overly rigid nature is one example. As such, designs that embrace the complexity of socio-cultural norms in addition to the needs of rational information exchange between users is necessary.

Design implication 2: Uncertainty affects user actions. The uncertainty, and the subsequent negative affect, may currently be influencing how users interact with each other. As we continue to interact with each other via virtual means, users' feelings about others can accumulate for the worse through such uncertain encounters. Such negative (or less positive) emotion accumulation (Résibois et al., 2017) can result in stress for the individual and strained interactions, thereby inducing a negatively reinforced cycle. Therefore, to enable more sustainable remote interactions, designing ways to incorporate rationale into our online nonverbal actions will be empowering for users.

Design implication 3: Room for embedding reasons. As there are a wide array of reasons behind users actions online, one design implication is that platforms should embed rationale in mediating user interactions with nonverbal actions in order to help users be clear about and therefore feel better about receiving nonverbal online actions.

The importance of knowing the rationale behind a certain action does not end at collaborative and social media platforms (Park & Lee, 2021), but also extends to the context of video-conferencing (Park & Whiting, 2020). In this work, authors found that elucidating the rationale behind why someone may be turning their video off

[6] https://about.fb.com/news/2016/02/reactions-now-available-globally/, https://slate.com/technology/2016/02/facebook-s-5-new-reactions-buttonsare-all-about-data-data-data.html. Accessed December 3, 2020.

can lead to less negative affect perceived by the observers. Enabling interactants to do so through the design of the platform will also help to ensure the comfort and psychological safety in doing so.

This may also be important in the context of fake news, which social media platforms have aimed to curb by implementing various design features (Elkind, 2020; Pennycook & Rand, 2020; Roth & Pickles, 2020). The spread of fake news, which has been found to be more due to humans than bots (Vosoughi et al., 2018), may be inhibited by designing a method for deliberation (Bago 2020). While users may not want to spread misinformation (Pennycook et al., 2019), they may want to share news that they perceive as being ridiculous or outrageous without espousing the validity of it. Embedding rationales could be one way of designing for such cases.

Design implication 4: Ethical considerations for online users. The design of nonverbal online actions also requires ethical considerations. Today's tools have undermined the communication landscape for some users—they can no longer count on nonverbal signals they expect to see in other settings-and this is an important reason for more attention to be paid to the careful design of nonverbal actions. Other users may be positively enabled by the current situation—people uncomfortable expressing their feelings sometimes thrive on mediated settings that may be avatar-based or anonymized—suggesting that a solution to this situation is not as simple as instituting a specific policy for all users. As a striking example of this issue, some software used for online lectures allows instructors to force all students to enable their video (Mouton, 2020), which frustrates students and exacerbates privacy concerns and power imbalances in the interaction. Because of the complex nature of this issue, we identify the design of nonverbal interaction as an area requiring thoughtful consideration and substantial further research—it has the potential to move us beyond reproducing an in-person status quo, but can only achieve that goal with the utmost care.

7 Limitations and Future Work

Extending the nonverbal online actions and their indicators studied to encompass a wider variety would allow us to build a more nuanced understanding of nonverbal actions that exist online and their consequences.

Our results were primarily derived from survey data based on platform users' reflections. The success of this approach requires users to have predictable interpretations of online action reasons, so we cannot ascertain if participant perceptions are consistent and that our construct validity holds. Continuing this work by conducting equivalent studies in the wild (i.e., on these social platforms) with greater ecologically validity will be important next steps that can help corroborate our findings.

Our work shows that nonverbal online actions ought to be better mediated than they are currently, as they result in reduced positive affect when interactants do not know the reasons for the actions. Uncovering nonverbal actions that lead to more negative affects as well as studying how to embed rationale into our nonverbal actions are invaluable directions for further research. Understanding how to parse nonverbal actions in more collaborative computer-supported cooperative work and computer-mediated communication settings also poses exciting areas of study.

8 Conclusion

From our investigation, we found that people associate different reasons to why others take nonverbal online actions, and their interpreted rationales influence how they feel about these actions. We also found that providing a rationale of a nonverbal online action is better than not doing so, regardless of what the action is. Thereby showing that nonverbal online actions themselves can be, in effect, louder than the words that describe the action (e.g., how we feel about the action of liking can feel more negative and less positive than the word "Like" itself), and the merit of providing a rationale for users' interactions. We derive four design implications for online platforms that provide user-user interactions, one of which is the implication for embedding a means for providing rationale—this could help interactants on social media platforms to gain more positive affect towards (i.e., feel better about) nonverbal online actions, which are usually ambiguous in intent. Furthermore, we have also developed a survey instrument and a study methodology that support our hypotheses. This instrument and methodology together can be used in other investigations to obtain even greater intellectual merit.

Online platforms for work are no longer our future, but our present. As misunderstandings and skewed signals proliferate in computer-mediated communication, understanding and properly scaffolding nonverbal actions with users' rationale are vital. Studied in the context of social media platforms, our findings have implications for design teams and collaborating teams as well, as collaborators require even more "good vibes" and clearer understanding of each others' intent in their actions. The importance of embedding rationale in nonverbal online actions are applicable to a wide range of system and service design problems in which peoples' remote interactions are computationally mediated. As users become more comfortable with and accustomed to providing rationales for their online actions, this may become a norm, even offline, and contribute towards greater clarity and empathy in communication and less communication breakdown. Therefore, in addition to making design teams more effective, this work may also cast a light on solving these societal challenges too.

Acknowledgements We are grateful for the work of hundreds of Amazon Mechanical Turk participants, remote study participants, and our Hasso Plattner Design Thinking Research Program workshop participants for making the data development possible.

We're also grateful for Larry Leifer for feedback in developing this work. This research was supported by the Hasso Plattner Design Thinking Research Program.

References

Andersen, P. A. (1991). When one cannot not communicate: A challenge to Motley's traditional communication postulates. *Communication Studies, 42*, 4, 309–325.

Antonijević, S. (2013). The immersive hand: Non-verbal communication in virtual environments. In *The Immersive Internet* (pp. 92–105). Springer.

Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General, 149*(8), 1608–1613.

Becker, B., & Mark, G. (1998). Social conventions in collaborative virtual environments. In: *Proceedings of Collaborative Virtual Environments* (pp. 17–19).

Benford, S., Bowers, J., Fahlén, L. E., Greenhalgh, C., & Snowdon, D. (1995). User Embodiment in Collaborative Virtual Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI'95)* (pp. 242–249). ACM Press/Addison-Wesley Publishing Co., USA. https://doi.org/10.1145/223904.223935.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (p. 127–149). American Psychological Association.

Dennis, A. R., & Kinney, S. T. (1998). Testing media richness theory in the new media: The effects of cues, feedback, and task equivocality. *Information Systems Research,* 9(3), 256–274.

DeSanctis, G., & Poole, M. S. (1994). Capturing the complexity in advanced technology use: Adaptive structuration theory. *Organization Science, 5*(2), 121–147.

Edwards, R., Bybee, B. T., Frost, J. K., Harvey, A. J., & Navarro, M. (2017). That's not what I meant: How misunderstanding is related to channel and perspective-taking. *Journal of Language and Social Psychology, 36*(2), 188–210.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32*(1), 88–106.

Elkind, E. (2020, November 11). Maria Ressa calls for "emergency intervention" in social media: It is untenable for democracy. https://www.cbsnews.com/news/maria-ressa-social-media-democracy/. Accessed November 16, 2020.

Goffman, E. (1978). *The presentation of self in everyday life.* Harmondsworth: Penguin Books.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis, 21*(3), 267–297.

Groshek, J., & Bronda, S (2016). How social media can distort and misinform when communicating science. *The Conversation, 30.*

Guerrero, L. K., DeVito, J. A., & Hecht, M. L. (1999). *The nonverbal communication reader*. Waveland Press Prospect Heights.

Guye-Vuillème, A., Capin, T. K., Pandzic, S., Thalmann, N. M., & Thalmann, D. (1999). Nonverbal communication interface for collaborative virtual environments. *Virtual Reality, 4*(1), 49–59.

Hackos, J. T., & Redish, J. (1998). *User and task analysis for interface design* (Vol. 1). Wiley.

Hayes, R. A., Carr, C. T., & Wohn, D. Y. (2016). One click, many meanings: Interpreting paralinguistic digital affordances in social media. *Journal of Broadcasting & Electronic Media*, 60(1), 171–187.

Ho, J., Tumkaya, T., Aryal, S., Choi, H., & Claridge-Chang, A. (2019). Moving beyond P values: Data analysis with estimation graphics. *Nature Methods, 16*(7), 565–566.

Kendon, A. (1986). Some reasons for studying gesture. *Semiotica, 62*(1–2), 3–28.

Kendon, A., Sebeok, T. A., & Umiker-Sebeok, J. (2010). *Nonverbal communication, interaction, and gesture: selections from Semiotica* (Vol. 41). Walter de Gruyter.

Knapp, M. L., Hall, J. A., & Horgan, T. G. (2013). *Nonverbal communication in human interaction.* Cengage Learning.

Kraut, R. E., Fussell, S. R., Brennan, S. E., & Siegel, J. (2002). Understanding effects of proximity on collaboration: Implications for technologies to support remote collaborative work. *Distributed Work*, 137–162.

Leavitt, A., Keegan, B. C., & Clark, J. (2016). Ping to win? Non-verbal communication and team performance in competitive online multiplayer games. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 4337–4350).

Lee, S. Y., Hansen, S. S., & Lee, J. K. (2016). What makes us click "like" on Facebook? Examining psychological, technological, and motivational factors on virtual endorsement. *Computer Communications, 73*(2016), 332–341.

Matsumoto, D., Frank, M. G., & Hwang, H. S. (2012). *Nonverbal communication: Science and applications*. Sage Publications.

Moore, R. J., Ducheneaut, N., & Nickell, E. (2007). Doing virtually nothing: Awareness and accountability in massively multiplayer online worlds. *Computer Supported Cooperative Work (CSCW), 16*(3), 265–305.

Motley, M. T. (1990). On whether one can (not) not communicate: An examination via traditional communication postulates. *Western Journal of Communication (Includes Communication Reports), 54*(1), 1–20.

Mouton, M. (2020). Using zoom to teach synchronously online. *Technology Tools for Teaching in Higher Education, The Practical Handbook Series.*

Nasby, W., Hayden, B., & DePaulo, B. M. (1980). Attributional bias among aggressive boys to interpret unambiguous social stimuli as displays of hostility. *Journal of Abnormal Psychology, 89*(3), 459.

Nguyen, D., & Canny, J. (2005). MultiView: Spatially faithful group video conferencing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 799–808).

Okada, K. I., Maeda, F., Ichikawaa, Y., & Matsushita, Y. (1994). Multiparty videoconferencing at virtual social distance: MAJIC design. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work* (pp. 385–393).

Park, S. Y., & Lee, S. W. (2021). Lost in Co-curation: Uncomfortable Interactions and the Role of Communication in Collaborative Music Playlists. In *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW1), 1–24.

Park, S. Y., & Whiting, M. E. (2020, August). Beyond zooming there: Understanding nonverbal interaction online. The New Future of Work Symposium.

Paulus, P. B., & Dzindolet, M. T. (1993). Social influence processes in group brainstorming. *Journal of Personality and Social Psychology, 64*(4), 575.

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. (2019). Understanding and reducing the spread of misinformation online. Unpublished manuscript https://psyarxiv.com/3n9u8.

Pennycook, G., & Rand, D. (2020, March 24). The right way to fight fake news. https://www.nytimes.com/2020/03/24/opinion/fakenews-social-media.html?smid=tw-nytopinion&smtyp=cur. Accessed November 30, 2020.

Pohl, H., Domin, C., & Rohs, M. (2017). Beyond just text: Semantic emoji similarity modeling to support expressive communication. *ACM Transactions on Computer-Human Interaction (TOCHI), 24*(1), 1–42.

Prelec, D. (2004). A Bayesian truth serum for subjective data. *Science, 306*(5695), 462–466.

Résibois, M., Verduyn, P., Delaveau, P., Rotgé, J. Y., Kuppens, P., Van Mechelen, I., & Fossati, P. (2017). The neural basis of emotions varies over time: Different regions go with onset-and offset-bound processes underlying emotion intensity. *Social Cognitive and Affective Neuroscience, 12*(8), 1261–1271.

Rinker, T. W. (2018). *textstem: Tools for stemming and lemmatizing text*. Buffalo. http://github.com/trinker/textstem version0.1.4.

Rodrigues, D., Lopes,D., Prada, M., Thompson, D., & Garrido, M. V. (2017). A frown emoji can be worth a thousand words: Perceptions of emoji use in text messages exchanged between romantic partners. *Telematics and Informatics, 34*(8), 1532–1543.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In *Advances in experimental social psychology* (Vol. 10, pp. 173–220). Elsevier.

Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology, 13*(3), 279–301.

Roth, Y., & Pickles, N. (2020, 11 May). Updating our approach to misleading information. https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.html. Accessed November 16, 2020.

Ruesch, J. (1953). Synopsis of the theory of human communication. *Psychiatry,* 16(3), 215–243.

Ruesch, J. (1955). Nonverbal language and therapy. *Psychiatry, 18*(4), 323–330.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: another look at the availability heuristic. *Journal of Personality and Social Psychology, 61*(2), 195.

Silge, J., & Robinson, D. (2016). tidytext: Text mining and analysis using tidy data principles in R. *JOSS, 1*(3). https://doi.org/10.21105/joss.00037.

Sumner, E. M., Ruge-Jones, L., & Alcorn, D. (2018). A functional approach to the Facebook Like button: An exploration of meaning, interpersonal functionality, and potential alternative response buttons. *New Media & Society, 20*(4), 1451–1469.

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation, 27*(2), 237–246.

Thompson. E. R. (2007). Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of Cross-Cultural Psychology, 38*(2), 227–242.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146–1151.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063.

Whiting, M. E., Blaising, A., Barreau, C., Fiuza, L., Marda, N., Valentine, M., & Bernstein, M. S. (2019). Did It have to end this way? Understanding the consistency of team fracture. In *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019) (pp. 1–23).

Whiting, M.E., Hugh, G., & Bernstein, M.S. (2019). Fair work: Crowd work minimum wage with one line of code. In: *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* (Vol. 7. 197–206).

Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., & Merget, D. (2007). The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. *CyberPsychology & Behavior, 10*, 1, 115–121.

The Neuroscience of Empathy: Research-Overview and Implications for Human-Centred Design

Irene Sophia Plank, Julia Petra Ariane von Thienen, and Christoph Meinel

Abstract Empathy is a central concept in design thinking. According to human-centred design, developers of novel products shall strive for a good understanding of product users, in order to design for their needs. This requires sophisticated cognitive capacities on behalf of the designers: being able to distinguish between their own knowledge states and needs versus that of the users. An IT expert who develops a banking app for elderly people must be able to imagine what it is like for an extreme user—such as an elderly person, who can barely use a mobile phone—to learn about online banking. What goes on in the design thinker's mind when he or she tries to understand a user? What sub-capacities are involved, where the design thinker may be more or less capable? What routes to "understanding others" are promoted and taught by means of design thinking empathy methods? Neuroscience has produced a cornucopia of research studies on the biological underpinnings of understanding others. In this chapter, we review insights from neuroscience on how humans understand fellow people. This includes an overview of conceptual distinctions and sub-capacities, such as *empathy* versus *compassion*, or *affective* versus *cognitive routes of social understanding*. We also review measurement approaches that can be used in design thinking research and human-centred design practice to assess people's abilities of understanding others. Moreover, the chapter discusses biases and pitfalls in understanding others, such as a natural tendency of the human brain to react less (to "empathise less") with persons who seem to be particularly different from us.

I. S. Plank (✉)
Faculty of Philosophy, Berlin School of Mind and Brain, Humboldt-Universität Zu Berlin, Luisenstr. 56, 10117 Berlin, Germany
e-mail: irene.sophia.plank@hu-berlin.de

Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

Einstein Center for Neurosciences, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

J. P. A. von Thienen · C. Meinel
Hasso Plattner Institute for Digital Engineering, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany
e-mail: julia.vonthienen@hpi.de

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_6

1 Empathy in Design Thinking and Human-Centred Design

Empathy is a fundamental concept in design thinking and human-centred design. While design thinking is understood in various ways by different communities, the view of design thinking as an empathic approach to innovation is emphasised by many. According to common views, design thinking invokes "a human-centred perspective, where innovators build empathy with users" (Verganti et al., 2019, p. 1). Design thinkers endorse a "human-centric point of view" (Meinel & Leifer, 2011, p. xv) as opposed to a technology-centric view, and design thinkers work to solve problems "in ways that satisfy human needs" (ibid.) (Meinel & Leifer, 2012). The mindset of design thinkers is to "focus on human values" (d.school, 2010, preface).

The key role that empathy plays in design thinking is also reflected in models of the creative process used in design thinking education, like the one developed and taught at Stanford University (see Fig. 1).

To arrive at valuable innovation, design thinkers shall first and foremost empathise with potential users, for whom novel solutions are to be developed. Bernard Roth, an academic director and co-founder of Stanford's design thinking institute *d.school* provides an example of how empathy can inform design process in the most beneficial ways.

> A four-person interdisciplinary team of Stanford Masters degree students were asked to create something that would change people's lives. They initially had difficulty finding anything suitable. Eventually they happened upon several janitors that cleaned the building at night. The janitors were new immigrants to America. They were from Mexico and spoke almost no English. [...] The students found out that the janitors had very little knowledge about financial matters and were being taken advantage of during almost every transaction. [...] The students undertook to develop and deliver Spanish-language lessons about financial

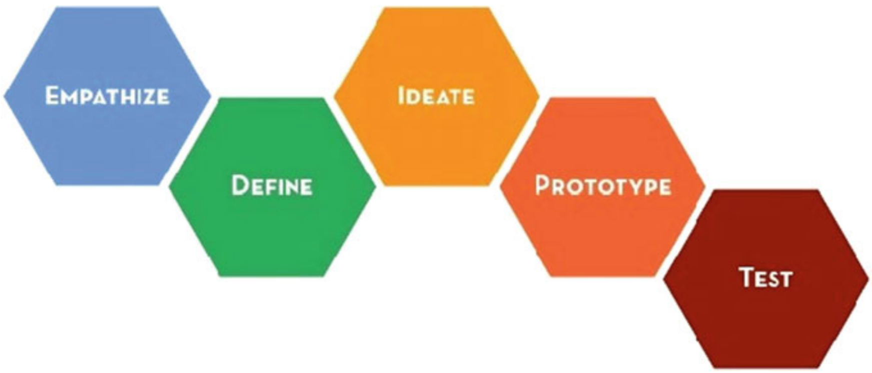


Fig. 1 Classic d*esign thinking process model*, as used at Stanford University: Design thinkers first empathise with potential users in order to develop worthwhile solutions for them. Subsequent steps in the creative process are the definition of a creative project, including the specification of a problem that shall be solved. In the ideation phase, solution ideas are thought up. One or more solution ideas get selected and implemented as prototypes. Finally, the prototypes get tested to see if they solve the creative problem, or whether the solution approach needs to be iterated

> planning and ways to conduct financial matters. The results were astounding. In six months two of the janitors was able to save 9 percent of their less than $40,000 annual salaries. One of the students was so inspired he went on to found a company, called Juntos, that allows people to use ordinary cell phones to learn about and deal with their finances. [...] In recognition of their potential for uplifting the world's poor, Juntos received the G20's Financial Inclusion Innovation Award. I still have the original project notebook from this group. Whenever I look at the notebook I am moved to tears by ***the empathy the students felt for the janitors***.
>
> (Roth, 2017, p. 82f, our emphasis)

In this example, the students seem to resonate emotionally with the janitors' unfortunate situation—so much so, in fact, as to gain ample motivation for a long-term engagement on behalf of the janitors and people in similar situations. The design thinking students react with compassion. In addition, they distinguish masterfully between their own knowledge state (being fluent in English and knowing about pitfalls of financial transactions in the USA) versus the knowledge state of their users. Due to this ability of understanding different knowledge states, the design thinking students can design suitable teaching materials for the janitors, so that they too build up knowledge on how to organise finances effectively in their new home country.

At the *Hasso Plattner Institute* in Germany with the associated *Digital Engineering Faculty* of *Potsdam University*, a sound understanding of users is essential in almost all projects. While the students themselves are IT experts, most users of digital engineering solutions are far less versatile in handling technology. If, for instance, a team of digital engineers develops a banking app, this app needs to be usable even by some elderly people, who have had little exposure to IT technology throughout their lives. For some elderly people, only the simplest and most straightforward dialogues will be understandable. In order to design usable software, the students therefore need to be able to imagine and understand what kind of knowledge and emotions the users will have when interacting with the novel product. This can be quite challenging as the users may be extremely different from the IT students who develop the product.

Design thinking helps developers understand potential users by highlighting the importance of empathy and by training empathy methods. Examples of standard techniques are the *Interview for Empathy*, *Why-How-Laddering* to elucidate user needs and values, or the *Empathy Map*—these and many more can be found in the d.school's method compilation *Bootcamp Bootleg* (2010). There are also design thinking classes specifically tailored to train participants in the use of empathy as a means to worthwhile innovation, such as the online classes *Inspiration for Design: A Course on Human-Centred Research* by Taheri, Schmieden and Mayer, *Human-Centred Design: From Synthesis to Creative Ideas* and *Human-Centred Design:*

Building and Testing Prototypes by Schmieden et al. (2017, 2018, 2019). Even apart from understanding user needs to arrive at worthwhile innovation, empathy plays a key role in design thinking. Another characteristic of the approach is to foster interdisciplinary teamwork and collaboration across multiple stakeholders (Plattner et al., 2009). Empirically, it is clear that different perspectives and training backgrounds induce communication difficulties compared to collaboration in monodisciplinary teams von Thienen et al. (2011). Thus, in order to collaborate fruitfully, design thinkers need marked skills of understanding others to deal with the perspectives of team members or further stakeholders, who typically come from highly diverse cultural and academic backgrounds. Research has found that even short design thinking trainings of only one week yield a significant increase in the student's self-rated confidence to work fruitfully in multiperspective teams (Traifeh et al., 2020). To what extent this increased confidence goes along with, or is even caused by, an increase in social-understanding competencies is still an open question, calling for further research.

Another pertinent question that surfaces specifically in the context of design thinking education concerns the source of inspiration that drives creative projects in one or another work direction. The classic approach is to let students empathise with potential users—as described by Roth above. More recently, however, different approaches have been probed, where the students start instead with their own personal passions and interests to kick off creative projects (von Thienen et al., 2021). In both cases, emotions play a key role in the creative process: In the first case, attention is directed towards the emotions and needs of others. In the latter case, attention is directed towards one's own emotions and needs. This too raises questions of empathy. In particular, how similar or different are processes of understanding someone else's emotions and intentions, compared to understanding one's own? When inventors use their own passions and visions as starting points to commence creative projects, do they—in a sense—empathise with themselves?

Given the importance of empathy in design thinking, clearly there is a need for related research to underpin established practices with sound scientific understandings. In 2017, Eva Köppen published her dissertation *Empathy by Design: Untersuchung einer Empathie-geleiteten Reorganisation der Arbeitsweise* (English: *Empathy by Design: Investigation of an Empathy-Led Reorganisation of Work*) (Köppen, 2017). Here, she discusses how increasing calls for empathy in the design of products, but also in terms of interactions at the workplace, impact and change dynamics in the private sector. In 2019, we added a dedicated design thinking research initiative on *human needs* (Hasso Plattner, 2019, 2020). Project work includes a review of literature and practices on what "human needs" mean, how people understand others and themselves, and what role this understanding plays in the development of innovation. Since collaboration is a major characteristic of human innovation capacities (cf. Corazza & Thienen, 2021; von Thienen et al., 2021), the relationship of empathy skills–collaboration–innovation has also been a major topic in this research field.

Overall, our design thinking research project on human needs aims to collocate scientific resources for the community of design thinkers and human-centred designers, so as to facilitate increased exchange between basic research and applied innovation projects. Among the resources that have been rendered available to date, lectures are accessible online. Irene Sophia Plank (Berlin School of Mind and Brain, Humboldt-Universität zu Berlin) reviews the neuroscience of empathy (Plank, 2020), Dr. Caroline Szymanski (Max Planck Institute for Human Development and HPI) explores how insights from social neuroscience can benefit the study and development of design thinking (Szymanski, 2019b). She also discusses the role of inter-brain synchrony as an indicator and potentially a causal mechanism of collaboration (Szymanski, 2019a). Dr. Marwa El Zein (Institute of Cognitive Neuroscience, University College London) elucidates neuroscientific underpinnings and psychological motivations behind collective decision-making, as opposed to single-person work and single-person decisions (el Zein, 2019). Dr. Laura Kaltwasser (Berlin School of Mind and Brain, Humboldt-Universität zu Berlin) discusses the role of emotions in social decision-making (Kaltwasser, 2019). Dr. Julia von Thienen (HPI) points out that a continuity of joint social norms fosters incremental innovation, while a change of social norms and values promotes radical innovation (Thienen & Santuber, 2020). Joaquin Santuber (HPI) examines joint social norms as a framework for efficient and smooth collaboration, and how this manifests in the team member's physiology (Thienen & Santuber, 2020). Dr. Julia Rodríguez Buritica (Biological Psychology and Cognitive Neuroscience, Free University of Berlin) explains how information about the behaviour and the experiences of others inform individual decisions, from childhood over adolescence to adulthood (Rodríguez Buritica, 2020).

In this chapter, our aim is to collocate pertinent resources on the neuroscience of empathy for the interested community. This includes a review of key competencies that humans use to understand others. Research shows that it makes sense to differentiate between an affective versus a cognitive route to understanding others (Sects. 2, 3, 4 and 5). Moreover, we review available tests and indicators that help to measure the empathy capacities of individuals (Sects. 6 and 7). One assessment approach is discussed in further detail, which suggests that understanding others is specifically challenging for people when the other person is perceived to be markedly "different from oneself" (Sect. 8). The chapter concludes with a brief review of how neuroscientific constructs and measurement approaches relate to design thinking practices and theory (Sect. 9).

2 Understanding Others as a Sophisticated Human Capacity

For humans, it seems natural and easy to coordinate and collaborate with conspecifics, in a myriad of different situations. This was already noticed by Aristotle, who addressed human social behaviour in his treatise on *Politics* as a "social instinct in all of us" (Jowett, 1985) and many more scholars after him came to make similar observations. Given outstanding characteristics of social behaviour in humans, scholars asked early on to what extent other species possess similar capacities of mutual understanding, learning and collaboration (Kaufman & Kaufman, 2015; Krupenye & Call, 2019; Premack & Woodruff, 1978; Price et al., 2017). All in all, social capacities and needs of humans seem so remarkable in the animal kingdom that they have been highlighted as a likely driver of humanity's unique evolution, including distinctive capacities of innovation and culture development (Corazza & Thienen, 2021; Dean et al., 2014; von Thienen et al., 2021). An influential early thought in this line of research and debate has been Dunbar's *Social Brain Hypothesis*. He argued that the size and complexity of the human brain increased to keep up with the increasing size and complexity of human social groups (Dunbar, 1998).

Even apart from questions of group size, social interactions require extensive cognitive capacities. Individual knowledge does not suffice, but mutual knowledge is needed: It is not enough for person *A* to know something. It is also vital to know whether the interaction partner *B* knows this as well, and in fact whether *B* knows that *A* knows it, and so forth. To merely lead a conversation in English, for example, both conversation partners need to know that the other person speaks English, and that the other person is familiar with the specific terminology used. Failing to keep in mind, the conversation partner's knowledge leads to miscommunication. This mutual knowledge that people need in order to interact successfully is often called their *common ground* (Lee, 2001).

However, knowledge and understanding are not the only social factors that have received extensive attention in studying the evolutionary development of humans in the animal kingdom. Frans de Waal highlights our emotional capacity for empathy and explains how this can motivate altruism. He argues that empathy coupled with increasing cognitive capacities among hominins advanced complex forms of social understanding that are characteristic for humans (Waal, 2008; Waal & Preston, 2017). In addition, some scholars describe a human "instinct" to socially align with each other. In particular when humans move together, they feel with each other and adjust their attitudes to the peers (Shamay-Tsoory et al., 2019).

In discussions of human interactions, one term that is often used is *empathy*. However, scholars have also noted that this term is "popular, yet fuzzy" (Hall et al., 2020, p. 1,), because it is used in many scientific disciplines and also in everyday life. Hall and colleagues asked laypeople for their definitions and intuitions, finding that empathy is a multifaceted concept for people, including strong variations of attributed meaning (Hall et al., 2021).

3 Two Routes of Social Understanding

To study any phenomenon, scientists typically begin by clarifying concepts, to specify more precisely what they intend to study. This approach has also been pursued in the study of empathy and social understanding.

Early on, a distinction has been proposed, according to which we understand others based on *affective* versus *cognitive* thought processes. Such a differentiation has already been prevalent in ancient philosophy. Plato, for instance, laid out an allegory of man as a charioteer. Drawing from his own reason and intellect, the charioteer tries to reach "the truth" or enlightenment. He does so by aiming to synchronise two passionate horses that pull the chariot, where one horse pursues moral impulses, while the other follows less noble appetites (Hackforth & Plato, 1972). The dualism

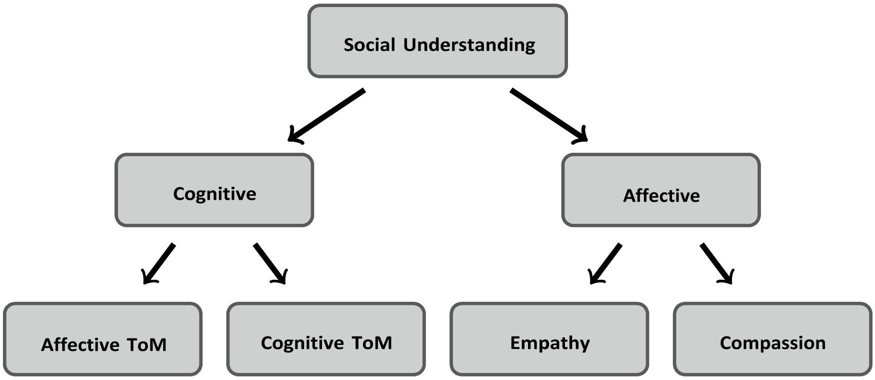


Fig. 2 Visualisation of different pathways towards "understanding others"—as studied in neuroscience

of reason (charioteer) versus affect (horses) was further advanced in philosophy over time, as reflected also in Descartes' dualistic account of human nature, and found its way into everyday parlance and world views. Nowadays, people often differentiate and focus on the contrast between "thinking versus feeling".

The dualism of reason versus affect has received a lot of criticism. It has even been called "Descartes error", with modern neuroscience indicating that reason and emotion are closely interlinked (Damasio & Descartes', 1994). Nonetheless, this distinction can be useful in terms of "more" versus "less" involvement of affect or reasoning across situations. Phenomenologically, this is well known to many people. There are situations in which social understanding is primarily a matter of affect, as when you comfort a friend who lost someone dear. In other cases, social understanding emerges primarily through reasoning, such as when a friend behaves unexpectedly and you find it hard to figure out the reason why your friend acted this way. Clearly, even in the first scenario, one does not only feel but also think, while it would also be unrealistic to consider the second scenario completely devoid of emotion. Therefore, neuroscience does not describe two unrelated systems. Instead, the affective and cognitive route to social understanding is considered to be active much at the same time and to interact in most real-life situations (Kanske et al., 2015a; Preckel et al., 2018; Shamay-Tsoory, 2011).[1]

Both routes group together several sub-processes. In neuroscientific research, special attention has been paid to *empathy* and *compassion* as part of the affective route to understanding others. On the cognitive side, *theory of mind* (*ToM*) has been introduced as a key concept (see Fig. 2).

Theory of mind is the process of drawing conclusions about another person's emotional state (affective ToM) or mental state (cognitive ToM) based on indicators.

[1] Sometimes "empathy" is used as the umbrella term for social understanding. The "two routes" are then referred to as cognitive versus affective/emotional empathy (Dvash & Shamay-Tsoory, 2014; Dziobek et al., 2008). However, this parlance has also been criticised as leading to confusion (Stietz et al., 2019).

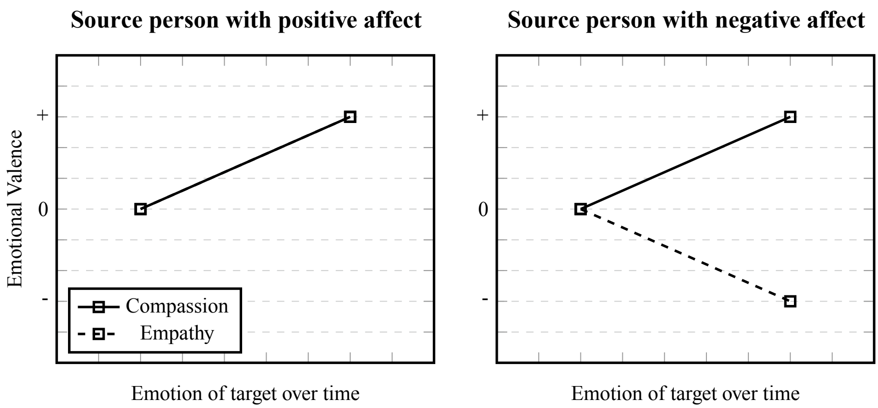


Fig. 3 Patterns of emotional reactions of a target person *B* to either positive (left side) or negative (right side) emotions of source person A. According to the definitions of compassion and empathy, person A and B have the same isomorphic affect in the case of positive source emotions. By contrast, when A has negative emotions and B reacts with an isomorphic, negative emotion, this classifies as empathy. When B reacts with positive feelings like warms and concern, it fulfils the definition of compassion

Conclusions in this realm can be more or less accurate or even incorrect. The other person may make an unhappy face and I believe she is unhappy, while in reality, she is just pretending. The other person may say she read the newspaper and I believe she knows the news of the day, but in reality, she read an old newspaper and does not know recent news at all.

Empathy and compassion are two different emotional reactions to another person. This way of understanding others cannot be described as correct versus incorrect or more versus less accurate. However, reactions can be more versus less appropriate or more versus less intensive. When another person is in pain and I feel happy, one might argue that I should feel discomfort in the light of the other person's suffering. Or maybe—when I am a doctor—I should better feel hardly any emotion at all, to be able to concentrate fully on helping the person without being distracted by emotions. In any case, my happiness or discomfort is not correct versus incorrect, but it may be (in)appropriate.

Notably, this neuroscientific distinction between theory of mind and empathy diverges from typical understandings in everyday parlance, where the term "empathy" is used in a much wider sense. However, clear definitions and fine-tuned concepts, as have emerged in neuroscientific research on social understanding, are essential to advancing scientific rigour (Hall et al., 2021).

Apart from real-life examples, philosophical and psychological theories, there has been another important line of research to inform frameworks of why and how we understand others. This research has looked at children, both "neurotypical" and "neurodiverging", and their evolving capacities of understanding others at different years of age.

Research indicates that some affective routes to understanding others develop early in life. One of the first phenomena that can be observed in infants is *emotional contagion*, where babies start crying when they hear another infant cry (Simner, 1971; Thompson, 1987). Researchers have argued that while this is not a case of "feeling with the other person", the behaviour is routed in sadness/discomfort, which starts with one infant and then passes on to others; so the emotion is contagious (Thompson, 1987). In recent years, this claim has been contested by researchers arguing that there is no indication that the emotion is relevant: subsequent infants could be crying because of the unpleasant sound of the first baby's crying, irrespective of the first baby's emotion (Ruffman et al., 2017). However, even when emotional contagion does not evolve in babyhood, there are many instances of preschool children showing some form of compassion towards other people (Trommsdorff et al., 2007).

In contrast, abilities of understanding others by cognitive means seem to develop later in life. A standard test builds on false beliefs (Baron-Cohen, 1985; Korkiakangas et al., 2016; Wimmer & Perner, 1983). In the test, children hear a story (either based on pictures or dolls) where there are two characters involved. One example is a story of Sally and Anne. Sally has a marble, which she places in a basket, with Anne watching. Then Sally leaves the room. While she is gone, Anne takes Sally's marble out of the basket and puts it into her own box. Then Sally re-enters the room. As a test question, the observing child is asked where Sally will look for her marble. To answer this question correctly, children need to be able to separate their own knowledge about where the marble actually is (in Anne's box) from Sally's false belief that the marble would still be in her basket. In the original study by Wimmer and Perner, most 4–5-year-old children failed the false belief test. They answered that Sally would look for her marble in Anne's box. By contrast, almost all of the 6–9-year-old children gave the correct answer that Sally would look for her marble in the basket where she had left it (Wimmer & Perner, 1983). This indicates that younger children are not able to differentiate between their own knowledge versus the (false) beliefs of others. With little variation, older children develop the cognitive capacity to track and disentangle knowledge states of different people.

In addition to elucidating cognitive developments in "neurotypical children", the Sally–Anne test is also used to elucidate impairments. In particular, children on the autism spectrum (which is characterised by persistent deficits in social communication and social interaction) seem to have problems with the kind of reasoning involved in this task (Association et al., 2013). Baron-Cohen and colleagues (1985) compared children and teenagers with autism spectrum disorder, between 6 and 17 years old, to healthy controls of 3 to 6-year-old children. In the false belief task, children on the autism spectrum showed marked impairments. By contrast, in studies on "affective routes to understanding others" children and adults with autism spectrum disorder do not seem impaired (Lockwood et al., 2013; Mazza et al., 2014).

The opposite pattern of impairment is found in people, who are colloquially termed as "psychopaths" or "sociopaths". DSM-V sets diagnostic criteria for antisocial personality disorder as a pattern of disregard for, and violation of, the rights of others (Association et al., 2013). Persons with this diagnosis show markedly reduced performance on tests of "understanding others affectively". At the same time, they show "normal performance" on tests of cognitive reasoning about others, like the Sally-Anne test (Lockwood et al., 2013).

This *double dissociation* of complementary impairments across different groups of mental disorders (autism versus antisocial personality disorder) provides another argument for the distinction of a cognitive versus an affective route to understanding others. Moreover, trends of selective functioning can also be observed in neurotypical people: Some persons seem to perform better on affective understanding, while others show higher performance on tests of cognitive understanding (Kanske et al., 2016).

Different routes to understanding others are often discussed to account for observable behaviour differences, e.g. between persons with different kinds of mental disorder or children of different age groups. In addition, the two routes are also discussed in relation to neural processing pathways (Kanske et al., 2015a; Preckel et al., 2018). Many meta-analyses have found distinct neural correlates to one versus the other way of social understanding (Jauniaux et al., 2019; Kogler et al., 2020; Schurz et al., 2014; Timmers et al., 2018). An especially insightful analysis was conducted by Schurz and colleagues in 2020, where they used hierarchical modelling (Schurz et al., 2020). This analysis included 188 studies that had used tasks on both routes to understanding others. The authors find that neural activations in these studies can be described best through a multilevel model with three groups of processes: cognitive, affective and intermediate. Tasks assigned to the cognitive cluster are false belief tasks (like the Sally–Anne test), trait judgements and strategic games. Tasks in the affective cluster include observing or sharing pain and other emotions, as well as tasks of reading emotions from the eye region of faces.[2] Tasks that have been grouped in the intermediate cluster involve more complex social and emotional stimuli, for example, social animations and emotions situated in context. The existence of the intermediate cluster does not necessarily counter the conceptual framework of two routes to understanding others. It can indicate how both routes work together in more complex situations.

4 Affective Route to Social Understanding

There are two main processes in the affective route to social understanding, which have received a lot of scientific attention: *empathy* and *compassion* (see Fig. 2). Both are affective reactions to another person's emotions, but they differ in one important detail.

In neuroscientific literature, *empathy* has been defined via four criteria. All four conditions need to be met, in order for an emotion to be classified as an "empathic reaction" (Vignemont & Singer, 2006): First, there needs to be an affective state in a target person *B*, for instance, sorrow. Second, this state is isomorphic to the

[2] Traditionally, the eye-reading test has been interpreted as a task of the cognitive route to understanding others, since it involves inferring emotions displayed in the eye region, where inferences can be right or wrong.

imagined or real affective state of person *A*. For instance, before *B* felt sorrow there was another person, *A*, who also felt sorrow. Third, *A*'s emotional state is the cause of *B*'s emotional state. For example, *B* feels sorrow because he imagined or observed *A*'s sorrow. Last, *B* has to be aware that her affective state was elicited by *A*'s affect. E.g. *B* feels sorrow and is aware that this sorrow is not caused by the rainy weather on that day, but because of an encounter with *A*, who had lost someone close and was therefore filled with sorrow. Simpler put, empathy occurs when a person is feeling *with* another person.

Compassion shares three of the four conditions with empathy: *B* is feeling something because of observing or imagining *A*'s affective state, and *B* is aware that *A*'s emotions were the source of her own affective state. However, the affective states of *A* and *B* can be of a different kind; they need not be isomorphic. Instead, compassion is always characterised by positive feelings, like warmth, concern or care de Vignemont and Singer (2006).[3] Therefore, compassion is often described as feeling *for* another person.

Depending on the affective state of source person *A*, compassion and empathy can lead to the same affective state in target person *B*. If *A* experiences positive emotions like happiness, then *B* will react with happiness, no matter whether *B* is being compassionate or empathic (see Fig. 3). The difference between the two kinds of reaction can only be observed when source person *A* feels something negative. Therefore, research has focused on negative emotions in source person *A* to allow for a clear differentiation of compassion versus empathy in target person *B*. However, it is important to note that compassion and empathy can (and sometimes do) occur at the same time. When consolidating a sad friend *A*, it is perfectly possible for *B* to experience some mixture of sadness and concern. Humans are capable of mixed and complex emotional states that seldom tick only one of the boxes that researchers attempt to separate. This does not make the distinction between compassion versus empathy less useful for research. Studies that have measured both compassion and empathy as a response to the same stimulus have shown that there are distinct neural correlates of both processes (Kanske et al., 2015b; Singer & Klimecki, 2014).

Potential consequences of empathy and compassion have received a lot of attention, both on the level of individuals and on the level of society at large. Notably, empathy is primarily discussed in critical terms in neuroscientific debates. This is very different from discussions in human-centred design, where empathy is treated as extremely favourable, indeed as a prerequisite for good design. However, the neuroscientific definition of empathy is highly specific. When human-centred designers speak of empathy, they typically address reactions that neuroscientists call compassion (cf. Sect. 9 below).

In the neuroscientific debate, it has been pointed out how empathy and compassion elicit fundamentally different dynamics when the emotions of source persons are negative (Singer & Klimecki, 2014). In that case, empathic reactions lead *B* to have negative feelings too, which is stressful. Thus, *B* experiences "empathic distress". If *B*

[3] Compassion is often also referred to as sympathy, for example in de Vignemont and Singer's paper (Vignemont & Singer, 2006).

experiences empathic distress too often, it poses a health risk and can lead to burnout. As a defence mechanism, *B* may try to withdraw from social contacts in the longer run. By contrast, in the case of compassion, emotional reactions are by definition of a positive and caring kind. This is assumed to be a predictor of good health and of continued pro-social behaviours. Against this background, neuroscientific debates have brought about calls for a development of methods to help people react with compassion instead of empathy when encountering people with negative affect.

Other dangers of empathy on the individual level can stem from either misdirected or excessive empathy. Breithaupt argues that empathy can even be selfish (Breithaupt, 2018). One example is the enjoyment of fiction, where consumers empathise with characters both in positive and negative situations. In fiction reading, the latter might even be more enjoyable (Hanich et al., 2014), as for instance feeling with the character's sadness can be cathartic (However, there is a difference between fictional versus real pain or sadness. Empathising with negative emotions of other people in real life is not cathartic for the empathiser.) In addition, empathy can go in "wrong directions", as in cases that Breithaupt calls "empathic vampirism". Here, person *B* empathises strongly with person *A*, but without concern for *A*'s well-being in real life. Examples can be obsessed fans, e.g. of a musician. The fan does not empathise with the musician as a real-life person, but with an idealised version of them. Another example is helicopter parents, who confuse their own well-being with that of their children, sacrificing their children's freedom in the process.

Other negative consequences of empathy are said to unfold on the societal level—so much so that empathy has been highlighted as an unsuitable ground for moral judgements and decisions (Prinz, 2011a, b). In particular, empathy tends to induce a "spotlight vision" (Bloom, 2017a). Thus, it can engender a bias, as people focus on specific individuals with whom they empathise, while neglecting larger groups, long-term solutions and "the bigger picture" (Bloom, 2017a). A pertinent example is the portrayal of refugees in the media, where dehumanising language has been used regularly, (Esses et al., 2013) and the distinction between migrants versus refugees was often disregarded.[4] However, some cases of single individuals have been highlighted and portrayed most favourably in the media. This holds for the case of a girl from Kosovo, Arigona Zogaj, whose family applied for asylum in Austria and was denied.[5] Arigona's classmates and the municipality started a petition to allow them to stay. Media coverage focused on Arigona after she ran away and publicly threatened to commit suicide if she was made to leave the country that she called her home. Many people empathised with her and campaigned for her right to stay, which she was granted officially in the end. However, nothing changed for migrants or refugees in general. The empathy was focused on her as an individual; it did not extend to other people in similar situations, even though suicide among refugees is

[4] UNESCO. Media and migration, covering the refugee crisis. https://en.unesco.org/ news/ media-and-migration-covering-refugee-crisis. Accessed: 2020-11-20

[5] Wikipedia. Asylfall familie zogaj. https://de.wikipedia.org/wiki/Asylfall_Familie_Zogaj/. Accessed: 2020-11-20

still a significant phenomenon.[6] Additionally, empathising with specific individuals or groups can entail strong negative feelings towards other groups. In extreme cases, this can induce sentiments of deep divides, even culminating in terrorist activities (Breithaupt, 2018).

While negative consequences of empathy have been discussed extensively in neuroscientific debates, positive effects should be mentioned as well. Feeling with another person is a strong bonding mechanism (Bloom, 2017b). It is hard to imagine intimate relationships without empathy (and compassion for that matter). Additionally, no emotion is inherently good or bad, functional or dysfunctional; it depends on the context and intensity (Zaki, 2017). In fact, in popular culture villains are often characterised by their lack of emotion and empathy. Examples of this are the Cybermen in the universe of "Doctor Who". These are humans stripped of all their emotions, who are then encased in robotic metal suits. Their lack of emotion—indeed their lack of capacity to feel with or for other people—coupled with their conviction to be the last step in evolution renders them horrifying enemies.

Moreover, empathy (and compassion) can be drivers of pro-social behaviour (Batson et al., 2015). In this sense, Pinker lays out how empathy—"to feel the pain of others"—can lead societies to become less violent (Pinker, 2011). For instance, when you watch someone being tortured and empathise with the other person's pain, it is unpleasant for you. Thus, you are motivated to help stop the torturing, in order to feel better yourself. Thus, empathy drives societies towards becoming more benevolent.

Considering all arguments, it seems both empathy and compassion can serve favourable purposes, as long as they are not misdirected or invoked excessively. This holds for individuals and societies at large.

5 Cognitive Route to Social Understanding

Neuroscientific discussions of the cognitive route to understanding others are very much in agreement about its benefits. The cognitive ability to understand the psychological state of another person based on indicators (facial expressions, behaviours, verbal statements…) allows us to "put ourselves in someone else's shoes", which is an extremely potent and useful skill to have. However, this skill can be used for ethical as well as less ethical purposes, as will become apparent below. In neuroscientific research, much attention has been paid to neurocognitive mechanisms that permit cognitive reasoning about the psychological state of others, and there are ongoing debates about respective capabilities in different animal species.

A key concept in the discussion of "cognitive routes to understanding others" is *theory of mind*. The concept was introduced by Premarck and Woodruff, who asked: "Does the chimpanzee have a theory of mind"? (Premack & Woodruff, 1978). Here, the term theory of mind means "that the individual imputes mental states to himself

[6] Salm, W., Blancke, D., & Maywröger, A. Suizid meldestelle. http://www.fairness-asyl.at/suizid-meldestelle/. Accessed: 2020-11-20

and to others" (p. 515) (Premack & Woodruff, 1978). The authors chose to speak of a "theory", because mental states are not directly observable but inferred based on indicators. Moreover, mental states function as variables to generate predictions, e.g. about what another individual will do next. To give an example, I might think you know that $2 + 2 = 4$, and I might believe you are motivated to answer maths questions to the best of your knowledge. Then I will predict your answer to be "4" when someone asks you "what is $2 + 2$?" By contrast, if I think you are in a mood for jokes and dislike overly easy maths questions, my prediction would likely be different.

The concept of a theory of mind interlinks with the notion of a "common ground" that individuals need to coordinate their social interaction (a mutual understanding of psychological states, cf. Sect. 2). Theory of mind is a mechanism, which serves to create and advance people's mutual ground. More precisely, it is a cognitive process, which uses various sources of information to infer the mental or affective state of another individual. The information used for the inference can be directly observed; for instance, because I have already heard you say $2 + 2 = 4$ before I attribute the respective knowledge to you. The information can also be indirectly observed, e.g. I heard your maths teacher say that you are very good at plus-tasks, so I assume this holds for $2 + 2 = 4$ as well. The information can also be imagined, e.g. I simply assume you have received regular maths education at school, even though I have not heard anyone address your school education explicitly.

In neuroscientific research, mental states that have received much attention are beliefs and intentions. By contrast, in human-centred design, people's "needs" and "values" are key concepts. They too are mental states. When product developers think about the needs and values of product users, they develop a theory of mind concerning the user. This theory can be right or wrong and it can be more or less accurate.

An overarching question in early neuroscientific studies was whether or not someone has a theory of mind—which suggests a dichotomous answer. The Sally–Anne test is constructed in such a way: according to expectation, the tested child will answer the false belief question either correctly or incorrectly. However, over time scholars began to undertake more gradual analyses. Papers came to focus on details of underlying mechanisms and elucidated a range of individual differences in neurotypical adults.

Regarding the explanation of how humans are able to arrive at theories of mind, debates are still ongoing. Two major lines of explanation have been pursued. They figure under the headlines of *theory theories* versus *simulation theories*.

According to explanations of the first kind, individuals literally *theorise* about the psychological states of others. A keyword in this tradition of explanation is "mentalistic folk psychology" (Gopnik & Wellman, 1992). According to this line of explanation, theories build up as interconnected systems, characterised by coherence. For instance, when I think you are smart and you know basic maths and I witness you answering the question of "$2 + 2 = ?$" incorrectly, I will try to re-establish coherence in my network of beliefs. I may add the novel belief that you are in a mood for jokes, or I assume you did not hear the question correctly. In the process of understanding others, the whole system of beliefs is used to predict the behaviour of someone else. Information is entered at one point, and after moving through the system of beliefs, it generates a conclusion, such as a prediction of what the other person will do next. If the prediction turns out to be incorrect, the belief system needs to be updated to restore coherence. This is similar to the way scientists develop theories of the word: They create theories to make predictions, test them, and in case predictions do not hold true, theoretical beliefs get amended or iterated (Gopnik & Wellman, 1992).

Another line of explanation is advanced under the headline of "simulation theories", which became popular upon the discovery of neurons that fire during observation and performing of an action (so-called *mirror neurons*) (Gallese et al., 1996; Pellegrino et al., 1992). Simulation theories assume people re-create or *simulate* mental and affective states of others by using overlapping activation networks in the brain. Thus, people (are said to) use an internal model of the other person's mind in order to understand this other person (Gallagher, 2001). Even though there are no neurons literally "mirroring" someone else in the human brain, there is ample evidence of overlap of general brain areas when observing versus imagining or experiencing something (cf. Sects. 7 and 8 below) (Buccino et al., 2001; Grèzes et al., 2003; Iacoboni et al., 1999).

Some researchers view simulation as a kind of narrative, which elucidates the other person's situation (Gallagher et al., 2006). Other scholars suggest that simulation is an automatic process of neural activation, which may even be unconscious. Yet, in all cases, it is assumed to facilitate the understanding of another individual (Gallagher, 2001). One such simulation theory was proposed by Gallese and Colleagues (2004). Their main claim is that the same neural circuits are active during first- and third-order experiences (cf. first-person vs. third-person stories). This, presumably, is the basis of an experiential understanding of other people, both affectively and cognitively (Gallese et al., 2004). Importantly, first-person experiences activate additional regions in the brain and body, allowing for a distinction between first-person versus third-person experiences. All in all, according to simulation theory, individuals do not only theorise about other people's psychological states, but simulations are run to achieve an experiential understanding of the other person (Gallese et al., 2004).

Even after decades of research, debates over different explanations for theory of mind endure. This might be due to the complexity of the phenomenon that is being studied. Moreover, explanations grouped together under the two headlines of "theory theory" versus "simulation theory" are rather diverse (Apperly, 2008). Thus, neuroscientific research in this field clearly has many more riddles to solve. In particular, the phenomenon of a "theory of mind" seems to cover multiple, diverse aspects, which may need to be studied with separate tests and methods (Schaafsma et al., 2015). Ultimately, great advances in neuroscientific understanding might ensue when different facets are understood by themselves and in their interaction.

At the same time, when reviewing earlier philosophical accounts, neuroscientists tend to agree on the importance of new work directions. Prominent thinkers like ancient Greek philosopher Aristotle and French philosopher Descartes believed that our knowledge of our own mental states is perfect: infallible, authoritative and privileged (Carruthers, 2013). However, these views are called into question and not only by real-life examples of people being confused about their own mental and affective states. Data from confabulation studies suggests that humans might use the same mechanism to understand the mental states of others and of themselves. (Carruthers speaks of an "interpretative sensory-access account" that humans use to understand psychological states, in their own case and that of others.). While the multiplicity of available information in understanding oneself leads to fewer errors, knowledge about one's own psychological state still seems to be imperfect and prone to errors, especially in situations where there is little or conflicting information available (Briñol & Petty, 2003; Wegner & Wheatley, 1999; Wells & Petty, 1980). Thus, in any case, neuroscientific studies into cognitive abilities of understanding psychological states highlight a range of empirical phenomena that research and theorising need to accommodate.

6 Tests and Questionnaires to Measure Capacities of Understanding Others

In this section, we review a number of tasks that are commonly used to measure the capacities of an individual to understand others. These tasks cover two major competency domains, namely affective and cognitive social understanding. This review can serve as a source of inspiration for human-centred design researchers: Such tasks can be used to assess individual differences of product developers, which may or may not predict specific qualities of design outcomes. Moreover, such tasks can be used to study the impact of education. For example, does design thinking training make product developers more empathic? How large is the impact on cognitive versus affective capacities of design thinkers, as they seek to understand users, teammates and other stakeholders?

One of the longest-established approaches for the measurement of social understanding in adults is a questionnaire developed by Davis, called the *Interpersonal Reactivity Index* (*IRI*) (Davis, 1980). It assesses social understanding on four sub-scales:

- *Fantasy*: tendencies of the test-taker to empathise with fictional characters (e.g. "I really get involved with the feelings of the characters in a novel".)
- *Perspective taking*: tendencies to adopt the perspective of other persons and use theory of mind (e.g. "When I'm upset at someone, I usually try to'put myself in his shoes' for a while".)
- *Empathic concern*: tendencies to feel compassionate towards others (e.g. "I often have tender, concerned feelings for people less fortunate than me".)
- *Personal distress*: tendencies to feel anxious and uncomfortable in emotional situations, regardless of whether alone or with others (e.g. "Being in a tense emotional situation scares me".)

It is important to note that this questionnaire intends to measure personality traits, which are presumed to be rather stable and independent of current contexts or situations. Thus, the scale is of limited use in experimental settings, where there are only short-term interventions. By contrast, long-term education effects might be investigated, when there is reason to believe that education has a lasting impact on the trainee's personality. It should also be noted that asking participants to rate their own social understanding bears a risk of answers being influenced by social desirability (King & Bruner, 2000). Nonetheless, for trait assessments, the *IRI* is a common measure that is often used. It is typically combined with other measurement approaches developed to assess (situation-specific) states of social understanding.

Another common measurement approach is the *Reading the Minds in the Eyes task (RMET)* by Baron-Cohen and colleagues, which has been developed to investigate theory of mind in adults (Baron-Cohen et al. 1997, 2001). Here, participants view pictures of faces with only the eye region visible. From a list of available options, test-takers are asked to choose which emotion best describes the affective state of the person in the image. To increase difficulty, complex emotional states and distractors are used. For instance, the image shows the eye region of a person expressing a feeling of seriousness. Answer options are relatively close in terms of typical facial expressions, such as "serious", "ashamed", "alarmed" or "bewildered" (Baron-Cohen et al., 2001). This task combines an emotional stimulus with a cognitive inference about someone else's mental state. Therefore, it has been deployed for both purposes, to measure affective and cognitive social understanding. A meta-analysis by Schurz and colleagues (discussed above) suggests that the test is better suited to assess affective social understanding (Schurz et al., 2020).

Tasks aimed at capturing primarily cognitive social understanding typically omit emotional stimuli. Examples include false belief tasks that are suitable for adults (Saxe & Kanwisher, 2003) and strategic games like the *Prisoner's Dilemma* (Kircher et al., 2009). However, many tasks that are assumed to assess cognitive social understanding often elicit brain activation that overlaps considerably with known affective networks of social understanding. A sample task, where this is the case, uses cartoons and asks participants to tell a story, drawing inferences about the fictional characters based on the available information (Schlaffke et al., 2015; Schurz et al., 2020). The fact that such tasks induce brain network activation associated with cognitive reasoning as well as affective understanding should not seem all too surprising. In real life, social situations are complex and people typically use different routes to social understanding in synergistic ways. Therefore, some authors argue that external validity is more important than a clear conceptional differentiation. This means that it would be more important that tasks measure people's real-life capacity of social understanding in valid ways, than to enforce disjunctive measures of affective versus cognitive social understanding. In the latter case, there is a danger of grasping only fragments of people's social understanding abilities, while under-representing the overall ability. One example of a test that strives to maximise external validity is the *Movie for the Assessment of Social Cognition* (*MASC*). Participants get to watch a short film. It is paused 46 times; and each time, participants are asked to answer questions about the mental states of the characters. This is similar to real-life situations, where we can use all kinds of information sources—affective and cognitive—to understand others (Dziobek et al., 2006).

Last but not least, some tasks attempt to measure and distinguish cognitive versus affective social understanding in one setting. The *Multifaceted Empathy Test* (*MET*) shows complex pictures and asks participants to draw inferences about the emotional state of the person in the picture. Moreover, participants are asked to rate their own feelings when viewing the image (empathy) and concern for the person in the picture (compassion) (Dziobek et al., 2008). Similarly, the *EmpaToM* shows short video clips of people telling a story (Kanske et al., 2015b). This story is followed by a question that either does or does not include an aspect of theory of mind. Moreover, the video clip is either emotionally charged or emotionally neutral. The valence of emotional stories is always negative. Here are translations of exemplary stories told in the videos and the correct answer for all four conditions from the original paper:

- *Neutral without ToM*: "Hm… well, … that evening, I cooked. I prepared one of those 3-course meals. And my boyfriend invited his sister. She brought a nice red wine from her vacation in France. And then there was my former flatmate. Well, I think we were sitting in the kitchen until one, one thirty."—"It is true that Anna's boyfriend was at the party".
- *Neutral ToM*: "My best friend Laura recently went to the movies with my brother. And she loves these cartoon movies, and my brother also thought the movie was great. He wants to watch another one of those with her right next week. That doesn't really sound like him… he used to be more into the action stuff."—"Anna thinks that her brother fell in love with her best friend and this is why he watches cartoon movies with her".
- *Emotional without ToM*: "We've been together for five years now, and it wasn't like we didn't like each other any more. But… at some point we just couldn't stop fighting. And once, he got so mad at me, he… hit me in the face. I just couldn't really go on after that."—"It is true that Anna met her ex-boyfriend at least five years ago".
- *Emotional ToM*: "My sister was diagnosed with bowel cancer a year ago and the odds aren't great. But you have to cling to something, don't you. Her doctor recently suggested a new treatment to her, but she refuses to try. It just makes me wanna cry."—"Anna thinks that her sister gave up hope and doesn't want treatment anymore".

After the clip, participants are asked to provide two ratings: how they feel altogether, and how much compassion they feel. Additionally, they are either asked a question that requires theory of mind or a factual reasoning question. While the measurement of compassion and theory of mind is similar to previous tasks, the measurement of empathy is computed indirectly, by calculating the difference between the participants' feelings after a neutral and an emotional video clip. The worse participants feel after the emotional clip, the higher their empathy score.

7 Physiological Indicators of Understanding Others

Alongside dedicated tests and questionnaires as outlined above, there is another approach to assessing people's social understanding. Here, the idea is to track the individual's psycho-physiological reactions in social situations. Research pursuing this approach has focused predominantly on the affective route of social understanding, or—in more general terms—it has used emotional stimuli. After all, there is a long history of theories highlighting the interplay of affective state and physiology (Barbalet, 1999; Cannon, 1927; Reisenzein, 1983). According to these theories, there is a marked physiological reaction when an individual experiences emotions. In the case of anxiety, typical physiological reactions include an increased heart rate and muscle tremor. Moreover, there is some quantitative correspondence: stronger emotions go along with stronger physiological reactions. Theories highlighting a close interplay between emotion and physiology are backed up by studies using various physiological assessments, including heart rate and skin conductance (Bradley & Lang, 2000; Khalfa et al., 2002; Nakahara et al., 2009). Thus, it should be possible to assess affective social understanding by confronting participants with the emotions of others and measure the participants' physiological responses. However, for researchers who would like to deploy such a measurement approach, it should be noted that physiological responses may be weaker than often expected (Mauss et al., 2005). Therefore, good measurement equipment and controlled measurement situations with well-chosen stimulus material and precise timestamps of stimulus onset or offset are important to obtain informative data.

Keeping these precautionary aspects in mind, physiological responses provide an interesting new perspective to social understanding. Electromyography is one approach that has been used to measure activity in facial muscles associated with emotion expressions. In a study by Achaibou and colleagues, participants watched videos in slow motion, where faces expressed increasing emotions; it was found that study participants automatically mimicked the facial expressions they saw in the video (Achaibou et al., 2008). Such mimicking does not necessarily indicate full-fledged affective social understanding, but at least study participants showed basic processes, which can be significant elements in affective social understanding. A follow-up study on adolescents combined measures of emotional mimicry (using electromyography while people watched emotional film clips) with ratings of state and trait social understanding (Graaff et al., 2016). To measure social understanding, participants were asked what emotion the person experienced that was shown in the film. Moreover, participants should indicate what kind of emotion they felt themselves after watching the video, and how intense their feeling was. In addition, participants were asked whether they felt sorry for the person in the video, and if yes, to explain why. Thus, the authors combined affective and cognitive aspects of social understanding. Trait social understanding was measured with the IRI (Davis & Franzoi, 1991). The authors found that state measures of affective social understanding correlate with emotional mimicry (Graaff et al., 2016).

Another physiological response that has been used in research on social understanding is skin conductance (*cf. electrodermal activity, EDA*). Skin conductance increases when people become aroused. The measure has a high time resolution, as skin reactions occur rapidly upon increased arousal. One traditional explanation for changes of skin conductance upon arousal refers to sweating processes, though present-day explanations are more complex. In research on social understanding, skin conductance is usually deployed to assess affective social understanding. However, interpreting and predicting skin response data in the context of empathy research is tricky, as a study by Deuter and colleagues shows. They used the Multifaceted Empathy Test (MET) while measuring facial muscle activity, skin conductance and heart rate to investigate the relationship between social understanding and these physiological parameters (Deuter et al., 2018; Dziobek et al., 2008). They hypothesised that participants would react with an increase of arousal when confronted with images of persons in strong emotional states. And this arousal should be accompanied by respectively strong physiological responses. However, in their study, no correlation was found between measures of heart rate and facial muscle activity. What was even more surprising, the authors found a negative correlation between skin conductance and affective social understanding: the higher participants scored in terms of affective social understanding, the lower their skin conductance (which roughly means, the lower was their sweating response) (Deuter et al., 2018). This highlights the importance of behavioural measurements when interpreting physiological responses.

To investigate the neural underpinning of social understanding, most researchers have used *functional Magnetic Resonance Imaging (fMRI)*, to locate parts of the brain uniquely activated during social understanding. Some studies have used *electroencephalography (EEG)* instead, which has a higher time resolution compared to the fMRI but a lower spatial resolution (Meinhardt et al., 2011; Zhouand & Han, 2021). Most tasks described in Sect. 6 to investigate cognitive social understanding have been used in the fMRI scanner. This multiplicity of studies has permitted excellent meta-analyses, shedding light on underlying neural correlates of theory of mind. The results suggest a core network that is activated in most theory of mind tasks. This core brain network includes the *medial prefrontal cortex* and the *bilateral posterior temporoparietal junction* (Schurz et al., 2014). Apart from the core network, studies yield different activation patterns depending on whether tasks explicitly ask participants to draw inferences about the mental state of others, versus tasks where inferences are implicitly required but not asked for (e.g. when reading a story without answering questions about the content). Influencing factors also include whether the stimuli are verbal or visual and whether participants are asked for inferences on affective or cognitive mental states (Molenberghs et al., 2016). Recently, some meta-analyses have combined studies on affective and cognitive social understanding to identify two respective core networks and investigate their differences (Kogler et al. 2020; Schurz et al., 2020).

Concerning the affective route of social understanding, there is one task that has been used extensively: *empathy for pain*. In 2004, Singer and colleagues published an experiment where volunteers either received painful stimulation themselves or knew that their loved ones would receive painful stimulation—all while in the fMRI scanner (Singer et al., 2004). The authors found a number of brain regions activated in both cases, in particular regions in the *anterior insula*, *anterior cingulate cortex*, *brainstem* and *cerebellum*. Moreover, activation in the *anterior insula* and the *anterior cingulate cortex* correlated with behavioural measurements of trait empathy (measured with the *IRI* and another empathy scale). The authors concluded that there is a shared neural network that becomes activated both when participants observe someone else in pain and when they actually feel pain themselves. This "revolutionary" and robust finding inspired many subsequent studies that would use pain to study empathy (often assessed with fMRI or EEG). To render the research less harmful, subsequent studies usually did not deliver actual painful stimulation (Bufalari et al., 2007; Gu & Han, 2007; Lamm et al., 2010). Instead, researchers used pictures or videos showing other people in pain—where this stimulus material is produced in Hollywood fashion without the people involved actually suffering. Notably, even with this fabricated material, the same brain activation patterns—reflecting people's empathy for pain—can be observed. Due to the robustness of the effect, studies have used the pain-paradigm to investigate empathy in detail, including the impact of mood, emotion regulation and many other processes (Naor et al., 2020; Qiao-Tasserit et al., 2017). Several excellent meta-analyses confirm a core network of empathy for pain comprising the *anterior insula* and *anterior cingulate cortex* (Jauniaux et al., 2019; Kogler et al., 2020; Lamm et al., 2011; Timmers et al., 2018). In an attempt to disentangle the function of these two brain regions in the context of empathy, Gu and colleagues asked participants to either make pain or laterality judgements of body parts in painful or neutral situations. Only the *anterior insula* was activated by painful situations regardless of the task (Gu et al., 2010). Thus, the authors suggest that the *anterior cingulate cortex* might not be specific to empathy, but rather might play a role in general information processing. By contrast, Singer and colleagues argue that the role of the *anterior insula* is to integrate different information including affective and sensory information, to produce a subjective feeling. According to this view, the *anterior insula* is vital for empathy (Singer et al., 2009).

8 Empathy Failure When the Other Person is "Not Like Me"

Finally, we shall review a number of studies using the empathy-for-pain research paradigm. Studies in this field can be specifically relevant for design thinkers and human-centred designers, because the research pinpoints unique difficulties that humans have when trying to understand others. A vast number of studies finds coherently that human empathy reactions are strongest when we are confronted with someone who is similar to us: the person belongs to the same social group, has the same ethnicity, etc. By contrast, when the other person is (perceived to be) markedly different from us, neural empathy reactions in the brain tend to be starkly reduced. This is an alarming pattern for human-centred designers. In effect, empathising with users is especially important when the user is markedly different from the designer. For example, the designer has proficient English-speaking skills and knows how to handle finances in the USA, while the intended users—as described in Bernard Roth's example from the d.school—belong to a group of immigrants from Mexico who hardly speak English and do not know pitfalls of financial matters in the USA. Or, in other cases, designers are IT experts in their twenties; they get to design digital banking solutions for elderly people who have little understanding or affinity towards IT technology. What pitfalls are to be expected when product developers try to empathise with persons who are very different from themselves?

One of the first studies in this line of research investigated the impact of "perceived fairness" on empathy for pain (Singer et al., 2006). Study participants were asked to play an economic game with a partner outside the fMRI scanner. This second person was actually a confederate of the experimenter, who either behaved fairly or unfairly in the game. Afterwards, participants witnessed these confederates receiving painful stimulation. When the confederate had behaved fairly in the economic game, study participants showed the typical brain activation pattern of people empathising with someone else in pain. However, when the confederate had behaved unfairly, only female study participants showed this kind of brain activation. By contrast, the neural empathy-for-pain reaction in male participants was markedly reduced. Additionally, men rated their desire for revenge towards the unfair confederates as higher compared to women. These study findings indicate two important dynamics: First, there is the possibility of gender differences in empathy-for-pain reactions. Second, empathy reactions can be different depending on how fair the other person behaves (Singer et al., 2006).

Similar effects were found when investigating the impact of group membership on empathy for pain (Hein et al., 2010). One influential study was conducted with soccer fans of two competing teams. Study participants either observed a fan of their own team (*ingroup*) or a fan of the competing team (*outgroup*) receive painful stimulation. Participants consistently reacted with compassion, empathy and a desire to help when an ingroup member received painful stimulation. However, when an outgroup member underwent the same kind of painful stimulation, participants reacted very differently. In that case, empathy for pain was reduced, and participants were more likely to choose to watch the outgroup member in pain instead of helping them. The authors conclude that affective social understanding can lead to pro-social motivation in some cases, especially when the other person is perceived as belonging to one's own social group. However, the authors also caution the audience: There can also be dynamics of enjoyment when outgroup members experience negative situations (Hein et al., 2010).

Subsequent studies have helped to clarify further details of how group membership impacts empathy reactions. Richins and colleagues worked with three groups: (i) students coming from the same university as the study participant (University of Exeter), (ii) students coming from another university, not in competition with the participant's university (University of Sussex) and (iii) students coming from a specific other university known to be rivalling with the participant's university

(Cardiff University) (Richins et al., 2019). Study participants got to see visual depictions of students from groups (i), (ii) and (iii), where the persons were either in pain or experienced a neutral situation. Participants showed the typical empathy-for-pain brain activation when seeing students out of groups (i) and (ii) in pain. However, participants showed no difference in brain reactions when they viewed the painful versus neutral situation of a person from group (iii). Thus, empathy for pain seems to be specifically reduced when outgroup members belong to a socially rivalling group.

Among the factors that people use—consciously or unconsciously—to estimate group membership, physiological factors indicating ethnicity seem to weigh in highly. In 2009, Xu and colleagues published the first paper to report a racial bias in empathy for pain (Xu et al., 2009). They measured brain responses in Chinese and Caucasian students in China. Participants watched video clips of either an Asian or a Caucasian model receiving painful versus neutral stimulation. Both Chinese and Caucasian study participants showed diminished neural response when the painful stimulation was delivered to a person of "the other-race". In particular, there was less reaction in the *anterior cingulate cortex*, one of the key regions involved in empathy for pain. Such study findings have been replicated multiple times, with different ethnicity groups coming from Italy, South Africa, Australia and the USA (Azevedo et al., 2013; Contreras-Huerta et al., 2013; Fourie et al., 2017; Mathur et al., 2010).

Notably, in the original study by Xu and colleagues as well as in many other studies, a marked dis-alignment is found between the neural reactions of study participants and their overt ratings of the stimuli. When asked to rate painfulness of the situation and self-unpleasantness, participants do not answer differently depending on ethnicity of the protagonist shown in the images or videos (Azevedo et al., 2013; Sheng & Han, 2012; Xu et al., 2009). Thus, participants may recognise that varying levels of empathy depending on "race" would be socially inappropriate, and they may align ratings with social norms of what one should be feeling when seeing others in pain. An indication of this is the positive correlation between implicit racial bias and increased neural empathic response for "own-race" compared to "other-race" pain found in some studies (Azevedo et al., 2013).

An interesting follow-up question is how the bias favouring persons of one's own ethnicity is biologically anchored in the brain. An EEG study on empathy for pain found that the *N170* is reduced when participants process painful facial expressions of "other-race" models. The *N170* is an oscillation measureable 170 ms after stimulus onset (i.e. after the picture of a person in pain appears on the screen). In terms of cognitive functions, the *N170* is associated with early on facial processing. This suggests that not only affective processing is reduced in case of "other-race" models. Even prior to affective processing, the mere perceptual processing of the other person's face seems to be reduced when the person is of another ethnicity (Sheng et al., 2017).

In ethical terms, it would seem to be desirable that human brains react with the same kind of thorough processing and empathy-propensity, no matter whose face is shown. Yet, the question about the biological reality remains. Why do human brains process faces of others differently, depending on physical ethnicity- cues? Some authors have proposed evolutionary explanations. In humanity's prehistory, it may have been specifically important to process faces of those persons in detail, who were closely around—ingroup members, usually all of the same ethnicity (Han, 2018). It has also been noted that the precise processing of other people's faces is computationally costly. The same can then be said of empathy reactions building on precise facial processing. In this sense, studies have been conducted where the processing capacities of study participants were limited, due to experimental interventions. Alongside seeing others in pain, participants had to tackle memory tasks, so that they had a high cognitive load. Under these circumstances, empathy reactions were markedly reduced (Morelli & Lieberman, 2013).

Does this mean that our human brain is inherently biased, prejudiced and discriminatory? Studies offer an outlook that may be interpreted alongside more optimistic lines, for a number of reasons.

To begin with, the studies discussed above reveal a bias on the neural level, part of which is not fully conscious and is hard to control. But this bias in processing does not always lead to a bias in behaviour. Ultimately, our concern will be that others are treated fairly. Early neural responses by themselves do not suffice to predict complex social behaviours and attitudes. Interestingly, one study showed higher overt ratings of empathy for "same-race" models among African Americans but not among Caucasian Americans (Mathur et al., 2010). However, both groups showed differential activation on a neural level when comparing same-race and other-race stimuli. Similar results were found in South Africa (Fourie et al., 2017). This pattern seems to reflect that most contemporary cultures oppose racial discrimination leading to the privileged ethnicity counteracting their implicit biases (Han, 2018).

Moreover, racial bias is not an unalterable processing mechanism in the human brain. There are ways to reduce or even eliminate this bias. Several effective means have been identified, in particular (i) exposure to people of different groups/ethnicities, (ii) guiding attention towards the individual person instead of the person-as-a-group member, (iii) working in mixed-ethnicity groups and (iv) dedicated empathy-compassion training.

Firstly, studies have compared neural empathy-for-pain reactions depending on how much exposure study participants have had with members of another ethnicity. The studies find that mere exposure to other ethnicities reduces the processing bias. For instance, Chinese people brought up in countries with a Caucasian minority show the same neural empathy-for-pain response towards Asian and Caucasian models (Zuo & Han, 2013). Another study assessed Chinese students living in Australia. The longer they had lived in Australia (i.e. the longer they were exposed to Australian faces), the smaller was their neural bias favouring Chinese over Australian faces in the empathy-for-pain paradigm (Cao et al., 2015). Importantly, in this study, the quantity of contact with Australians correlated with neural empathy measures: Longer exposure to Australian faces predicted smaller differences between the neural empathy reactions to Australian (other-race) compared to Chinese (same-race) people. By contrast, the quality of contact—to what extent the Chinese study participants liked their experiences with Australians—had no such impact on the neural processing.

Secondly, researchers have guided the attention of study participants to the painful experiences of others shown in pictures and videos. This directed attention also reduced racial biases in the empathy-for-pain paradigm, as has been demonstrated with fMRI and EEG (Sheng & Han, 2012; Sheng et al., 2014).

Thirdly, studies found that work experiences in heterogeneous groups reduced the bias as well. In one study, Chinese students were assigned to mixed-ethnicity groups for a competitive game; after learning who was on their team, they showed significantly reduced neural racial bias as measured by EEG (Sheng & Han, 2012).

Finally, social understanding can be improved by training. Several studies have used a special kind of meditation, which focuses on compassion and empathy, to increase affective social understanding (Mascaro et al., 2013; Trautwein et al., 2019).

Together, these studies suggest that there are ways to increase social understanding and overcome neurocognitive biases that favour empathy towards people who are similar to us. Moreover, teaching and communicating values of compassion and empathy can have a positive impact on the way we approach others.

9 Implications for Human-Centred Design

In this last section, we discuss how neuroscientific concepts relate to design thinking culture, and what can be learned for human-centred design practice or research.

To recapitulate, empathy is a crucial topic in design thinking. Students are trained to empathise with potential users of the innovative products to be developed. This serves to ensure that the emerging products are tailored directly to the user's core needs. In order for that to be the case, design thinkers need to be able to distinguish between their own personal needs, knowledge, skills and values versus those of the users'.

In parallel, neuroscience has become very interested in topics of empathy and how humans understand each other. Neuroscientific research yields a rich set of resources for design thinkers and human-centred designers. This allows for more scientific rigour in the analysis of what designers actually do when they empathise with users, and how this informs eventual product outcomes. For instance, one design thinker may be extremely good at empathising with users, while another design thinker may actually find it difficult to put himself in someone else's shoes. How does this personal capacity impact creative processes and work outcomes? Is there a way to build up personal capacities by means of empathy training, in order to help design thinkers empathise with users and to develop products that users really need? Moreover, neuroscience has advanced a number of helpful tools to assess multiple sub-capabilities that humans use to understand others. These measurement approaches can be very serviceable in design thinking research too, to advance an increasingly sophisticated understanding of the role empathy plays in the development of innovation.

However, it is a well-known phenomenon that neuroscientific attempts to elucidate psychological concepts (like "empathy") based on physiological studies lead to radical re-definitions of the original psychological terms (Bennett & Hacker, 2005; Thienen & Kausalniveaus, 2017). So, in order to apply neuroscientific work results properly in the context of design thinking and human-centred design, it is important to clarify what each side means with their key concepts.

What design thinkers call "empathy" is called "compassion" in neuroscience. In design thinking, the term "empathy" is used in a very broad and positive sense. Already one of the founding fathers of design thinking, John Arnold, elaborated on this point. Similar to empathy-for-pain studies, where participants are confronted with the unfortunate condition of others, according to John Arnold, some unfortunate condition of others or oneself is the typical starting point of innovation projects. One example he gives concerns the area of mobility. Due to traffic accidents, traffic deaths and traffic jams the situation at the outset can be experienced as "bad". However, the creative person uses experiences of problems as sources of inspiration, e.g. to imagine new ways how mobility might be re-designed in the future, so that traffic solutions would become safer and more enjoyable. This contrasts to persons with less of a creative mindset. They would either remain in their negative sentiments would try to accept situations as they were, or they might try to look away from discomforting realities to not feel bad about them. Arnold raises the question himself as to whether creative persons might experience "discontent, frustration and unhappiness" (p. 63) regularly, given that something dissatisfying and problematic in life is seen as the typical starting point of innovation projects (Arnold, 2016). Arnold answers that creative minds do focus on problems, yet not with negative sentiments, but in a solely positive spirit.

> An attitude of healthy skepticism in place of complacent acceptance is essential to the creative personality. The highly imaginative person is one who [. . .] is constantly asking himself how he can improve the things he sees. He is concerned with how the basic needs of man can be better satisfied. If this is discontent, then part of the question must be answered in the affirmative. I feel, however, that the word discontent connotes a rather definite negative quality and, therefore, should not be used. The spirit of the innovator is wholly positive. (Arnold, 1959/2016, p. 63)

In neuroscientific jargon, the term "empathy" has come to mean that the empathiser experiences the same emotion that another person experiences. In design thinking contexts, such a usage of the term would imply that design thinkers feel unhappy basically in all of their innovation projects. After all, design thinkers work until they find the "pain points" of potential users: areas in life where core needs of the users are not met, so that the users face major problems they are unhappy with. According to neuroscientific parlance, empathising with users in such situations would mean that design thinkers feel unhappy too. Moreover, neuroscientific theories would predict that design thinkers withdraw from social engagement and innovation projects sooner rather than later, to experience respective unhappiness less often.

By contrast, neuroscientists introduce the term "compassion" to characterise socially understanding persons, who acknowledge someone else's negative situation and react with complementary positive sentiments of concern and care, aiming to alleviate the other person's situation. Such a positive reaction to other people's problems is exactly the phenomenon described and sought out in design thinking. Thus, in neuroscientific parlance design thinkers do not empathise with users but develop compassion with them. Moreover, in design thinking education the primary, sensible teaching aim would seem to be the cultivation of compassion—and not empathy—in students.

Should design thinkers change their process models and parlance? One way to look at this is to simply acknowledge vocabulary differences. What design thinkers call "empathy" is called "compassion" in neuroscience. Another line of thought can ask about further details of design thinking education and empirical relationships. What role do negative emotions of the users play in innovation projects? Does more user unhappiness regarding their initial problem imply greater prospects of innovation success later on? How precisely do emotional and cognitive reactions of the innovators inform innovation projects? For instance, is it better when innovators react with a pure spectrum of positive sentiments (neuroscientific compassion) to the users' unhappiness, or might it be helpful to have some moments of shared unhappiness (neuroscientific empathy), as to acknowledge the problem thoroughly before embarking on a search for solutions?

Understanding oneself is a matter of interpretation, as is understanding others. Another aspect of neuroscientific theorising may be endorsed more straightforwardly in design thinking. Neuroscience has found that humans use very much the same brain regions when experiencing emotions first hand versus when understanding the same kinds of emotions in others. All in all, there is data to indicate that the processes of understanding oneself and understanding others are rather similar (while the person may have a broader and deeper spectrum of information available in her own case, e.g. concerning her thoughts and body sensations). This view contrasts to some earlier philosophical models, which suggested that individuals have a perfect and infallible first-person access to their own emotions and further psychological states.

In this regard, the design thinking view of social understanding is very close to the neuroscientific view. Design thinking methods for understanding the needs of users highlight that users usually cannot tell by themselves what their key unmet needs are. These needs must be identified by means of interpretation. For instance, in the method of *Empathy Maps* design thinkers analyse and interpret user interviews and field observations by reviewing what people said and did, besides reflecting on what people might have thought and felt. This is what design thinkers spell out when they create an *Empathy Map*:

> SAY: What are some quotes and defining words your user said?
>
> DO: What actions and behaviors did you notice?
>
> THINK: What might your user be thinking? What does this tell you about his or her beliefs?
> FEEL: What emotions might your subject be feeling?
>
> Note that thoughts/beliefs and feelings/emotions cannot be observed directly. They must be *inferred* by paying careful attention to various clues. Pay attention to body language, tone, and choice of words. [. . .] One way to identify the seeds of insights [about user needs] is to capture "tensions" and "contradictions" as you work. (d.school, 2010, p. 15, our emphasis)

Thus, users are not seen as possessing infallible self-knowledge about their beliefs, values and unmet needs. The work strategy is not to ask users directly about these things. Instead, the method encourages and trains design thinkers how to make careful observations, interpret clues and draw conclusions. This is very much in line with the ways in which neuroscientists describe processes of understanding others.

Moreover, the neuroscientific discussion of large-scale similarities between processes of "understanding others" and "understanding oneself" is also interesting for design thinking in another direction. As mentioned in Sect. 1, two different approaches have emerged in design thinking education. The classic and more common approach is to train students in empathising with users, so as to start creative projects. In some other design thinking classes (specifically those concerned with neurodesign), students learn instead to start creative projects by exploring their own passions and needs. In light of neuroscientific research, the two approaches might be more similar than is often assumed. At least, both approaches may stimulate somewhat similar brain activities of cognitive and affective processing.

Many design thinking empathy methods encourage theory of mind. Design thinking methods very often invoke theory of mind strategies to help innovators understand users. *Empathy Maps* are a good example. This method calls for cognitive reasoning, an interpretation of clues and the detection of inconsistencies. The same can be said about many empathy methods used in design thinking. For instance, the method of *Interviews for Empathy* provides similar instructions:

> **Look for inconsistencies.** Sometimes what people say and what they do are different. These inconsistencies often hide interesting insights.
>
> **Pay attention to nonverbal cues.** Be aware of body language [. . .]. (d.school, 2010, p. 10, emphasis in original)

Methodologies like this encourage rather analytic means of cognitive reasoning in order to understand other people. Of course, in the overall design thinking portfolio, using emotions in the realm of empathy work and when testing prototypes is encouraged. However, respective methodologies seem less well elaborated. Thus, neuroscientific differentiation can inspire further developments in design thinking methodology, so as to become even more systematic and focused. According to neuroscientific analyses, there are two routes to understanding others, based on (i) cognitive reasoning: theory of mind versus (ii) affect (see Fig. 2 above).

What methodologies might be available to teach the use of affect as a route to understanding others in design thinking? As is often the case, psychotherapy has ample methods in place that can easily be adopted for design thinking purposes (Thienen et al., 2012). These include dedicated methods on how people (psychotherapists/design thinkers) can use their own emotional reactions in response to others (patients/users) as means to a better understanding of the other person.[7] In addition, neuroscientific studies have successfully deployed meditations focused on empathy and compassion towards other people (Mascaro et al., 2013; Trautwein et al., 2019). Such meditations could also be adapted for design thinking purposes.

[7] In particular, the psychotherapeutic use of *countertransference* and related approaches of *plan analysis* would seem to be good candidates for methods that teach a use of affect as a route to understanding others.

Neuroscientific measurement approaches can be used to test empirical claims about the role of empathy in innovation projects. Empathy with users in creative processes is a hallmark of design thinking. A range of hypotheses are implied in design thinking theory. They include predictions of products becoming more innovative, more user-friendly and more economically successful in the case of sound empathy work and a good understanding of user needs. To test such empirical claims, scientifically sound methods are needed to assess just how well designers understand users (prior to achieving this or that kind of an innovation outcome, success or failure). Based on neuroscientific test batteries, it becomes feasible to elucidate empirical relationships in details, including individual strengths and weaknesses. For instance, one design thinker may be specifically good at an affective understanding of others, whereas another design thinker may be better at cognitive understanding. How do such individual strengths and weaknesses impact the innovation projects people engage in?

Moreover, design thinking trainings convey a range of methods to help design thinking students understand users. Can the impact of these trainings be measured in scientific terms, so that students get better in quantifiable ways at core processes of understanding others? What impact does each method have on the students' capacities of understanding users?

Overall, with its established test batteries and physiological indicators, neuroscience provides valuable resources for ever more sophisticated design thinking studies on the role of empathy in innovation projects.

Empathising with others seems specifically challenging for the human brain when the other person is perceived to be "different from me". This finding has been replicated by a range of neuroscientific studies. It points to a challenge that is uniquely pertinent to approaches of design thinking and human-centred design. After all, product developers need to empathise with users, and this is specifically important when the difference between product developers and users is large. In that case, product developers cannot use their own knowledge and needs as a point of reference to design a good product, but the point of view of users must be thoroughly understood.

This challenge remains a potential hot-spot for subsequent design thinking research. Maybe expert design thinkers have found ways to facilitate empathy with others, regardless of how different these other persons may seem upon first impression. Then it would seem interesting what methods or cognitive strategies expert design thinkers use in order to facilitate their own empathic reactions.

Moreover, neuroscientific research has suggested some strategies that may help design thinkers understand users when there are major differences in the social groups to which people belong (cf. Sect. 8). One of these strategies is to focus on people as individuals and not as group members. This approach is reminiscent of the already established design thinking method of creating *Personas* (Nielsen, 2013). Here, design thinkers learn to focus their attention on an individual (a real or fictitious character), instead of imagining a larger user group, in order to facilitate empathy and promote user-centred solutions. Another way in which design thinking methodology could facilitate the understanding of radically different others is to search for their "basic needs". It has long been noted that people may seem very different from each other in terms of their transient and culture-specific needs or concerns (Maslow, 1943). By contrast, in terms of their basic needs, all people seem very much the same, needing to eat, drink, breath and sleep to stay healthy, longing for safety and social acceptance, etc. Thus, the design thinking concern for people's basic needs (and method such as *Why-How-Laddering* (d.school, 2010) to carve out basic needs) may be another important strategy to help innovators understand diverse others. In this sense, design thinking may already possess a number of approaches to facilitate the understanding of users that appear to be very different from product developers. However, the relative contribution of different methods in this regard needs to be further clarified in future research and it might be helpful to add more tailor-made methods to facilitate empathy in situations that antagonise intuitive understandings.

In the meantime, being mindful of potential biases and pitfalls can be an important first step in design thinking and human-centred design: Our abilities to understand users may be specifically dependent on how similar or different they appear to be from us. When we perceive others to be very different from us, we may need to caution ourselves to better think twice and invest extra effort in order to ensure a thorough understanding of the other person's knowledge, values and needs.

All in all, neuroscientific research has engendered ample resources that can be serviceable for design thinking and human-centred design, to reflect on the role of empathy in the context of innovation projects in ever more sophisticated ways.

References

Achaibou, A., Pourtois, G., Schwartz, S., & Vuilleumier, P. (2008). Simultaneous recording of EEG and facial muscle reactions during spontaneous emotional mimicry. *Neuropsychologia, 46*(4), 1104–1113.

Apperly, I. A. (2008). Beyond simulation–theory and theory–theory: Why social cognitive neuroscience should use its own concepts to study "theory of mind." *Cognition, 107*(1), 266–283.

Arnold, J. E. (2016). Creative engineering. In *Creative engineering: Promoting innovation by thinking differently* (pp. 59–150).

Association, A. P., et al. (2013). *Diagnostic and statistical manual of mental disorders* (*DSM-5*®). American Psychiatric Pub.

Azevedo, R. T., Macaluso, E., Avenanti, A., Santangelo, V., Cazzato, V., & Aglioti, S. M. (2013). Their pain is not our pain: Brain and autonomic correlates of empathic resonance with the pain of same and different race individuals. *Human Brain Mapping, 34*(12), 3168–3181.

Barbalet, J. M. (1999). William james' theory of emotions: Filling in the picture. *Journal for the Theory of Social Behaviour, 29*(3), 251–266.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or asperger syndrome. *Journal of Child psychology and Psychiatry, 38*(7), 813–822.

Baron-Cohen, S., Leslie, A. M., Frith, U., et al. (1985). Does the autistic child have a "theory of mind." *Cognition, 21*(1), 37–46.

Baron-Cohen, S., Wheelwright, S. J., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(2), 241–251.

Batson, C. D., Lishner, D. A., & Stocks, E. L. (2015). The empathy–altruism hypothesis. In D. A. Schroeder & W. G. Graziano, (Eds.), *The oxford handbook of prosocial behavior*, chapter 13 (pp. 259–268). Oxford University Press.

Bennett, M. R., & Hacker, P. M. S. (2005). *Philosophical foundations of neuroscience*. Blackwell.

Bloom, P. (2017a). Empathy and its discontents. *Trends in Cognitive Sciences, 21*(1), 24–31.

Bloom, P. (2017). Empathy, schmempathy: Response to zaki. *Trends in Cognitive Sciences.*

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. *Cognitive Neuroscience of Emotion*, *25*, 49–59.

Breithaupt, F. (2018). The bad things we do because of empathy. *Interdisciplinary Science Reviews, 43*(2), 166–174.

Briñol, P., & Petty, R. E. (2003). Overt head movements and persuasion: A self-validation analysis. *Journal of Personality and Social Psychology, 84*(6), 1123.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13*(2), 400–404.

Bufalari, I., Aprile, T., Avenanti, A., Di Russo, F., & Aglioti, S. M. (2007). Empathy for pain and touch in the human somatosensory cortex. *Cerebral Cortex, 17*(11), 2553–2561.

Cannon, W. B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *The American Journal of Psychology, 39*(1/4), 106–124.

Cao, Y., Contreras-Huerta, L. S., McFadyen, J., & Cunnington, R. (2015). Racial bias in neural response to others' pain is reduced with other-race contact. *Cortex, 70*, 68–78.

Carruthers, P. (2013). Mindreading the self. In S. Baron-Cohen, H. Tager-Flusberg & M. V. Lombardo, (Eds.), *Understanding other mind: Perspectives from developmental social neuroscience*, chapter 26, (3rd edn., pp. 467–486). Oxford University Press, UK.

Contreras-Huerta, L. S., Baker, K. S., Reynolds, K. J., Batalha, L., & Cunnington, R. (2013). Racial bias in neural empathic responses to pain. *PLoS ONE, 8*(12), e84001.

Corazza, G. E., & von Thienen, J. P. A. (2021). *Invention*. Springer.

Damasio, A. R. (1994). *Descartes' error. Emotion, Reason, and the Human Brain.* Putnam Publishing.

Davis, M. H. (1980). A mulitdimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85.

Davis, M. H., & Franzoi, S. L. (1991). Stability and change in adolescent self-consciousness and empathy. *Journal of Research in Personality, 25*(1), 70–87.

de Vignemont, F., & Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences, 10*(10), 435–441.

De Waal, F. B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59*, 279–300.

De Waal, F. B., & Preston, S. D. (2017). Mammalian empathy: Behavioural manifestations and neural basis. *Nature Reviews Neuroscience, 18*(8), 498–509.

Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E., & Kendal, R. L. (2014). Human cumulative culture: A comparative perspective. *Biological Reviews, 89*(2), 284–301.

Deuter, C. E., Nowacki, J., Wingenfeld, K., Kuehl, L. K., Finke, J. B., Dziobek, I., & Otte, C. (2018). The role of physiological arousal for self-reported emotional empathy. *Autonomic Neuroscience,214,* 9–14.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91*(1), 176–180.

d.school. (2010). *Bootcamp Bootleg*. Institute of Design at Stanford.

Dunbar, R. I. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews, 6*(5), 178–190.

Dvash, J., & Shamay-Tsoory, S. G. (2014). Theory of mind and empathy as multidimensional constructs: Neurological foundations. *Topics in Language Disorders, 34*(4), 282–295.

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., Kessler, J., Woike, J. K., Wolf, O. T., & Convit, A. (2006). Introducing MASC: A movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders*, *36*(5), 623–636.

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the multifaceted empathy test (MET). *Journal of Autism and Developmental Disorders, 38*(3), 464–473.

el Zein, M. (2019). Shared responsibility in collective decisions. https://www.tele-task.de/lecture/video/7776/.

Esses, V. M., Medianu, S., & Lawson, A. S. (2013). Uncertainty, threat, and the role of the media in promoting the dehumanization of immigrants and refugees. *Journal of Social Issues, 69*(3), 518–536.

Fourie, M. M., Stein, D. J., Solms, M., Gobodo-Madikizela, P., & Decety, J. (2017). Empathy and moral emotions in post-apartheid South Africa: An fMRI investigation. *Social Cognitive and Affective Neuroscience, 12*(6), 881–892.

Gallagher, S. (2001). The practice of mind. Theory, simulation or primary interaction? *Journal of Consciousness Studies*, *8*(5–6), 83–108.

Gallagher, S., et al. (2006). The narrative alternative to theory of mind. *Radical enactivism: Intentionality, phenomenology, and narrative*, (pp. 223–229).

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*(2), 593–609.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8*(9), 396–403.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*,*7*(1 & 2).

Grèzes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: An fMRI study. *NeuroImage, 18*(4), 928–937.

Gu, X., & Han, S. (2007). Attention and reality constraints on the neural processes of empathy for pain. *NeuroImage, 36*(1), 256–267.

Gu, X., Liu, X., Guise, K. G., Naidich, T. P., Hof, P. R., & Fan, J. (2010). Functional dissociation of the frontoinsular and anterior cingulate cortices in empathy for pain. *Journal of Neuroscience, 30*(10), 3739–3744.

Hackforth, R. (1972). *Plato: Phaedrus*. Number 119. Cambridge University Press.

Hall, J. A., Schwartz, R., & Duong, F. (2021). How do laypeople define empathy? *The Journal of Social Psychology, 161*(1), 5–24.

Han, S. (2018). Neurocognitive basis of racial ingroup bias in empathy. *Trends in Cognitive Sciences, 22*(5), 400–421.

Hanich, J., Wagner, V., Shah, M., Jacobsen, T., & Menninghaus, W. (2014). Why we like to watch sad films. The pleasure of being moved in aesthetic experiences. *Psychology of Aesthetics, Creativity, and the Arts*,*8*(2), 130.

PI Julia von Thienen. (2019). Hasso Plattner Design Thinking Research Program. Innovation modelling based on human needs. https://hpi.de/en/dtrp/projects/projects-201920/innovation-modelling-based-on-human-needs.html.

PI Julia von Thienen. (2020). Hasso Plattner Design Thinking Research Program. Probing different concepts of human needs. https://hpi.de/en/dtrp/projects/projects-2020-21/probing-different-concepts-of-human-needs.html.

Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron, 68*(1), 149–160.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*(5449), 2526–2528.

Jauniaux, J., Khatibi, A., Rainville, P., & Jackson, P. L. (2019). A meta-analysis of neuroimaging studies on pain empathy: Investigating the role of visual information and observers' perspective. *Social Cognitive and Affective Neuroscience, 14*(8), 789–813.

Jowett, B. (1985). *The politics of Aristotle*. (Vol. 1). Clarendon.

Kaltwasser, L. (2019). *We feel therefore we are? About emotions and cooperation*. https://www.tele-task.de/lecture/video/7863/.

Kanske, P., Böckler, A., & Singer, T. (2015). Models, mechanisms and moderators dissociating empathy and theory of mind. In *Social behavior from rodents to humans*, volume 30 of *current topics in behavioral neurosciences*, (pp. 193–206). Springer, Cam.

Kanske, P., Böckler, A., Trautwein, F. M., Lesemann, F. H., & Singer, T. (2016). A restrong empathizers better mentalizers? Evidence for independence and interaction between the routes of social cognition. *Social Cognitive and Affective Neuroscience,11*(9), 1383–1392.

Kanske, P., Böckler, A., Trautwein, F. M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and Theory of Mind. *NeuroImage, 122*, 6–19.

Kaufman, A. B., & Kaufman, J. C. (2015). *Animal creativity and innovation.* Academic Press.

Khalfa, S., Isabelle, P., Jean-Pierre, B., & Manon, R. (2002). Event-related skin conductance responses to musical emotions in humans. *Neuroscience Letters,328*(2), 145–149.

King, M. F., & Bruner, G. C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology and Marketing, 17*(2), 79–103.

Kircher, T., Blümel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., van Os, J., & Krach, S. (2009). Online mentalising investigated with functional MRI. *Neuroscience Letters, 454*(3), 176–181.

Kogler, L., Müller, V. I., Werminghausen, E., Eickhoff, S. B., & Derntl, B. (2020). Do I feel or do I know? Neuroimaging meta-analyses on the multiple facets of empathy. *Cortex*.

Köppen, E. (2017). *Empathy by Design. Untersuchung einer empathie-geleiteten Reorganisation der Arbeitsweise.* BoD–Books on Demand.

Korkiakangas, T., Dindar, K., Laitila, A., & Kärnä, E. (2016). The Sally-Anne test: An interactional analysis of a dyadic assessment. *International Journal of Language & Communication Disorders, 51*(6), 685–702.

Krupenye, C., & Call, J. (2019). Theory of mind in animals: Current and future directions. *Wiley Interdisciplinary Reviews: Cognitive Science, 10*(6), 1–25.

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage, 54*(3), 2492–2502.

Lamm, C., Meltzoff, A. N., & Decety, J. (2010). How do we empathize with someone who is not like us? A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience, 22*(2), 362–376.

Lee, B. P. (2001). Mutual knowledge, background knowledge and shared beliefs: Their roles in establishing common ground. *Journal of Pragmatics, 33*(1), 21–44.

Lockwood, P. L., Bird, G., Bridge, M., & Viding, E. (2013). Dissecting empathy: High levels of psychopathic and autistic traits are characterized by difficulties in different social information processing domains. *Frontiers in Human Neuroscience*, *7*, 1–6.

Mascaro, J. S., Rilling, J. K., Tenzin Negi, L., & Raison, C. L. (2013). Compassion meditation enhances empathic accuracy and related neural activity. *Social Cognitive and Affective Neuroscience*, *8*(1), 48–55

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*, 370–396.

Mathur, V. A., Harada, T., Lipke, T., & Chiao, J. Y. (2010). Neural basis of extraordinary empathy and altruistic motivation. *NeuroImage, 51*(4), 1468–1475.

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion, 5*(2), 175.

Mazza, M., Pino, M. C., Mariano, M., Tempesta, D., Ferrara, M., De Berardis, D., Masedu, F., & Valenti, M. (2014). Affective and cognitive empathy in adolescents with autism spectrum disorder. *Frontiers in Human Neuroscience, 8*, 791.

Meinel, C. & Leifer, L. (2011). Design Thinking research. In H. Plattner, C. Meinel & L. Leifer (eds.), *Design thinking. Understand – improve – apply* (pp. xiii–xxi). Heidelberg: Springer.

Meinel, C., & Leifer, L. (2012). Design thinking research. In *Design thinking research*, (pp. 1–11). Springer.

Meinhardt, J., Sodian, B., Thoermer, C., Döhnel, K., & Sommer, M. (2011). True- and false-belief reasoning in children and adults: An event-related potential study of theory of mind. *Developmental Cognitive Neuroscience, 1*(1), 67–76.

Molenberghs, P., Johnson, H., Henry, J. D., & Mattingley, J. B. (2016). Understanding the minds of others: A neuroimaging meta-analysis. *Neuroscience and Biobehavioral Reviews, 65*, 276–291.

Morelli, S. A., & Lieberman, M. D. (2013). The role of automaticity and attention in neural processes underlying empathy for happiness, sadness, and anxiety. *Frontiers in Human Neuroscience, 7*, 1–15.

Nakahara, H., Furuya, S., Obata, S., Masuko, T., & Kinoshita, H. (2009). Emotion-related changes in heart rate and its variability during performance and perception of music. *Annals of the New York Academy of Sciences, 1169*(1), 359–362.

Naor, N., Rohr, C., Schaare, L. H., Limbachia, C., Shamay-Tsoory, S. G., & Okon-Singer, H. (2020). The neural networks underlying reappraisal of empathy for pain. *Social Cognitive and Affective Neuroscience*, (pp. 1–12).

Nielsen, L. (2013). *Personas-user focused design*. Springer.

Pinker, S. (2011). *The better angels of our nature: The decline of violence in history and its causes*. Penguin.

Plank, I. S. (2020). Neuroscience of empathy. https://www.tele-task.de/lecture/video/8304/.

Plattner, H., Meinel, C., & Weinberg, U. (2009). *Design thinking*. MI-Wirtschaftsbuch.

Preckel, K., Kanske, P., & Singer, T. (2018). On the interaction of social affect and cognition: Empathy, compassion and theory of mind. *Current Opinion in Behavioral Sciences, 19*, 1–6.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences, 1*(4), 515–526.

Price, E. E., Wood, L. A., & Whiten, A. (2017). Adaptive cultural transmission biases in children and nonhuman primates. *Infant Behavior and Development, 48*, 45–53.

Prinz, J. (2011). Against empathy. *The Southern Journal of Philosophy, 49*, 214–233.

Prinz, J. (2011). Is empathy necessary for morality. *Empathy: Philosophical and psychological perspectives*, *1,* 211–229.

Qiao-Tasserit, E., Quesada, M. G., Antico, L., Bavelier, D., Vuilleumier, P., & Pichon, S. (2017). Transient emotional events and individual affective traits affect emotion recognition in a perceptual decision-making task. *PLoS ONE, 12*(2), e0171375.

Reisenzein, R. (1983). The Schachter theory of emotion: Two decades later. *Psychological Bulletin, 94*(2), 239.

Richins, M. T., Barreto, M., Karl, A., & Lawrence, N. (2019). Empathic responses are reduced to competitive but not non-competitive outgroups. *Social Neuroscience, 14*(3), 345–358.

Rodríguez Buritica, J. (2020). Examining social influences on brain and behavior across development. https://www.tele-task.de/lecture/video/7977/.

Roth, B. (2017). Design thinking in Stanford. In *Design thinking live. 30 perspectives on making innovation happen*, (pp. 80–86). Murmann.

Ruffman, T., Lorimer, B., & Scarf, D. (2017). Do infants really experience emotional contagion? *Child Development Perspectives, 11*(4), 270–274.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporoparietal junction in "theory of mind" . *NeuroImage, 19*(4), 1835–1842.

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences, 19*(2), 65–72.

Schlaffke, L., Lissek, S., Lenz, M., Juckel, G., Schultz, T., Tegenthoff, M., Schmidt-Wilcke, T., & Brune, M. (2015). Shared and nonshared neural networks of cognitive and affective theory-of-mind: A neuroimaging study using cartoon picture stories. *Human Brain Mapping, 36*(1), 29–39.

Schmieden, K., Taheri, M., & Mayer, L. (2019). Human-centered design: Building and testing prototypes. https://open.hpi.de/courses/prototype2019.

Schmieden, K., Taheri, M., & Mayer, L. (2018). Human-centered design: From synthesis to creative ideas. https://open.hpi.de/courses/ideas2018.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews, 42*, 9–34.

Schurz, M., Radua, J., Tholen, M. G., Maliske, L. Margulies, D. S., Rogier, B., Sallet, J., & Kanske, P. (2020). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin.*

Shamay-Tsoory, S. G. (2011). The neural bases for empathy. *The Neuroscientist, 17*(1), 18–24.

Shamay-Tsoory, S. G., Saporta, N., Marton-Alper, I. Z., & Gvirts, H. Z. (2019). Herding brains: A s. *Trends in Cognitive Sciences, 23*(3), 174–186.

Sheng, F., Du, N., & Han, S. (2017). Degraded perceptual and affective processing of racial outgroups: An electrophysiological approach. *Social Neuroscience, 12*(4), 479–487.

Sheng, F., & Han, S. (2012). Manipulations of cognitive strategies and intergroup relationships reduce the racial bias in empathic neural responses. *NeuroImage, 61*(4), 786–797.

Sheng, F., Liu, Q., Li, H., Fang, F., & Han, S. (2014). Task modulations of racial bias in neural responses to others' suffering. *NeuroImage, 88*, 263–270.

Simner, M. L. (1971). Newborn's response to the cry of another infant. *Developmental Psychology, 5*(1), 136.

Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences, 13*(8), 334–340.

Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology, 24*(18), R875–R878.

Singer, T., Seymour, B., O'Doherty, J. P., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science, 303*(5661), 1157–1162.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature, 439*(7075), 466–469.

Stietz, J., Jauk, E., Krach, S., & Kanske, P. (2019). Dissociating empathy from perspective-taking: Evidence from intra- and inter-individual differences research. *Frontiers in Psychiatry,10*(126).

Szymanski, C. (2019a). Neural team dynamics. https://www.tele-task.de/lecture/video/7952/.

Szymanski, C. (2019b). Social neuroscience and teamwork. https://www.tele-task.de/lecture/video/7720/.

Taheri, M., Schmieden, K., & Mayer, L. (2017). Inspirations for design: A course on human-centered research. https://open.hpi.de/courses/insights-2017.

Thompson, R. A. (1987). Empathy and emotional understanding: The early development of empathy. In *Empathy and its development* (pp. 119–145).

Timmers, I., Park, A. L., Fischer, M. D., Kronman, C. A., Heathcote, L. C., Hernandez, J. M., & Simons, L. E. (2018). Is empathy for pain unique in its neural correlates? A meta-analysis of neuroimaging studies of empathy. *Frontiers in Behavioral Neuroscience, 12*, 1–12.

Traifeh, H., Nicolai, C., Refaie, R., Meinel, C., et al. (2020). Engaging digital engineering students in design thinking. In *DS 101: Proceedings of NordDesign 2020, Lyngby, Denmark, 12th-14th August 2020* (pp 1–14).

Trautwein, F. M., Kanske, P., Böckler, A., & Singer, T. (2019). Differential benefits of mental training types for attention, compassion, and theory of mind. *Cognition, 194*, 2020.

Trommsdorff, G., Friedlmeier, W., & Mayer, B. (2007). Sympathy, distress, and prosocial behavior of preschool children in four cultures. *International Journal of Behavioral Development, 31*(3), 284–293.

Van der Graaff, J., Meeus, W., de Wied, M., van Boxtel, A., van Lier, P. A., Koot, H. M., & Branje, S. (2016). Motor, affective and cognitive empathy in adolescence: Interrelations between facial electromyography and self-reported trait and state measures. *Cognition and Emotion, 30*(4), 745–761.

Verganti, R., Dell'Era, C., & Swan, K. S. (2019). *Motivation for and objectives of special issue.* Wiley Online Library.

von Thienen, J., Kolodny, O., & Meinel, C. (2021). Neurodesign: The biology, psychology and engineering of creative thinking and innovation. In N. Rezaei (Ed.), *Thinking: Bioengineering of science and art*. Springer Nature.

von Thienen, J., Noweski, C., Meinel, C., Lang, S., Nicolai, C., & Bartz, A. (2012). What can design thinking learn from behavior group therapy? In *Design thinking research* (pp. 285–302). Springer.

von Thienen, J., Noweski, C., Meinel, C., & Rauth, I. (2011). The co-evolution of theory and practice in design thinking–or–"mind the oddness trap!" In *Design thinking* (pp. 81–99). Springer.

von Thienen, J., & Santuber, J. (2020). Normative aspects in creativity, collaboration and culture development. https://www.tele-task.de/lecture/video/7936/.

von Thienen, J., Weinstein, T., Szymanski, C., & Meinel, C. (2021). *Neurodesign education.* Springer.

von Thienen, J. P. A. (2017). *Kausalniveaus: Eine Methodenanalyse zur Kausalforschung der Psychologie.* Pabst Science Publishers.

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist, 54*(7), 480.

Wells, G. L., & Petty, R. E. (1980). The effects of over head movements on persuasion: Compatibility and incompatibility of responses. *Basic and Applied Social Psychology, 1*(3), 219–230.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103–128.

Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *Journal of Neuroscience, 29*(26), 8525–8529.

Zaki, J. (2017). Moving beyond stereotypes of empathy. *Trends in Cognitive Sciences.*

ZhouandS, Y., & Han. (2021). Neural dynamics of pain expression processing: band synchronization to same-race pain but desynchronization to other-race pain. *NeuroImage, 224*, 117400.

Zuo, X., & Han, S. (2013). Cultural experiences reduce racial bias in neural responses to others' suffering. *Culture and Brain, 1*(1), 34–46.

Creation of Models for Prototyping

Toward Patterns of Exploratory Programming Practice

Marcel Taeumel, Patrick Rein, and Robert Hirschfeld

Abstract Patterns document best practices in many domains. For a long time, practitioners in the field of software engineering have been collecting and using such patterns too, to approach recurring design challenges. However, the challenges of efficient problem understanding and solution revising have no such form for efficiently communicating programming practice. It takes a long time to discover and learn such exploratory skills when using programming tools as is, without thorough reflection. We want to apply the idea of patterns to capture traditional and modern practices of exploratory programming. In this chapter, we begin to draft a pattern language, starting with four patterns to enable and control exploration, which we extracted from personal programming practice and experience.

1 Introduction

Software development often has the characteristics of a wicked problem (DeGrace & Stahl, 1990). Creating "good" software requires fulfilling external requirements relevant to the users and internal requirements relevant for the developers of the system. Users value easy-to-learn interfaces and useful features, and developers appreciate code that can be understood by others and architectural designs that can be adapted easily. Such mediation between two sometimes quite different sides entails constant communication efforts as depicted in Fig. 1. Often, we discover some of these (external and internal) requirements *only after* we build a version of the software. Given the complexity of many software systems, such timing can be fatal: we

M. Taeumel (✉) · P. Rein · R. Hirschfeld
Hasso Platter Institute, 14482 Potsdam, Germany
e-mail: marcel.taeumel@hpi.uni-potsdam.de

P. Rein
e-mail: patrick.rein@hpi.uni-potsdam.de

R. Hirschfeld
e-mail: robert.hirschfeld@hpi.uni-potsdam.de

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_7

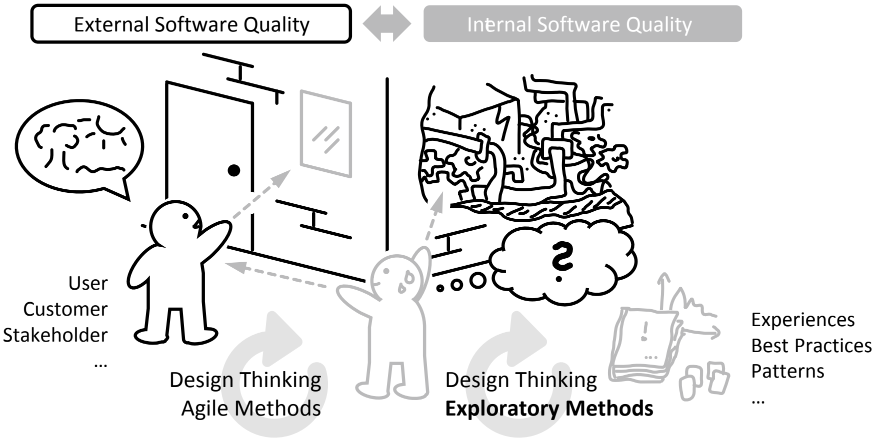


Fig. 1 Software developers mediate between user requirements and technical implementation. The goal is to master the implementation side and find a program design that can quickly adapt to changing user requirements

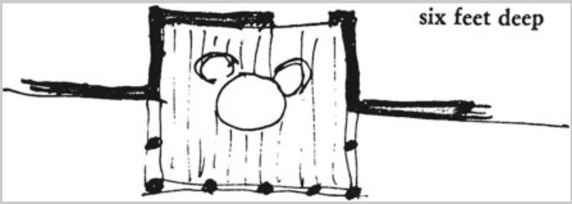
will only understand the full consequences at a point where systems become very difficult to change.

To tackle such "wickedness," some programmers engage in an intensive form of exploring the problem and solution space. When reasoning about their designs, they follow a programming style called "exploratory programming" (Sheil, 1983; Trenouth, 1991). In this style, programmers do not try to construct the perfect solution right away, but aim to deeply understand the problem at hand, including possible solutions. They achieve such immersion by creating various prototypical implementations (Kery & Myers, 2017; Trenouth, 1991), which they can directly try out and refine to eventually look at the system from many different perspectives. Thereby, the exploration process consists of many small, yet insightful, experimental changes to the prototype, whose consequences programmers can directly observe and learn from (Trenouth, 1991).

However, the actual skill set around exploratory programming is still difficult to learn. Proficient programmers typically invest years to acquire a number of best practices to support exploration: How to keep the overhead of iteration low? How to avoid breaking the system? How to get detailed feedback for answering any particular question?—Experienced programmers do not only know the steps to be taken, but also which ones work best in a particular situation, and how to adapt them to new ones. This knowledge is the result of years of practice or direct observation of other programmers during their explorations. There is no form to efficiently pass on this specific knowledge to the next generation of programmers.

We want to ease the learning of exploratory programming style by uncovering and documenting best practices as *patterns*. Patterns area concise form of communicating the core of a solution obtained through experience (Alexander, 1979; Alexander et al. 1977). Patterns typically describe the *problem* they are tackling, the *context* in which they are to be applied, the core of the *solution*, and its *consequences*. Depending on a pattern's subject (or domain), some of these parts are left out or described in more detail. A common property of all pattern collections is their *generative* nature. By adapting the essential part of the solution, they can be used to generate solutions that are tailored to new, unforeseen challenges. We think that the pattern form is suitable to describe exploratory programming practices. Their structure can provide guidance to help us describe all relevant aspects of such practices, while at the same time being flexible enough to cover a variety of practices.

Where does the pattern form originate? The first-ever published pattern collection describes solutions for creating and shaping the living environment, including towns, houses, and individual rooms (Alexander et al., 1977). One such pattern is "six foot balcony." This pattern first describes the *context* in which the pattern is to be used: It then goes on to describe the main problem it tackles: "Balconies and porches which are less than six feet deep are hardly ever used." The *problem* is discussed further by describing the observation of how people make different use of balconies that are narrow and ones that are deep. The pattern then discusses the *solution* including variations such as enclosing balconies or recessing them into buildings:



"Whenever you build a balcony, a porch, a gallery, or a terrace always make it at least six feet deep. If possible, recess at least part of it into the building so that it is not cantilevered out and separated from the building by a simple line, and enclose it partially." —Alexander (1977, p. 784)

The patterns in our collection are based on the experience of our research group. This experience is the result of more than a decade of research and education around exploratory programming systems. For both aims, we employ two exploratory programming environments: Squeak/Smalltalk (Goldberg & Robson, 1983; Ingalls et al., 1997) and Lively (Ingalls et al., 2016; Lincke et al., 2012, 2017). In research, we create tools to support exploration of software designs by allowing programmers to change running systems, quickly adapt their tools to the task at hand, and gain more insights into the actual behavior of their system (Lincke et al., 2017; Rauch et al., 2019; Taeumel, 2020). In education, we conduct lectures on software architecture and software engineering, as well as other courses on software design, end-user programming, and tool building. All these courses include practical work in one of these environments. This provides us with ample opportunity to observe the typical struggles of beginners in such environments.

This chapter presents a first draft of four relevant patterns on exploratory programming, which we observed in object-oriented system programming. Our contribution is as follows:

- "Tangible Names" and "Tangible Pixels"—Two patterns to *enable* exploration, which clarify the textual and graphical aspects of environments as entry points.
- "Configurable Constraints" and "Reliable Recovery"—Two patterns to *control* exploration, which clarify the aspect of trust in the environment as to avoid getting lost or wreaking havoc.

In Sect. 2, we provide more detail on how and where we found the patterns for this chapter, that is, our experience with teaching and research in object-oriented programming systems. The following two sections describe the patterns: Sect. 3 for *enabling* exploration and Sect. 4 for *controlling* exploration. We close those descriptions with a reflective discussion on their maturity in Sect. 5 as this is the first iteration of a pattern collection for exploratory programming. We conclude our thoughts in Sect. 6.

2 From Experience to Pattern Form

We precede the presentation of this chapter's patterns on exploratory programming practice with more background on our expertise in object-oriented programming systems. At the end, we sketch the actual pattern form that we use, which deviates from existing pattern catalogs but fits our needs for capturing those practices.

2.1 Our Programming Experiences

The patterns described in this chapter result from our own experience with exploratory programming. To provide you with a background for the patterns, we will briefly outline our experience. We are a research group of 14 people focusing on programming tools and experience. Our experience with exploratory programming stems from engaging in it ourselves, from research around corresponding tools and environments, and from teaching undergraduate and graduate courses in exploratory environments.

Our research revolves around exploratory and live programming (Rein et al., 2019). As part of this research, we design and create new tools and programming environments, designed to support specific exploratory practices. Many of the ideas and environments we build upon originated from the Learning Research Group at Xerox PARC (Kay & Goldberg, 1977). The systems and ideas, which have stood the test of time, form the foundation for many of the following projects.

Exploratory programming entails new situations that are not supported by existing tools. The Vivide environment (Taeumel et al., 2012, 2014, 2020) supports exploratory programmers in creating new tools or adapting their tools quickly to such new situations. The Babylonian Programming systems (Niephaus et al., 2020; Rauch et al., 2019; Rein et al., 2019) allow programmers to annotate their source code with examples that are then used to display the results of expressions directly within the source code. Thus, they can get feedback on dynamic behavior anywhere in the environment. The Lively environments (Ingalls et al., 2016; Lincke et al. 2017) bring the ideas of exploratory programming to Web programming by allowing programmers to develop a Web application from within itself. As modern applications are often based on several languages, the Squimera and the TruffleSqueak environment support exploratory programming for such systems through polyglot exploratory tools (Niephaus et al., 2018, 2020). Finally, as exploration involves the creation of alternative solutions, the CoExist environment supports programmers with managing and switching between multiple variations (Steinert et al., 2012).

These systems, and similar ones created by others, are designed to support exploratory practices. To make use of them, programmers need to have basic knowledge of exploratory practices, so that they can recognize how the tool supports them and to recognize situations in which the tool or environment is applicable.

We have experience in teaching exploratory programming on the undergraduate as well as the graduate level. For most of our courses, we use Squeak/Smalltalk. As we teach undergraduate courses and many students continue their graduate studies at our university, we can work with the same students several times during their studies. At the undergraduate level, we teach two lectures, one on software architecture and one on software engineering, each spanning three months. During these lectures, the students work on projects in Squeak/Smalltalk. At the graduate level, we teach seminars on software design, programming tools, execution environments, and modularity. All of the seminars focus on project work in Squeak/Smalltalk or Lively. Throughout this time, we are able to observe how they acquire exploratory programming practices. While these observations are not empirically verified, they serve as a starting ground to determine which practices beginners pick up by themselves and which ones need to be taught explicitly, for example through patterns. A general observation is that students progress from learning the language Smalltalk, to learning individual tools of Squeak/Smalltalk, to making use of the whole environment, and eventually learning more general practices of exploratory programming.

2.2 A Purely Object-Oriented Programming System

We describe the solution and examples of each pattern from the perspective of purely object-oriented, exploratory programming systems (Wegner, 1990), namely Squeak/Smalltalk (Goldberg & Robson, 1983; Ingalls et al., 1997). To make the pattern descriptions accessible, we provide a short background on this perspective by briefly introducing the basic concepts.

The main element of such systems is *objects*. Objects are used to represent entities relevant to the system, for example, a domain-specific entity such as a person, a system-specific entity such as a file, or basic information such as a number. In particular, an object stores the data relevant for that entity, for example, a person's name and date of birth, or a file's most recent modification timestamp. Beyond the data, an object also has behavior, which can be invoked by sending a message to the object. The sum of all behaviors of all objects defines the behavior of the system. The behavior of objects is typically the same for all objects of the same kind, for example, all chat message objects can be sent to a person and can create an object representing a reply message. This common behavior of one kind of object is captured in an abstraction called a *class*, for example, a chat message class. All objects of one kind have the same class. Thereby, the behavior described in the class is re-used for all those objects.

Beyond this basic principle, the Squeak/Smalltalk perspective takes a different angle on the notion of systems and programs. In Smalltalk systems, the program or application to be created is part of the running programming environment. So, programmers do not create a program or system outside of the environment, but change the running environment itself from within to make it behave the way they want. As programmers can access anything in that environment, and the environment is the running program or application they want to modify, they can access all parts of the running program or application. To store such a system and share it with other, Smalltalk systems can be saved into an *image*. This feature is similar to hibernation in operating systems; all state and all running processes is saved into a file. The system can be restarted from that image file and will be in the exact same state as it was when the programmers saved it.

2.3 Pattern Audience

Our pattern collection is motivated by making exploratory programming practices learnable by novices. Thereby, programmers are users of an environment and try to employ the practices during programming. However, programmers can also be the builders of their environments. As such, they might want to make use of the patterns to get guidance in how to shape an environment for exploratory programming.

For learning the practices, the patterns are useful for exploratory programming novices and experts alike. When talking about novices, we refer to programmers new to exploratory programming. This includes programming novices, who are unfamiliar with programming in general, as well as programming experts, who are already familiar with programming, but not with exploratory programming. Both benefit from the pattern representation of practices. So far, learning exploratory programming either required a lot of time to build up personal experience, or an experienced teacher, regularly demonstrating practices by example. With the patterns, novices can now learn the practices by themselves.

Experts of exploratory programming may still benefit from the practices. As there are no written, in-detail descriptions of the practices of exploratory programming, most programmers only have their individual experience to go by. Through the patterns, they can contrast their implicit techniques with the experience of others.

Further, they may discover variations within and commonalities between practices they are not aware of, which might make their practices more effective.

Independent of a programmer's skill level, the patterns may help programmers building tools and languages fit for exploration. Tool builders may refer to the exploratory programming patterns to determine what their language or environment needs are to support certain practices. Further, they might choose to support particular practices, and the patterns may provide some background on when the practice is used or how it may be altered by programmers.

2.4 Pattern Form

Since the introduction of patterns (Alexander, 1979; Alexander et al. 1977), different communities and authors have taken up the idea and created their own pattern collections. While they all agree on the idea of a pattern as the description of the core of a solution, they differ in the form they use to describe the patterns. The main difference between these forms is the list of aspects described for each pattern.

The original pattern descriptions by Christopher Alexander consist of the name of the pattern, the context, the problem described as a set of forces, the solution, trade-offs in the solution, and a set of related patterns. The "Gang-of-Four" book, which popularized patterns in the software development community, uses a more form that includes several detailed sections describing the solution (Gamma et al., 1995). Yet another form was used in the learning and presentation patterns (Iba, 2014; Iba & Sakamoto, 2011), which featured a summary of the pattern consisting of only one line.

For this chapter, we use the following form:

- **Intent** is a short summary of the pattern including the fundamental challenge as well as a glimpse of the solution.
- **Motivation** describes the problem programmers might encounter during their exploration. It describes the domain and the context in which the problem occurs. This section concludes with a summary of why programmers may develop a "desire for exploration" in this situation.
- **Forces to Resolve** describes the different constraints and considerations when applying the practice. Whenever adapting the pattern to a specific situation, these forces may influence the specific adaptations.
- **(Toward a) Solution** describes the specific techniques making up the practice, including variations.

- **Consequences** describe what is required from a programming system to support this practice. It also points out technical challenges that may arise during the particular practice used in exploration.
- **Notes on Squeak/Smalltalk** illustrate how the practice would be applied in the Squeak/Smalltalk system.

3 Patterns to Enable Exploration

Our first collection of two patterns is about *enabling exploration.* In purely object-oriented systems, there is much structured information available. When investigating bugs or adding features, programmers access the object graph to understand what is there and what is missing. When explicating thoughts, programmers benefit from naming objects and then organizing those *tangible references* in spaces. Since modern programming tools offer graphical interfaces, programmers also have to make sense of a program's visual output. In sum, this combination of typing (names) and clicking (on shapes) represents an entry point to exploratory practice.

3.1 Tangible Names

Maybe also known as "Object Bindings" or "Names in Spaces."

Intent

Names help people denote accessible meaning of otherwise transient thoughts. Therefore, programmers should use names to organize not only code artifacts but all relevant objects. Programming environments should allow for flexible attachment of such names. Consequently, the use of established vocabulary should yield access to the underlying artifacts.

Motivation

In object-oriented environments, all structured information is represented as objects that have relationships to other objects. Those structures can be very deep and thus hard to follow and abstract. Code objects typically have intrinsic names to be easily identified. Many other objects, especially those that occur at run-time, may not have a (derived) textual representation that helps programmers in their understanding.

Programmers are in a constant learning process. They communicate with domain experts (or customers) to understand the rules and requirements that should be somehow represented in code. Along the way, programmers make all kinds of observations—such as computational results—that need to be documented to not get lost. Ideas emerge and become clearer. Consequently, such emergent clarity needs to be denoted *before* becoming program code, like sticky notes in the programming environment.

Programmers work with names on a regular basis because source code is filled with such textual identifiers for classes, methods, and all kinds of variables. Names help explicate thoughts; they encode meaning. In an environment where all kinds of objects can be materialized thoughts, names play an important role in keeping track.

The question is whether programmers are willing to write down and handle names when talking (and reasoning) *about* programs and their informational trails.

Today's programming tools are full of textual labels. There are code browsers or object inspectors, which employ text fields or lists with labels. Programmers rely on their recognition of an object's intrinsic names to look up and find information of interest. Also, text-based search is a common entry point in program understanding. Programmers just type (part of a) name into a text field and expect interesting objects to show up in a (text-based) result list. Name it, spell it, type it, find it.

Programmers' Desire for Exploration. An object's inherent structure does not yield a name appropriate for the current task. The programmer wants to reduce cognitive load by explicating and working with new names in the environment:

- Attach a name to an object for later reference.
- Look up the object structure for any name that is visible on screen.
- Share names between several tools (or scopes).
- Organize thoughts on different levels such as domain, task, or personal.

Such names may change. They can be mere nicknames (or mnemonics) in the beginning.

Forces to Resolve

Programmers usually understand the importance of good names in source code, but they might hesitate to bring the same attention to names that appear in the entire programming environment:

- Names may not be reachable outside a certain tool or other scope.
- In the "offline" world, taking notes is *very* easy.
- Arbitrary name lookup is not possible for arbitrary labels in (graphical) tools.
- A good name is hard to find.
- Recognizing a name on screen "feels good" and reduces cognitive load.
- Extra references to objects consume extra resources in the environment.
- The same name can change meaning over time.

(Toward a) Solution

Programmers have to come up with and refer to names all the time when writing or reading source code. This very habit (or custom) is the base for working with *tangible names* during exploration. Names can be very helpful to organize objects, even outside a program's code base.

Know about and rely on the system's vocabulary. At the beginning of an exploratory session, programmers work with the names that already exist and are accessible for technical artifacts. For example, there are package names, class names, or method names. Besides such language artifacts, there can be names for run-time objects such as the current process and active method. The environment should make this basic level accessible through names to serve as entry points for exploration. The programmers' thoughts may spin around those technical artifacts, which triggers the desire to learn more about what's going on.

Just write down that name. Once an idea or observation starts to gain clarity, programmers will try to describe it to explicate meaning—or at least give it a token to further think about it, to not forget about it. In the programming environment, programmers should be able to just type names for further reference. There might not even be an object attached to such names yet. Still, programmers can now go looking for objects that deserve such names.

Attach (more appropriate) names to objects. Many programming tools display characteristic object structure in textual form. Given a programmer's current task, a name may come up that would serve as a more appropriate identifier. Programmers should attach such a name to the particular object, so that its meaning can be recalled more easily. Programming environments should allow for adding any number of extra names to objects.

Organize names in spaces. Programmers should find "an empty sheet of paper" or "a clean whiteboard" to document their thoughts in the exploration process with little friction loss. For example, windows with big empty text fields are common metaphor to represent such spaces on screen. Such spaces can directly represent task scopes; they may be even expanded to document overall domain knowledge. Programmers can collect names and attach them to objects on the fly. Like sticky notes on whiteboards, names can be moved around to influence each other in the overall process of program understanding.

Resolve names to reveal structure. Within a certain space, programmers should be able to directly resolve the names they have just typed or observed in a tool's graphical display. The environment should keep track of the names' connection to the underlying object and hence the structured information. As an effect, the object's (intrinsic) textual form can appear or a more sophisticated tool can offer a means to explore structure. In text widgets, such name lookup resembles code evaluation. In list widgets, the connection of any visual label to an underlying model (and thus object) might be more challenging if not supported by the tool framework.

Combine spaces to integrate exploration paths. Programmers should reflect on the spaces they currently use for name collection. Related themes may emerge, which requires to combine spaces (or at least bring names over from one space to another.) Programmers should avoid connecting "loose ends" in offline notes. Instead, they should employ the means in the (digital) environment such as shared clipboards or drag-and-drop gestures. Consequently, the environment should offer a basic model for tool (and thus name) integration without compromising data quality.

Dismiss the spaces you no longer need. Programmers should reflect on their current task's progress. Once finished, spaces should be dismissed. Adding extra names to objects may interfere with the environment's automatic cleanup mechanism. Since resources are usually limited, discarding names (or entire spaces) is part of the exploratory process. Note that there are usually means to recover from mistakes. Some environments may offer automatic dismissal of no longer needed spaces, which programmers may have to configure to accommodate their working habits.

Consequences

Being able to name objects requires object representations for information in the environment. External data can usually be imported as generic structures such as maps and dictionaries. Materializing low-level language (and run-time) concepts, however, can be more challenging for the provider of the programming system. Yet, programmers are likely to include "behavior" or "execution stack" in their thoughts when thinking about specification and implementation.

Having the freedom of reasoning about any accessible object with new names in custom spaces, programmers can easily break abstractions. There are environments without a certain compilation boundary, that is, source code access to all parts in a system, which demands a certain discipline for information hiding. Programmers need to be aware of not "leaking" usually hidden information into new source code, that is, after the exploratory session.

Aliasing is already a challenge in object-oriented architectures. During exploration, programmers add even more names to the same objects, which makes identity a rather intangible, hardly explicable concept. As different (work)spaces support overlapping names, tools for overview might mitigate this consequence. On the other hand, working with external data (and distributed structures in general) implies a comparable challenge outside the context of exploratory programming practice.

Names can point to outdated structure without programmers' being aware of it. Programs (under observation) only manage *their* point of view. Programmers hold on to certain objects by chance, but have often no means to notice when related objects "lose interest" in their direct neighbors.

Notes on Squeak/Smalltalk

In Squeak/Smalltalk, programmers can write notes into workspaces, which are interactive text buffers that support code evaluation like a read-eval-print loop (or REPL). The Smalltalk language can be used as a scripting language in almost any other tool's text fields to set up new, but tool-local, name bindings. The combination of such tool spaces is possible, for example, through global variables. There are globals (and reserved keywords) that reference basic run-time information such as thisContext for the active method (context) and ActiveWorld for the topmost GUI object.

Many graphical tools in Squeak retain a (more or less direct) connection between visual label and underlying object. This connection allows programmers to explore underlying structure through simple pop-up menus or drag-and-drop gestures. Vivide (Taeumel, 2012, 2020) is a tool-construction framework on top of Squeak/Morph that preserves such a direct connection in the GUI by design. Consequently, programmers can resolve names not only in text fields but other interactive widgets too.

3.2 Tangible Pixels

Maybe also known as "Meta Menu" or "Shape Halo" or (more generic) "Direct Manipulation Interface."

Intent

Visual shapes can raise attention and trigger curiosity to explore. Programmers use interactive spaces to organize graphical representations on screen. Programming environments should allow to "look behind" visual shapes to explore the underlying objects and relationships. In practice, programmers can point and click to manipulate such shapes directly.

Motivation

Using high-resolution, graphical displays, programmers can create convincing illusions of tangibility merely through colorful pixels on a two-dimensional plane. Even if a program-under-construction has no elaborate visuals itself, today's programming tools (and environments) can offer visualization to clarify (code) structure. The question is whether programmers accept graphical interfaces only from a user's perspective or whether they also try to work with *visual shapes* as a tangible medium under construction.

The shared (programming) environment uses objects to represent everything, including graphical primitives. There can easily be extra gateways to connect "what is visible" to "what it is made of." Sometimes, a widget's affordances guide programmers to shorten the feedback loop in their exploratory journeys—such as clicking a button nearby to reveal a pop-up menu. Yet, extra (hidden) gestures may have to be learned to enable exploration.

While the connection between a visual shape to *any* underlying object may be simple, finding *useful* paths to descriptive model data may not be. That is, *spatial distance* can be reduced with elegant software design, while *semantic or temporal distance* often remains part of the exploration efforts.

Programmers' Desire for Exploration. A visual shape on screen makes the programmer curious because it may indicate a bug or place for a new feature. The programmer wants to reduce cognitive load by directly navigating from the pixels to objects and hence structured information:

- Understand the structure behind flat pixels.
- Open tools to explore that structure, to make it tangible.
- Keep the connection between tools and visuals on screen.

- Organize thoughts on different levels such as domain, task, or personal.

Direct manipulation (for exploration) helps shorten the feedback loop.

Forces to Resolve

Programmers usually design graphical interfaces for usage only, not for exploratory (debugging) practice. Still wanting to understand how the underlying objects enable the program's purpose, programmers might hesitate to even try using the same interface to also "look behind the curtains," that is, the visual shapes. The following forces emerge:

- The program's GUI has no extra code to enable debugging.
- Visual objects should directly relate to model data in the domain.
- The scene graph is too deep and complex.
- The visuals are too small to point at.
- There are no distinct, steady shapes; it is more like animation.

(Toward a) Solution

Graphical output is often the "result" of the system. Starting the exploration from there means programmers start from something that they can grasp and that is already tied to a purpose.

Point and hover. Many visual shapes on screen offer extra information when programmers hover the mouse cursor over them and wait for a bit. Then, descriptive tooltips (or balloon texts) appear as over lays nearby. Such user interaction strengthens the blending of pixels into tangible compounds (or graphical objects). In programming tools, the revealed insight can indeed help programmers to look at object structure. In other (regular) programs, such information might be targeted toward its users, not programmers who what to "look behind the curtain."

Look for and click on meta buttons. There are UI elements that do not invoke immediate side effects on the system (or program). Such elements are often clickable buttons that offer possible actions through pop-up menus. Looking at such actions, programmers can get a better understanding of what object is actually displayed in pixels nearby. Similar to hover effects, the emerging visual compounds support the connection between pixels and underlying structure.

Employ reserved (meta) input gestures. Three-button mice render click-on instructions ambiguous. The primary click on, for example, buttons or list elements, is part of the common bi-manual interaction mode—keyboard plus mouse. Yet, there can be many other input gestures (such as keyboard shortcuts) that encode special modes or means to interact with visual shapes. Users may want to *talk about* a thing on screen when they perform a secondary click and expect a pop-up menu to show up. Programmers should know such meta gestures as they might exploit underlying objects. There are environments that make the entire scene graph tangible.

Enter gateways to reach (meta) tools. Programmers should follow the shortest path available to explore the connection between the visuals on screen and the underlying object structure. Programming environments should allow for such short paths through reserved input gestures. That is, there should be a connection between a program's run-time objects and the objects that make up the source code (or other resources).

Organize programs and tools in spaces. When the environment offers the means to explore running programs and structure-revealing tools side by side, programmers should organize those in (visual) spaces. Such spaces help document exploration paths (and overall progress) in a tangible way.

Consequences

The visual design may be in conflict with serving both user and programmer. Allocating extra screen space for buttons (or similar) might confuse users, which would defeat the primary purpose of that program. Also, increasing shapes' sizes so that programmers can click on and "look behind" the surface might not be a viable option either.

Extra input gestures—dedicated to exploratory programming—would not be useful for regular users. Already, there is often a dispute on supporting common keyboard shortcuts for common (user) operations. Mouse buttons are limited and so are keys on the keyboard. Taking away more options would interfere with this discussion from a new perspective.

Programmers would have to learn about extra interface elements and how to use them while running the program-under construction. It can already be challenging to organize non-visual objects, and separate essential from supportive. Visual objects further aggravate this issue. User interface and programming interface might blend, which could be okay for programmers, but frustrating for users.

Depending on the system's rendering pipeline, preserving a pixel-to-object mapping can be challenging. If not supported by the underlying graphics framework by design, extra programming effort may be required to at least offer such a connection for selected programs.

Notes on Squeak/Smalltalk

Squeak has always supported one-button mice in making point-and-click interfaces discoverable and simple to use. For example, there is a button for a list widget's menu, placed in the scrollbar. There is no need to learn a secondary click: rather the user first clicks on a list element, then clicks on the menu button to show available actions for that element. Note that three-button mice are also supported. In that case, the secondary click avoids extra mouse movement.

In Squeak/Morphic, all graphical objects—so-called morphs—can be selected through a special gesture. Then, a "context menu" appears in the form of a halo around that object. While this menu can be used as part of the regular user interface, it also offers a gateway to programming tools such as object inspectors and code browsers. The halo concept originates in the *outliner* in the Self system (Ungar & Smith, 2007).

The collection of pixels that represent a graphical object can be difficult to see. The Morphic halo appears as a rectangular "outline," which can be invoked whenever the programmer has a reference to such a graphical object (or morph). Consequently, there is also a direct connection from object to pixels—not just the other way around.

4 Patterns to Control Exploration

Our second collection of two patterns is about *controlling exploration*. Programmers have to trust their programming environment and tools. While learning about a problem domain, implementation strategies, and personal preferences, programmers will gain trust in their tools if those can mirror that progress. If one's mindset can be observed on the screen, programmers will get a feeling of being in control. First, they can set up boundaries to avoid making mistakes and derailing, but staying focused instead. Second, they can establish an area to safely work within, which includes reliable recovery and cleaning up after the exploration task.

4.1 Configurable Constraints

Maybe also known as "Configurable Guides" or (more generic) "Domain-specific Environments."

Intent

"With great power comes great responsibility." Being deep in an exploration activity, programmers benefit from meaningful limitations while they progress—to stay in focus and avoid mistakes. Programming environments should allow for configurations that constrain or guide the tangible notion of names and pixels.

Motivation

To foster the programmer's mindset for exploration, the environment should take care of traps that would otherwise distract or intimidate. When programmers have to fear drastic consequences, they might resort to unchecked hypotheses instead of exploring and learning about what is really happening. In self-sustaining systems, such consequences could entail broken tools or lost data, which in turn means higher costs in the software development process.

Distraction may come from standard tools showing irrelevant information such as low-level code in debuggers or irrelevant modules in browsers. Intimidation may come from a sheer overwhelming amount of possibilities. Luckily, many tools can be tailored to specific exploration strategies. Programmers can reduce cognitive effort when screen contents match their mental model as closely as possible.

Domain-specific tools can help guide programmers actions in a generic fashion. That is, a tool's specificity can complement its expressiveness. Programmers should always stay in control; they decide how they want to proceed. Tools (and environments), however, help programmers remember and apply best practices.

Programmers' Desire for Exploration. Programming environments offer many complementary tools. Programmers have to explore their choice of tools as well as the information visible through these tools:

- Keep going and stay in focus.
- Explore *relevant* object structure.
- Hide *irrelevant* implementation details.
- Use non-limiting support for the current task.

Meaningful constraints can promote a state of "flow" (Csikszentmihalyi, 2008) in exploratory programming.

Forces to Resolve

When not following *a single plan* but exploring possibilities to gain understanding, programmers may hesitate to freely embrace exploration within the programming system:

- It is hard to recover; therefore, mistakes must be avoided.
- It is hard to focus because generic programming tools "leak" implementation details.
- It is hard to proceed because domain-specific tools impede general-purpose programming if needed.

(Toward a) Solution

Programmers can avoid many mistakes and stay focused within exploration by mastering the means of representing information in the environment. That is, they have to choose the right tools, tweak tool parameters, and know when to change plans.

Choose tools appropriate for exploration. Programming environments usually have many different tools for many different programming tasks. There is often no "one fits all" solution; a notable overlap in tool features can occur. Programmers should choose from tools that fit the desired exploration strategy. Selection criteria include accessibility and representation of relevant software artifacts.

Configure to accommodate specific needs. Programming tools are often "general purpose" but also offer configuration parameters to accommodate specific domains, tasks, or personal preferences. Therefore, programmers should schedule extra time to tweak those parameters. Especially at the beginning of exploration, known characteristics of the current problem domain can already be included.

Reflect and realize when to change strategies. Being deep within an exploration path, programmers should account for extra time to reflect on the current working mode. Different tools might be more appropriate to continue, including generic code browsers. Different configurations of the tools in use might yield more promising results. That is, exploring the problem and solution spaces includes exploring the available means to do so.

Start exploration in empty spaces. Programmers should avoid interfering with the results of other work in the system. A new exploration (path) should start in a rather empty space such as a new instance of a tool window. Programming environments should account for having enough space to follow many different hunches.

Migrate progress to new tools. When programmers choose to switch tools, they should also try to bring existing insights along. That is, all names and meaningful objects, including visuals, should remain (somewhat) accessible in the other tool's interface. There will be compromise because different tools have different strengths and levels of data support.

Consequences

Tool configuration may blend into tool construction, which may take unexpected time and effort. Especially in open systems where programmers can access and modify the entire codebase, one has to carefully "timebox" any attempt to change the status quo. Thus, the matter of "staying focused" becomes double-edged: using tools and also configuring them.

Switching tools *and* transferring (intermediate) results may only work within a certain environmental boundary. If the underlying representation of structured information differs fundamentally, programmers might have to compromise and serialize parts of this information as they see fit. If such a reduction in quality is not an option, programmers can try to integrate external tools directly into the programming environment.

Trial-and-error remains part of the exploration process. Programmers cannot always know when to switch tools. In any case, the overall programming task may still be "timeboxed," leaving only limited resources for out-of-plan exploration. However, such a limitation can be an obvious trigger for programmers to "just try something different" in any remaining period.

Notes on Squeak/Smalltalk

Squeak comes with tools that are tailored to the Smalltalk language. Class browsers show source code; object inspectors show instance variables; debuggers accurately display the context of method activations. Consequently, guidance comes from the (hopefully descriptive) names of code artifacts. While there are simple filters, programmers have to selectively disregard unrelated information. There are no on-board means for higher-level, domain-specific perspectives that could guide exploration.

Luckily, there are frameworks and libraries that build on top of Squeak/Morphic, which programmers can install to support exploration. These includes projects that aim to improve programming education and programming experience in general. Yet, they can play part of their role in (general purpose) exploratory practice. Etoys (Freudenberg et al., 2009) and Scratch (Maloney et al., 2010), for example, both hide textual code complexity through visual shapes—meant to be composed and explored through click, drag, and drop. Then, there is Babylonian programming (Niephaus et al., 2020; Rauch et al., 2019; Rein et al., 2019), which embeds concrete values into abstract code, so that programmers do not have to stray and lose time in breakpoint triggered debuggers. There are also object-focused, script-based means to construct new tools for exploration with the Vivide framework (Taeumel et al., 2012, 2014, 2020). Programmers therefore have many alternatives to choose from.

4.2 Reliable Recovery

Maybe also known as "Safety Net" or "Back to the Start" or "Checkpoints."

Intent

Programmers leave traces during exploration. Those traces may need to be altered when backtracking or removed when finishing. Programming environments should allow for configurations that manage (or constrain) side effects on software artifacts (including the tangible notion of names and pixels).

Motivation

Programmers consume many different kinds of information when trying to understand programs and possibilities. Yet, consumption can entail change such as disassembling a closed box. It is thus advisable to take extra care to scope the effects of such exploration. That is, programmer's do not just observe, but they actually "poke around" to learn how specific objects react.

The most obvious solution—known from "traditional" programming practice—is typically too costly: throw away everything and start over. There can be a nondeterministic state, which is hard to replicate for another round of exploration. When programmers are continuously modeling artifacts in a running system, restarting might also imply tediously retyping source code or remodeling other essential resources. Luckily, there have been approaches that shorten the cycle of recovery to try again or continue work.

Programmers' Desire for Exploration. Programming environments can be both messy and tidy at the same time. It is very easy to create empty spaces; it is "just" digital software. Programmers want to dive into the exploration task:

- Keep going and stay in focus.
- Backtrack when hitting a dead end.
- Clean up when finished exploring.
- Quickly recover when having broken something by accident.

Programmers can easily forget about that cleanup, which can later become a reason for unnecessary recovery.

Forces to Resolve

Having the system's state made of interconnected objects, programmers have to take care of those objects and their relationships during exploration. Like cleaning up your study may be not worth the effort, programmers may hesitate to follow exploratory practices:

- Living with "brittle" (run-time) state around for too long feeds the urge to "reboot" and start afresh.
- It is hard to disseminate the "broken" from the useful state.
- It is costly to throw "everything" away.
- It is hard to anticipate the effects of exploration tools (and actions) upfront.

(Toward a) Solution

Even within a constrained and guided setup, programmers can make mistakes and need to recover. A system's object graph may just be too complex to foresee the effects of every possible action.

Use tangible (and easily discardable) scopes. During exploration, programmers grow a collection of (perhaps newly) named objects. This collection should represent a scope that can easily be dismissed when finished. The environment's resources are typically limited; automatic cleanup works only through computational, user-independent rules. Thus, programmers must explicitly indicate the state of exploration as they see fit. A tangible scope can be pointed to and thus helps with such indication.

Establish distinct steps on a path. Programmers should modularize their exploration path. At best, an obvious (only linearly dependent) sequence of steps (or tools or scripts) can be re-evaluated repeatedly while the program (under observation) keeps running. Along such paths, programmers can easily backtrack and revise their choices.

Create checkpoints for safe retreat. Programmers should replicate (or copy) a specific setting before experimenting with unknown side effects. This has a similar effect to the way that children can have repeated fun by coloring (by numbers) on a photocopy, rather than on the original. Programming environments should offer clear guidelines to specify and duplicate (part of) the object graph. Clear boundaries, like shielded sandboxes, can further help to establish trust between programmers and their environment.

Hit the pause button to take a break. Exploration can be time-consuming. Programmers have to consider their working schedule and thus maybe interrupt a session. Thus, programming environments should offer a means to pause all action in the running system—or selected modules. On the one hand, programmers can then take a closer look at such "snapshot of time" to better understand the objects and messages in situ. On the other hand, programmers can actually take a break and rely on the system to continue running—exactly where it left off—the next day.

Consequences

Modularity in the exploration path largely depends on guides and constraints offered through tools and their interfaces. If programmers would be forced to put much effort into refactoring existing steps, chances are that they would not do it. Such extra effort would interfere with their focus and thus interrupt the "flow." Consequently, the modular description of exploration steps is one of the primary challenges in domain-specific tool construction.

At the same time, there can be *too many* checkpoints, outliving past exploration tasks and demanding extra resources. Programmers might hesitate to discard (even tangible) scopes because these form new objects of value, that is, documentation for later use. There can always be new but similar challenges in the near future; one cannot know upfront. Yet, the actual value of such (maybe outdated) checkpoints can be difficult to assess, even in retrospect.

Programmers might avoid creating *complete* checkpoints for reasons of cost. It might even be impossible to strive for completeness. External resources can be especially difficult to grasp; stubbing them can interfere with trust in the exploration's outcome. In other words, working with *real data* is a problem force that is not addressed through this pattern.

Notes on Squeak/Smalltalk

Squeak's tools (and associated windows) can represent tangible scopes to organize exploration and clean up after it. For example, programmers close *workspace windows* to dismiss bindings and thus *tangible names*. They also organize multiple windows in *projects* (or "desktops"), which can easily be closed to dismiss open tools and thus *tangible pixels*.

Within a single workspace (window), programmers modularize (partial) scripts through text lines of source code. Consequently, they are in charge of orchestrating simple inspection or effectual experimentation. At best, programmers can re-evaluate the entire code in a workspace without breaking things or "polluting" the environment with useless data.

Programmers can hit the key combination [CMD] + [.] at any point to suspend the currently running process. That is, they can pause message passing for a specific portion in the system, usually the UI process. After inspection, suspended processes can then be resumed—or terminated to free resources. In combination with Squeak's image, programmers are basically in control of (execution) time. Yet, there is ongoing research on how to offer more elaborate tools for immediate recovery in Squeak. For example, CoExist (Steinert et al., 2012) offers fine-granular revisions for code changes without needing programmers' anticipation of mistakes.

5 Discussion and Future Work

A pattern's success is measured through relevance, quality, and impact. Its mere discovery is of less importance. In this chapter, we attempt to describe aspects of exploratory practice in pattern form for the first time. The result is a collection of "drafts" that need to be polished and revised. Yet, the sole artifact of a pattern is not useful unless applied in practice. That is, fellow programmers who use the practice of exploratory programming should see value in our work. Therefore, like many patterns and pattern-authors before us, we seek feedback from both practitioners in the field and the pattern community—which takes time and several pattern-writing workshops.

The patterns we drafted are very broad and leave many questions unanswered.

We made an attempt to formulate not only actions for programmers but also advice for tool builders. From experience, we know that both tool usage and construction go hand in hand. Indeed, it may be the same programmer who switches between roles many times during the same programming task. Consequently, our patterns can reveal shortcomings in programming environments regarding its tools and means for construction. Since time is always a scarce resource in software development, some patterns may thus not be feasible to apply. We want to address such situations in our next revisions in also offering more paths to enable exploration.

In our next steps, we will tackle verbosity to make each pattern's intent more clear. Especially, the solutions we propose in each pattern are likely to be split up into patterns of their own, leaving the current form as possible categories for orientation. Of course, when discovering complementary patterns or new perspectives as a whole, the entire organization can change. In the process, we will also investigate practices beyond object-oriented systems, because exploration happens in every programming environment. Our vision is to collect and materialize an accessible catalog of patterns—maybe even create a pattern language (Alexander et al., 1977)—that can serve as a reliable reference in daily programming practice.

6 Conclusion

In this chapter, we described typical approaches of exploratory programming practices as they occur in education and research through the Squeak/Smalltalk system. We gained many of our own experiences with this system's concepts, which already originated in the 1970s and hold up splendidly for today's challenges. The system's purely object-oriented design offers many interesting perspectives on program understanding and debugging with short feedback loops. First, we covered patterns to *enable* exploration, which unpacks the role of textual labels and visual shapes. Second, we addressed patterns to *control* exploration, which emphasizes not only avoiding mistakes but also embracing them through trusted means for recovery.

This chapter is only the first step toward a more substantial collection of patterns, maybe a whole pattern language, that can serve programmers in many domains. Even at this early stage,we implymany valuable aspects of exploratory programming practice to be further unpacked in pattern form. We believe that such a comprehensive, accessible catalog can help connect many overlapping efforts in contemporary programming language and tool research.

Acknowledgements We gratefully acknowledge the financial support of the HPI Research School on Service-oriented Systems Engineering (www.hpi.de/en/research/research-schools) and the Hasso Plattner Design Thinking Research Program (www.hpi.de/en/dtrp).

References

Alexander, C. (1979). *The timeless way of building*. Oxford University Press.

Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., & Angel, S. (1977). *A pattern language—Towns, buildings*. Oxford University Press.

Csikszentmihalyi, M. (2008). *Flow:The psychology of optimal experience*. Harper Perennial Modern Classics

DeGrace, P., & Stahl, L. (1990). *Wicked problems, righteous solutions: A catalogue of modern software engineering paradigms*. Yourdon Press Computing Series. Yourdon Press. URL https://books.google.de/books?id=__omAAAAMAAJ

Freudenberg, B., Ohshima, Y., & Wallace, S. (2009). Etoys for one laptop per child. In *2009 Seventh International Conference on Creating, Connecting and Collaborating Through Computing* (pp. 57–64). IEEE

Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1995). *Design patterns: Elements of reusable object-oriented software*. Addison-Wesley Longman Publishing Co.

Goldberg, A., & Robson, D. (1983). *Smalltalk-80: The language and its implementation*. Addison Wesley Longman Publishing Co.

Iba, T. (2014). *Presentation patterns: A pattern language for creative presentations*. Lulu. com.

Iba, T., & Sakamoto, M. (2011). Learning patterns III: A pattern language for creative learning. In L. B. Hvatum (Ed.) *Proceedings of the 18th Conference on Pattern Languages of Programs, PLoP 2011* (pp. 29:1–29:8), Portland, Oregon, USA, October 21–23, 2011. ACM. https://doi.org/10.1145/2578903.2579166.

Ingalls, D., Felgentreff, T., Hirschfeld, R., Krahn, R., Lincke, J., Röder, M., Taivalsaari, A., & Mikkonen, T. (2016). A world of active objects for work and play: the first ten years of lively. In E. Visser, E. R. Murphy-Hill, & C. Lopes (Eds.) *2016 ACM International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software* (pp. 238–249), Onward! 2016, Amsterdam, The Netherlands, November 2–4, 2016. ACM. https://doi.org/10.1145/2986012.2986029

Ingalls, D., Kaehler, T., Maloney, J., Wallace, S., & Kay, A. (1997). Back to the future: The story of squeak, a practical smalltalk written in itself. In *Proceedings of OOPSLA, 32*, 318–326. ACM. http://doi.acm.org/https://doi.org/10.1145/263698.263754

Kay, A., & Goldberg, A. (1977). Personal dynamic media. *Computer, 10*(3), 31–41.

Kery, M. B., & Myers, B. A. (2017). Exploring exploratory programming. In A. Z. Henley, P. Rogers, & A. Sarma (Eds.) *2017 IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC 2017* (pp. 25–29), Raleigh, NC, USA, October 11–14, 2017. IEEE Computer Society (2017). https://doi.org/10.1109/VLHCC.2017.8103446.

Lincke, J., Krahn, R., Ingalls, D., Röder, M., & Hirschfeld, R. (2012). The lively partsbin-a cloud-based repository for collaborative development of active web content. In *45th Hawaii International International Conference on Systems Science (HICSS-45 2012), Proceedings* (pp. 693–701), 4–7 January 2012, Grand Wailea, Maui, HI, USA. IEEE Computer Society (2012). https://doi.org/10.1109/HICSS.2012.42.

Lincke, J., Rein, P., Ramson, S., Hirschfeld, R., Taeumel, M., & Felgentreff, T. (2017). Designing a live development experience for web-components. In L. Church, R.P. Gabriel, R. Hirschfeld, & H. Masuhara (Eds.) *Proceedings of the 3rd ACM SIGPLAN International Workshop on Programming Experience, PX/17.2* (28–35), Vancouver, BC, Canada, October 23–27, 2017. ACM. https://dl.acm.org/citation.cfm?id=3167109

Maloney, J., Resnick, M., Rusk, N., Silverman, B., & Eastmond, E. (2010). The scratch programming language and environment. *ACM Transactions on Computing Education (TOCE), 10*(4), 1–15.

Niephaus, F., Felgentreff, T., Pape, T., Hirschfeld, R., & Taeumel, M. (2018). Live multi-language development and runtime environments. *The Art, Science, and Engineering of Programming, 2*(3). https://doi.org/10.22152/programming-journal.org/2018/2/8

Niephaus, F., Rein, P., Edding, J., Hering, J., König, B., Opahle, K., Scordialo, N., & Hirschfeld, R. (2020). Example-based live programming for everyone: Building language-agnostic tools for live programming with lsp and graalvm. In *Proceedings of the 2020 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software* (pp. 1–17), Onward! 2020. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3426428.3426919

Rauch, D., Rein, P., Ramson, S., Lincke, J., & Hirschfeld, R. (2019). Babylonian-style programming design and implementationof an integration oflive examples into general-purpose sourcecode. *Programming Journal, 3*(3), 9. https://doi.org/10.22152/programming-journal.org/2019/3/9

Rein, P., Lincke, J., Ramson, S., Mattis, T., Niephaus, F., & Hirschfeld, R. (2019). Implementing Babylonian/s by putting examples into contexts: Tracing instrumentation for example-based live programming as a use case for context-oriented programming. In *Proceedings of the Workshop on Context-oriented Programming* (pp. 17–23).

Rein, P., Ramson, S., Lincke, J., Hirschfeld, R., & Pape, T. (2019). Exploratory and live, programming and coding—A literature study comparing perspectives on liveness. *Art Science Engineering Program, 3*(1), 1. https://doi.org/10.22152/programming-journal.org/2019/3/1.

Sheil, B. (1983). Power tools for programmers. *Datamation Magazine*

Steinert, B., Cassou, D., & Hirschfeld, R. (2012). Coexist: overcoming aversion to change. In A. Warth (Ed.) *Proceedings of the 8th Symposium on Dynamic Languages, DLS '12* (pp. 107–118), Tucson, AZ, USA, October 22, 2012. ACM. https://doi.org/10.1145/2384577.2384591.

Taeumel, M. (2020). *Data-driven tool construction in exploratory programming environments* (Ph.D. thesis). University of Potsdam, Digital Engineering Faculty, Hasso Plattner Institute. https://doi.org/10.25932/publishup-44428

Taeumel, M., Perscheid, M., Steinert, B., Lincke, J., & Hirschfeld, R. (2014). Interleavingofmodification and use in data-driven tool development. In *Proceedings of the 2014 ACM International Symposium on New Ideas, New Paradigms, and Reflections on Programming & Software* (pp. 185–200), Onward!. ACM, New York, NY, USA. http://doi.acm.org/https://doi.org/10.1145/2661136.2661150

Taeumel, M., Steinert, B., & Hirschfeld, R. (2012). The VIVIDE programming environment: Connecting run-timeinformationwithprogrammers'systemknowledge. In G. T.Leavens & J.Edwards(Eds.) *ACM Symposium on New Ideas in Programming and Reflections on Software* (pp. 117–126), Onward! 2012, part of SPLASH '12, Tucson, AZ, USA, October 21–26, 2012. ACM. https://doi.org/10.1145/2384592.2384604

Trenouth, J. (1991). A survey of exploratory software development. *The Computer Journal, 34*(2), 153–163. https://doi.org/10.1093/comjnl/34.2.153

Ungar, D., & Smith, R. Self. In *Proceedings of the Conference on History of Programming Languages (HOPL) 2007, HOPL III* (pp. 9–1). ACM, New York, NY, USA. https://doi.org/10.1145/1238844.1238853

Wegner, P. (1990). Concepts and paradigms of object-oriented programming. *OOPS Messenger, 1*(1), 7–87. https://doi.org/10.1145/382192.383004.doi:10.1145/382192.383004

Designing Photography Guidance for Rapid In-Camera Iteration

Jane L. E. and James A. Landay

Abstract Designers have long known the benefits of iteration and rapid prototyping. Many experienced photographers follow a similar process, in particular of iterating in camera: trying out different angles and compositions, varying lighting, adjusting a subject's pose, etc. However, amateurs often do not realize the benefits of capturing variations of a single shot. Inspired by the parallels between the design thinking process and the photographic process, we design new interfaces that provide contextual in-camera feedback to aid users in learning visual elements of photography. We interactively visualize results of image processing algorithms as additional information for the user at capture time. In this chapter, we explore ways to encourage stages of the design thinking process, specifically through the design of guided photography interfaces that aid in iterating through the exploration of three different photographic concepts: lighting, composition, and decluttering.

1 Introduction

As cameras become smarter and more pervasive, more people want to learn to be better content creators. However, currently cameras provide limited aid in improving the esthetic quality of the user's photographs. For an amateur who is interested in photography, but has limited training and equipment, the prospect of trying to take a "good" photograph can be quite daunting. One common mistake that amateurs frequently make is taking too few photographs in the moment and relying on editing to improve their photographs. However, changes that can be made at the editing stages are significantly limited by the already captured content, and often many mistakes cannot be fixed at all without returning to the photograph location. As supported by our formative interviews, experienced photographers will often capture

J. L. E. (✉) · J. A. Landay
Department Computer Science, Stanford School of Engineering, 353 Jane Stanford Way, Stanford, California 94305, USA
e-mail: ejane@stanford.edu

J. A. Landay
e-mail: landay@stanford.edu

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_8

many photographs of a given scene: they know how to consider different options in composition, lighting, pose, etc., and recognize the challenges of not having options upon editing.

We see similarities in this photographic process and the design process. Designers have long known the benefits of iterating quickly and having many prototypes. Bill Buxton describes the benefits of ideating through sketching due to its flexibility "enabl[ing] ideas to be explored quickly and cheaply" (Buxton 2010). Similarly while still in the process of taking photographs, it is cheaper to iterate and generate more "prototypes," or photographs. Photographers can get feedback from users, whether it be from the photographer themselves, or possibly customers/subjects in the photographs, and integrate this feedback into future prototypes. Upon leaving the location, these photographs become "high-fidelity prototypes," and changes are limited and more expensive.

We are interested in designing camera interfaces that can encourage photographers to ***incorporate stages of the design process into their photographic process***. In particular, we want to take advantage of the strengths of computation to generate visualizations that provide additional information in the form of in-camera feedback, providing users with new lenses of sorts with which to see the camera viewport.

In this chapter, we begin by describing some challenges (Sect. 8.2) of trying to learn photography. We then summarize some findings from our formative studies (Sect. 8.3) to understand people's current photography practices and learning experiences. Based on these findings, we propose a set of design goals (Sect. 8.4) for designing camera interfaces that provide contextual feedback to help users become more aware of the artistic choices they are making in the camera. We then describe some related work (Sect. 8.5) and present three examples of interfaces (Sect. 8.6) in the context of the proposed design goals. These guidance photography interfaces aim to aid users in iterating in-camera through the exploration of three different photographic concepts: lighting, composition, and decluttering.

2 Learning Photography

People have always loved taking photographs. In particular, we want to document memories through photographs. In the past, this required carrying around a designated camera whenever there might be moments worth documenting. Even before that, it involved hiring an expert to bring a camera to a specific location to take photographs. In recent years, a large population has gained access to relatively high-quality cameras directly built into their phones that travel with them throughout the day.

These cameras have made way for a new type of more casual photography. People no longer have to spend a fortune on a separate device (and multiple lenses), spend hours learning the knobs and dials of a complex camera, and lug around a heavy device in order to be able to capture photographs. The number of photographs taken per year tripled from 2015 from 2010, with 75% captured using a phone as compared to 40% in

2010 (Heyman 2015). People are taking more photographs, and as a result, more and more people are interested in photography (Cakebread 2017; Heyman 2015; Zhang 2015). However, there is still limited access to effective opportunities for learning how to improve at photography. For a novice who is interested in photography, but has limited training and equipment, the prospect of trying to take a "good" photograph can be somewhat daunting.

2.1 Photography Resources

Many photography learning resources are readily available to a wide population, ranging from books to videos, etc. However, this content is often static and therefore needs to rely on describing these concepts in the context of existing examples. In practice, it can be difficult to take these abstract concepts and apply them in the context of a new photograph. In education literature, this is referred to as transfer. It can be challenging to bridge the gap when transferring knowledge you have learned in one context to a completely different context (Ambrose et al. 2010; National Academy of Sciences-National Research Council 1999).

Additionally, in the moment of trying to capture a new photograph, there are numerous considerations to manage simultaneously. Unlike many other art forms, photography has no "blank canvas." Any image seen through the camera viewfinder has the potential to be a candidate for the final photograph. How does one transition from just "taking pictures" to actually "doing photography"? Given the popularity of photography, taking this step to try to stand out, and put in the intentional effort of "crafting" a photograph, can be particularly intimidating (Batista 2015). It can be hard for a novice photographer to recognize which concepts to try to apply, much less how to apply them in a way that communicates their artistic intentions.

2.2 In-Person Photography Classes

In-person, practical art classes, on the other hand, are often more effective in addressing this knowledge gap. However, they are also harder to make broadly available as they require the time and involvement of expert photographers/instructors. One unique aspect of in-person classes is that they focus heavily on in-class critiques—in which the teacher and other classmates provide feedback directly in the context of the student's work. Commonly this feedback translates abstract photographic concepts into descriptions of how they relate to the student's work—often taking the form of annotations directly in the context of the students' work (e.g., sketched or air-drawn in gestures by the teacher). These annotations can sometimes be used to highlight characteristics of the photograph or to point out mistakes to fix. These concrete visualizations of abstract concepts can feel more concrete and actionable, narrowing the knowledge gap.

2.3 Limitations of Current Practices

However, the teacher's instruction and feedback is still a scarce resource. Teachers will usually only be able to provide feedback during class-time, and their attention is spread across the students in class. Therefore, each student will usually only get feedback on a select subset of their work. It is not feasible for teachers to be around to give feedback on each and every photograph as each student is out in the field capturing it. Thus, when the feedback is provided, the student is already more committed to the image. This can make receiving feedback more intimidating, but also more difficult to apply. The student has already left the location where this photograph was taken, and so the effort required to capture new images that apply most feedback at this point is significant. It would be preferable to receive this feedback in the moment while the student is taking the photograph and able to explore new ideas with the context of the feedback within the environment (Ambrose et al. 2010).

This is somewhat unique to photography due to the immediacy of this art form as compared to other slower art forms like sketching, painting, sculpting, etc. Here, much of the art process is happening in the classroom making it easier for teachers to walk around and give feedback in the moment as the student progresses through each step of creating their art piece. Having feedback in-the-moment is crucial to encouraging creativity through iteration and exploration (Dubberly 2004; Jansson and Smith 1991; Plattner et al. 2009); therefore, we aimed to design ways to encourage this exploration in photography. Specifically, we aim to further close this feedback loop by providing photographers with artistic feedback directly in the camera viewport, so photographers can use this information in the moment as they make decisions about their photographs.

3 Formative Studies

> I've found that shooting lots of pictures, experimenting, and reviewing results has been the best method [for learning photography].—survey participant.

In order to inform the design of our camera interfaces, we conducted a survey to understand people's current photographic process and their learning experiences, including what was effective about the different learning methods and whether any of it could be helpful at capture time. We additionally did nine in-depth interviews with experienced photographers about their current practices and tools.

3.1 Photography Practice Survey

Our first question involved understanding if people wanted in-camera guidance at all, and if so, what type of feedback would be helpful or not to their current photography practice.

To answer this question, we designed a survey asking about people's existing photography practice and past learning experiences. We surveyed adults with an interest in photography and received a total of 127 responses from participants (74 male, 51 female, 1 non-binary, 1 preferred not to say), 19–63 years old with a range of photography experience.

Many of our participants had had some sort of photography learning experience (only 23 had never used any resources such as classes, videos, books, etc., to learn photography)—several of the ones who had taken in-person classes (22 of 45) mentioned the effectiveness of the "direct, immediate feedback in the moment" (P7), that those provided. Specifically, many mention the benefits of the feedback being "individualized" (P111) and "hands-on" (P73) for making immediate adjustments and correcting mistakes. Without this guidance, it can be difficult to "transfer knowledge to other conditions" (P10). Thus, we saw benefit in pursuing the direction of trying to provide contextually based in-camera guidance, so users can make better creative decisions during capture.

When asked why they would or would not use capture-time guidance in their camera, many mentioned concern with an app being too "distracting" (P51), "disruptive" (P53), or "intrusive" (P93). This was due to either concern about missing a moment (P12), intruding on others' time when in a group (P14), or feeling less artistic freedom (P86). Guidance should thus prioritize being minimally distracting and more about providing suggestions than insisting on specific artistic choices.

Our survey respondents supported the idea of providing suggestions for larger changes in the camera. On a Likert scale from 1 (strongly disagree) to 5 (strongly agree), participants rated their editing practices as significantly more around small framing adjustments (Mdn = 4, IQR = 3–5) than substantial cropping to change the image's composition (Mdn = 3, IQR = 2–4) [Wilcoxon signed-rank test Woolson (2007) $V = 2217$, $p < 0.001$], which further supports the idea of providing composition guidance at capture time to reduce the need for more drastic changes while editing later.

Overall, people have a preference for getting feedback on the framing of their current photograph (Mdn = 4, IQR = 3–4) rather than thinking of possible photographs to take (Mdn = 3, IQR = 2–4) [Wilcoxon signed-rank test $V = 2183$, $p < 0.001$]. Thus, we should additionally focus our guidance on helping people refine their current image's composition. Of our survey respondents, 89% are willing to spend up to 5 s on capture-time guidance to get a high-quality result.

3.2 Experienced Photographer Interviews

To go into more depth on what types of guidance could be helpful in photography practices, we interviewed nine experienced photographers about their current photography practices and tools. All had formal training in photography, and five additionally had teaching experience. Interviews were structured around the following questions:

- Describe your typical photography process(es).
- What photography tools do you use and what guidance does it provide?
- Would composition guidance be helpful for you?

Six of the interviewees expressed that they already consistently use overlays of sorts such as focus dots, light meters, levels, or composition grids (primarily the rule-of-thirds). This suggests that photographers are okay with some amount of their camera view being obstructed when that information is useful to their overall ability to take better photographs.

When asked what guidance might be helpful for them, some (5) expressed interest themselves in something that might give them new perspectives, like an "experienced photographer on my shoulder saying try this, try that" (P7), or even providing random composition guidelines to swipe through and try (P0). In general, these experienced photographers were very open to having any feedback that might help them try out different ideas in taking a photograph in order to have more choices to pick from when going back to edit in the future.

Some suggested they are able to use guidance as needed while maintaining creative freedom to break the rules and follow their intuition. For example, explicitly disregarding the composition grid:

> might have a rule-of-thirds overlay, but don't follow it super closely. I gravitate towards the bottom two eyes in the rule-of-thirds… (P1).

However, others mentioned conforming to a certain style, whether it be due to the prominence of the rule-of-thirds overlays, realizing an unexpected theme across many photographs, or just having a

> tendency to view the world in a specific way, but someone else might be different, always open to try a different type of shot or idea (P6).

In describing their own photography processes, these photographers recounted both a process of searching for good compositions, as well as pre-composing and waiting for a shot—the latter was mentioned as particularly important in street photography. One described this waiting as a "necessary tension" in street photography (P5).

From these interviews, we learned that even experienced photographers could benefit from feedback that encourages them to try new ideas. In certain scenarios, they need to capture a shot immediately and thus rely on their instincts to quickly frame the photographs, while in other scenarios they are willing to spend more time and devote both screen real-estate and their attention to a tool that helps them achieve a higher quality image. We aim to support both of these scenarios, for novices and experts alike.

4 In-Camera Guidance Design Goals

According to our formative survey and interviews, we came up with three design goals for our photography guidance:

- **Context-Aware**. Guidance should adapt to the current scene and appear overlaid on the viewfinder.
- **Encourage Exploration**. It should help photographers discover new ideas while executing existing intentions.
- **Maintain Flexibility**. It should not restrict photographers from pursuing other creative choices.

Context-Aware. It can be difficult to apply an abstract photographic concept. Experts have experience doing so in many contexts and have developed patterns and established styles that they actively seek (methods/approaches that they use). Providing concrete suggestions and feedback in the context of the current image can make applying these abstract concepts approachable for non-experts. This includes both understanding how to apply a concept more broadly as well as refining the execution of an idea.

Many survey participants expressed that they appreciated the individualized and immediate feedback of in-person classes or even going on a photograph walk with more experienced friends. This enables reviewing and adjusting while still in the context of the current photograph.

Encourage Exploration. It can be hard to come up with an initial idea; just as it can be easy to fixate on perfecting a specific photograph idea. Encouraging exploration can help with both. However, it can be difficult to know how to explore. Experts know to take many different shots of any given subject and have a set of knobs and dials in mind to manipulate to generate drastically different ideas, and yet still expressed interest in guidance that could help them generate more ideas to explore. Making the space easier to explore and making different options apparent can make it easier to come up with new ideas to tryâĂŤ–especially for non-experts.

Both survey participants and experts were interested in new ideas or perspectives, in particular suggestions for bigger changes, as they preferred to only make smaller framing adjustments while editing.

Maintain Flexibility. Our goal is to assist the user in their creative process. Thus, the interface should allow users to have creative agency. It should help users better achieve their own artistic intentions, and not distract them by providing restrictive guidance.

Both survey participants and experts also expressed concern that the interface might be too distracting and disruptive. Survey participants were concerned that it would slow down their photographic process. This was especially important in situations with time pressure, such as when capturing a fleeting moment or when traveling with a group. Experts similarly worried about flexibility, but put more emphasis on the need to maintain creative freedom.

5 Related Work

In this section, we review some high-level areas of related work in the context of our design goals (described in Sect. 8.4).[1]

5.1 Automatic Photograph Improvement

Recent developments in graphics and vision have been quite helpful in producing esthetically improved images by introducing a range of post-processing algorithms. Examples include automatically cropping a photograph (Wei et al. 2018), removing distractors (Fried et al. 2015), or transferring lighting styles (Shu et al. 2017). However, much of the related work in this area focuses on automated workflows, which can somewhat detach users from the creative experience of taking a photograph. We can see that each of these methods is indeed context-aware. In fact, a lot of work is put into trying to recover the context in the image. These methods also allow users to quickly explore a wide range of options for their image. However, the options are still limited by the already fixed content of the image.

5.2 Guided Photography Interfaces

Additionally, some guided photography interfaces do exist, such as those shown here. These tools guide the user to a specific "better" option. Mitarai et al. (2013) propose an interface that guides a user to a specific secondary composition through navigational guidance. Li and Vogel (2017) propose an interface that guides a user to achieve what it defines as a "good" selfie, as determined through crowdsourcing. These tools have an internal model of esthetics that are mostly hidden from the user and apply this model without providing much explanation. By imposing a specific suggestion from the tool, they limit exploration and creative flexibility.

[1] For more detailed related work, see E et al. (2019) and (2020).

5.3 Existing In-Camera Tools

In our formative studies, many experts also described consistently using a number of existing in-camera overlays, from leveling assistance to light metering, etc., and not feeling like these limited their creative flexibility. The light meter, for example, attempts to evaluate if the current image is properly exposed. The photographer uses this information to adjust the camera to achieve better overall lighting. However, also note that a photographer can choose to intentionally over or underexpose the image based on this feedback. In fact, photographers will often quickly take a set of photographs with different light meter readings. This can be thought of as a quick succession of iterating to generate a number of protoypes. The photographer can then quickly review these "prototypes" and then further iterate to achieve the image they want. We take inspiration from these existing tools to design guidance interfaces that can similarly encourage this rapid prototyping/iteration cycle.

6 Photography Interfaces

Here, we present interface designs that tackle three different visual concepts of photography: lighting, composition, and decluttering. In particular, we focus here on how we applied the aforementioned design goals (Sect. 8.4) to the challenges associated with each of these photography concepts.[2]

6.1 Lighting

One of the most challenging and impactful considerations in photography is lighting. In a portrait studio, it is common to have a main light, fill light, and background light, as well as rim lights, hair lights, kickers, etc., positioned in a way to achieve a specific lighting style (Hunter et al. 2015). Non-experts generally do not have access to such equipment nor the knowledge of how to arrange them. However, in many environments, even when just relying on available light, the lighting on a face can vary drastically by just rotating the subject (Fig. 1).

We leverage this observation to design and implement an interface that shows the photographer a gallery of possible lighting styles achievable in the current environment, thereby helping the photographer orient the subject to capture their selected look (Fig. 2). Determining this orientation requires knowledge of the environment—specifically, the position of the subject relative to lights in the scene. We do this by capturing the environment with a 360° camera (E et al. 2020).

Thinking about our design goals, for portrait lighting, the important **context** is the lighting in the environment. To encourage **exploration**, we want to make sure

[2] For more details on these interfaces, see E et al. (2019) and (2020).

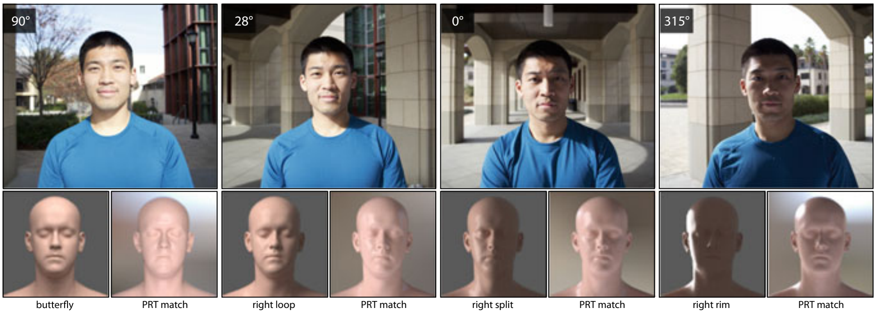


Fig. 1 In a fixed lighting environment, photographers can produce many different lighting styles (e.g., butterfly, right loop, right split, and right rim) just by rotating the subject in place without changing their location. Given an HDR environment map from a 360° camera at some initial orientation and a target lighting style (bottom left), our tool automatically identifies the optimal angle for reorienting the subject to match the desired lighting—e.g., 90° for butterfly lighting. We visualize an approximation of the best orientation match on a generic virtual scene (bottom right)

Fig. 2 Upon loading an HDR environment map of the current scene, the interface of our portrait lighting tool computes the optimal orientation for each target lighting style and displays a gallery of target appearances (left). The tool grays out unattainable targets. Selecting a target brings up the reorientation guidance screen, which displays how far the photographer should rotate clockwise or counterclockwise around the subject (right). It also shows the background at the target orientation and the current view from a webcamera at the current location of the photographer

users are aware of lighting style options in the context of their current environment, and that they have **flexibility** to choose between these different options. **Context-Aware.** In this case, we captured the important context, the lighting in the environment. By representing portrait lighting as a discrete set of lighting styles and corresponding reorientation angles considering the lighting conditions, we were able to allow participants to consider varied lighting styles in the context of the current scene.

Encourage Exploration. By presenting the user with these many options at once, the tool allows them to easily scan the gallery to preview what the end result might look like. Using the tool, participants were able to explore through a gallery of lighting styles, pick their desired style, and ultimately achieve results that they believed were of better quality.

Maintain Flexibility. The user has flexibility to choose between these styles. While we picked a predefined set of lighting styles, one can easily imagine adjusting the set of lighting styles based on personal preferences. We also included a painting interface for making custom adjustments to the lighting styles we provided.

6.2 Composition

Artists have developed many ways to describe composition. One way of interpreting composition is by looking at alignment of visually important elements in an image with the lines/intersection points of a composition grid, such as the rule-of-thirds (Ghyka 1977; Glover 2014; Northrup 2014). In addition to thirds, there are other proportions that are known to be visually pleasing. The harmonic armature is a composition grid that encompasses a variety of these proportions such as halves and quarters, formed by intersecting diagonals. This can describe a wider range of compositions, but can be overwhelming to interpret.

Currently in articles, books, etc., photographers will describe photographic composition by manually highlighting individual lines from such a grid to emphasize alignment choices of the image. We want to automate this process to enable an interactive guidance interface that provides these annotations as feedback to help the user discover such alignments directly in the context of the current camera image (E et al. 2020). Specifically, we heuristically define the set of relevant lines in the harmonic armature for a given image and display the composition grid with these lines interactively highlighted on the camera viewport (Fig. 3). Since composition is about the relative positioning of elements that attract attention, we use a saliency model to represent this visual attention and select the subset of lines that most closely align with these salient regions.

Considering our design goals for photographic composition, the **context** to consider is the relative positioning of objects within frame. **Exploring** composition can involve moving around the camera to try out different angles and framing, as well as moving the objects in frame. Users should have creative **flexibility** in choosing whatever compositions they find appealing.

Context-Aware. For composition, we are using saliency to capture the relevant context of the image. We aimed to reflect the composition of the current camera image with the highlighting of the most relevant composition lines in the grid, or the adaptive armature.

Encourage Exploration. By showing grid lines, we hoped that users could more easily think about possible composition options, and that the highlights especially, might help them understand their current composition or consider new composition options, while still attending to other aspects of the photograph. These high-level grid representations make it easier to explore the space of possible compositions and also highlighted new composition ideas for users to try. Figure 4 shows a range of

Fig. 3 To design our interactive composition guidance interface, we were interested in better understanding people's ability to recognize composition and to annotate them on a composition grid. We collected annotations from both experienced photographers as well as novices on Mechanical Turk. Inspired by these results, we developed an algorithm for heuristically computing these lines, or adaptive armatures. We display these adaptive armatures as an overlay in an in-camera composition guidance tool and study how it impacts how people take photographs. Here you see this guidance tool in action, highlighting the relevant lines in this image of the lounge chair

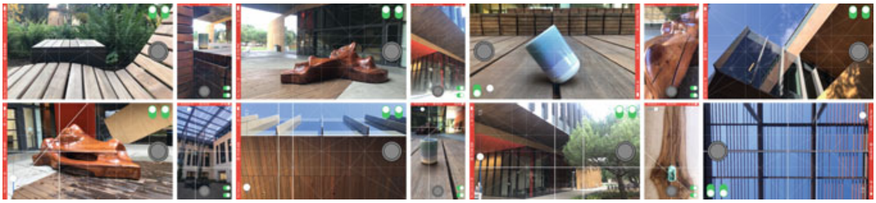


Fig. 4 Some participant photographs from a user study. These are shown with the overlays that the participants saw while doing the user study. Top: static guidance, bottom: adaptive guidance. Participants often found grid lines to align to elements or edges in the image, but also sometimes used them as looser guidelines for leveling the image, splitting the image into regions (e.g., thirds), or occasionally even disregarded the grid lines

photographs captured by participants in one of our user studies.

Maintain Flexibility. Finally, the highlights are meant to represent the user's current composition choices, rather than impose a new composition. Users have flexibility in how they respond. They can choose to better align to highlighted lines, align with different lines, or intentionally misalign elements relative to the grid.

Nonetheless, we chose the harmonic armature as the underlying grid to be able to support alignment to a wider range of compositions.

6.3 Decluttering

Once the photographer has decided on an approximate orientation and framing, a range of adjustments in framing can still significantly impact the quality of the final shot. With the photographer mainly focusing on the subject and the action, it can be easy for some unwanted objects in the background to go unnoticed. It can be incredibly frustrating to take a photograph only to realize after the fact that there is some unwanted clutter in the background that distracts from the main focus of the image (Glover 2014).

Our eyes are drawn to regions of high contrast—light foreground on a dark background, and dark foreground on a light background. Thus, contrast around the subject will help clarify and declutter the overall image, we refer to this as subject-background separation (or more formally known as figure ground relationship), whereas contrast in other regions especially the border of the image, will distract, causing the eye to be attracted away from the focal subject, which we call imageborder (or edge) flicker.

Photographers recommend a number of methods to be able to more easily see the contrast in an image. For example, they recommend squinting at the image to see a blurred and higher contrast version of the image, or looking at the image in grayscale to focus on contrast without aspects of color (Block 2013). Glover (2014) suggests drawing a line around the subject where areas where contrast is maybe too low between the subject and background are shown as gaps.

Inspired by this idea of using outlines to highlight contrast and lack of contrast, we hoped to recreate this outlining as an overlay directly in the camera, also extending it to contrast along the image borders. To do so, we detect edges throughout the image, and then use an object-based saliency to determine possible subjects in the image. We interactively show the edges along the borders of these detected subjects to represent subject-background separation, and then, the edges around the border of the image to represent image border flicker (Fig. 5).

For decluttering, the important **context** to consider is what will draw the viewers attention. **Exploring** how to declutter an image involves considering how elements of the image contribute to the overall story, and what might focus or distract. In particular, it might involve trying out different backgrounds that better highlight the focal subject, or removing objects that take attention away from the focal subject. Users should have **flexibility** to decide how they want to present the story.

Context-Aware. We use edge detection to capture edges of objects in the frame. We additionally use saliency to try to capture the context of how these edges relate to important subjects in the photograph. Our edge-based overlay distilled the concept of declutter down to the principles of subject-background separation and image border

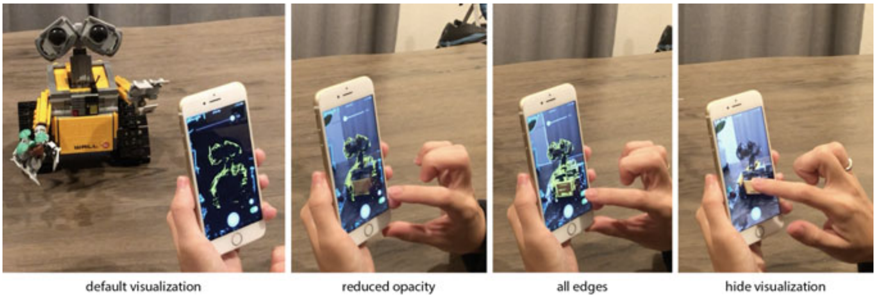


Fig. 5 A few states of the decluttering visualization. Left to right: The default visualization shows the edges around the detected subject border as well as around the image border over an opaque black background. The user can choose to reduce the opacity of the background to show some of the camera view. The user can also toggle on all edges to see edges within the subject and background as well. Finally, the user can either use the bottom toggle or touch anywhere on the screen to hide the visualization completely

flicker specifically in the context of the current image, and helped draw focus away from the subject itself, to these surrounding regions.

Encourage Exploration. By drawing attention to the subject-background relationship and the image borders, we hope this might encourage users to consider these when exploring different ways offraming their subject. This visualization helped participants quickly explore different backgrounds and angles for a capturing a subject, and ultimately feel more confident in their photographs.

Maintain Flexibility. Users can use this information to decide whether or not highlighted objects should be in the final image. In many cases we saw participants notice highlighted edges and consciously decide that they felt like it was okay to keep this potential noise.

7 Conclusion

We presented a set of design goals for encouraging rapid prototyping in the camera, as well as three example interfaces that target a set of important photographic concepts, each with their unique challenges. A common thread amongst these interfaces is the use of computation to generate a visualization that we display to the user in real-time. These visualizations provide the user with an alternative context-aware representation of the image that helps focus their awareness on particular aspects of the image. We also find that these representations help to break down the search space into a number of discrete options, making it less intimidating to explore the space of possible options. While we specifically target photography in this chapter, we are generally excited about the space of building guidance directly into creativity support tools in a way that further encourages aspects the design process to be embedded in the creative process.

References

Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., & Norman, M. K. (2010). *How learning works: Seven research-based principles for smart teaching*. Wiley.

Batista, N. (2015). Work is regularly published in lifestyle FSWNB, men's publications all over the world. Photography is difficult. https://fstoppers.com/editorial/photography-difficult-55962.

Block, B. (2013). *The visual story: Creating the visual structure of film*. CRC Press.

Buxton, B. (2010). *Sketching user experiences: Getting the design right and the right design*. Morgan kaufmann.

Cakebread, C. (2017). People will take 1.2 trillion digital photos this year—Thanks to smartphones. https://www.businessinsider.com/12-trillion-photos-to-betaken-in-2017-thanks-to-smartphones-chart-2017-8.

Dubberly, H. (2004). How do you design. A compendium of models.

Fried, O., Shechtman, E., Goldman, D. B., & Finkelstein, A. (2015). Finding distractors in images. In *Proceedings of the IEEE Conference on Computer Vision and pattern Recognition* (pp. 1703–1712).

Ghyka, M. C. (1977). *The geometry of art and life*. Courier Corporation.

Glover, T. L. (2014). *Canon of design: Mastering artistic composition*. Tavis Leaf Glover.

Heyman, S. (2015). Photos, photos everywhere. https://www.nytimes.com/2015/07/23/arts/international/photos-photoseverywhere.html.

Hunter, F., Biver, S., & Fuqua, P. (2015). *Light science & magic: An introduction to photographic lighting*. CRC Press.

E, J. L., Fried, O., & Agrawala, M. (2019). Optimizing portrait lighting at capturetime using a 360 camera as a light probe. In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, UIST"9* (pp. 221–232). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3332165.3347893.

E, J. L., Fried, O., Lu, J., Zhang, J., Mech, R., Echevarria, J., Hanrahan, P., & Landay, J. A. (2020). Adaptive photographic composition guidance. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3313831.3376635.

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies, 12*(1), 3–11

Li, Q., & Vogel, D. (2017). Guided selfies using models of portrait aesthetics. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, pp. 179–190.

Mitarai, H., Itamiya, Y., & Yoshitaka, A. (2013). Interactive photographic shooting assistance based on composition and saliency. In *International Conference on Computational Science and Its Applications* (pp. 348–363). Springer.

National Academy of Sciences-National Research Council D Washington. (1999). *How people learn: Brain, mind, experience, and school*. ERIC Clearinghouse.

Northrup, T. (2014). *Tony Northrup's DSLR book: How to create stunning digital photography*. Tony Northrup.

Plattner, H., Meinel, C., & Weinberg, U. (2009). *Design-thinking*. Springer.

Shu, Z., Hadap, S., Shechtman, E., Sunkavalli, K., Paris, S., & Samaras, D. (2017). Portrait lighting transfer using a mass transport approach. *ACM Transactions on Graphics (TOG), 36*(4), 1

Wei, Z., Zhang, J., Shen, X., Lin, Z., Mech, R., Hoai, M., & Samaras, D. (2018). Good view hunting: Learning photo composition from dense view pairs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5437–5446).

Woolson, R. (2007). *Wilcoxon signed-rank test. Wiley Encyclopedia of clinical trials* (pp. 1–3).

Zhang, M. (2015). There are now 8× more people taking pictures than 10 years ago. https://petapixel.com/2015/12/15/there-are-now-8x-morepeople-taking-pictures-than-10-years-ago/.

Haptic Guidance to Support Design Education and Collaboration for Blind and Visually Impaired People

Alexa F. Siu, Elyse D. Z. Chase, Gene S.-H. Kim, Abena Boadi-Agyemang, Eric J. Gonzalez, and Sean Follmer

Abstract Designers create sketches, diagrams, and other visual media to both externalize a specific design concept as well as to explore design spaces. The largely visual and spatial nature of these diagrams used to support design activities poses several challenges for blind and visually impaired (BVI) designers to participate along sighted peers. This challenge includes the creation of tools to support the teaching of tactile graphics and their collaborative use in the context of design education. In efforts to address several of these challenges, we present PantoGuide, a low-cost system that provides audio and haptic guidance, via skin-stretch feedback to the dorsum of a user's hand while the user explores a tactile graphic overlaid on a touchscreen. This system allows programming of haptic guidance patterns and cues for tactile graphics that can be experienced by students learning remotely or that can be reviewed by a student independently.

1 Introduction

Designers create and share sketches and diagrams throughout the entire design process. These external representations of information serve as cognitive aids for spatial reasoning tasks and help communicate concepts and collaborate with others. For example, concept maps and mind maps are typically used in the define and ideate phases while customer journeys, mockups, and feedback grids are used in the later prototype and test phases. The largely visual and spatial nature of these diagrams used to support design activities poses several challenges for blind and visually impaired (BVI) designers to participate alongside sighted peers. Most of the information is presented through inaccessible means, forcing BVI designers to work through sighted intermediaries, reducing agency and creativity. In this work,

A. F. Siu (✉) · E. D. Z. Chase · G. S.-H. Kim · A. Boadi-Agyemang · E. J. Gonzalez · S. Follmer
Department Mechanical Engineering, Stanford School of Engineering, 416 Escondido Mall, Room 155, Stanford, CA 94305, USA
e-mail: afsiu@stanford.edu

S. Follmer
e-mail: sfollmer@stanford.edu

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_9

we investigate interactive and collaborative diagram communication tools that are accessible and critical for the inclusion of BVI designers.

People who are BVI frequently rely on alternative text and tactile media to access graphical content. Tactile graphics, in particular, are considered most optimal and suitable for presenting spatial graphics to BVI students (Gorlewicz et al. 2018). The ability to effectively read and interpret tactile graphics and charts is an essential part of a tactile learner's path to literacy. However, recent studies have found that BVI students struggle to work with tactile graphics (Beal et al. 2018; Morash and Mckerracher 2014; Penny Rosenblum et al. 2020). In a survey of BVI students in mainstream classrooms, 50% of students reported receiving insufficient instruction on graphics, how to use them and interpret them (Zebehazy and Wilton 2014a). Many teachers of students with visual impairments (TVIs) have also reported their students often rely on individualized instruction and are not always able to use mathematical graphics independently (Zebehazy and Wilton 2014b). One challenge toward increasing access to tactile diagrams is not only supporting tactile access but also supporting teaching through appropriate guidance. This will be important also for supporting design education.

Reading tactile graphics is not immediately meaningful to a novice student, rather it is a skill that requires instruction and training. TVIs need to help students develop interpretation skills by orienting students to the graphics, asking them questions to guide their exploration, and encourage forming a spatial understanding of the graphic (Zebehazy and Wilton 2014a). Teachers report an important aspect of tactile literacy as (/to be) encouraging students to consistently adopt a systematic exploration by having a clear process on how to approach reading a graphic (Penny Rosenblum et al. 2019). This process requires one-on-one instruction to either provide feedback by moving a student's hand, tracing along paths or individualized verbal encouragement, and guidance depending on the student's skills (Fusco and Morash 2015; Muehlbradt et al. 2018; Penny Rosenblum et al. 2018).

Recent work has begun to explore methods in which BVI users can obtain spatially relevant feedback about tactile graphics through audio-augmented interfaces (Catherine M Baker et al. 2014; Fusco and Morash 2015; Landau et al. 2006; Li et al. 2019; Melfi et al. 2020; Shen et al. 2013). These systems enable graphics to not only include Braille labels but also audio annotations. While audio is an effective tool for communicating details on demand and contextual annotations, it is less effective for guidance especially when trying to move a user's attention to specific areas of interest (Wai and Brewster 2002; Wai et al. 2001). These tools have also mostly been explored in asynchronous use without considering collaboration. Moreover, in collaborative group settings, audio can be particularly disruptive (Bennett et al. 2019). In design collaboration, where teams' attention can quickly shift to different areas of a diagram, providing relevant feedback of attentional or deictic cues is important to help orient the user (Pölzer et al. 2013).

Haptic feedback has been used as an alternative to provide directional signals in a variety of scenarios. Wrist-worn vibrational devices have been used for navigation (Scheggi et al. 2014), handheld skin-stretch devices for providing translational and rotational cues (Walker et al. 2019), and commercial haptic devices guidance in exploring data charts (Doush et al. 2010). Some systems have also explored the use of haptic guidance for conveying shapes and movement in videos to BVI students (Muehlbradt et al. 2018). Unlike audio cues, haptic cues can be more discrete and suitable for group settings (MacLean et al. 2017) and can provide better spatial awareness.

In an effort to address several of these challenges, we see the potential for a tool that provides contextual audio and haptic feedback to guide a BVI student in exploring a tactile graphic. This could help support not only tactile literacy in design education but also facilitate collaboration. To investigate the role of guidance in design education, we introduce PantoGuide, a wrist-worn device that provides haptic guidance cues to the dorsal (back) side of the user's hand while the user explores a tactile graphic mounted onto a touch screen for additional audio feedback (Fig. 1a).

PantoGuide has a tactor that stretches the user's skin (skin-stretch) to convey directional cues. While there are many types of haptic feedback, skin-stretch has been used in devices to provide precise directional guidance (Chinello et al. 2017; Norman et al. 2014; Walker et al. 2019; Yem et al. 2015). Moreover, skin-stretch feedback, unlike graspable and handheld haptic devices, does not interfere with a user's typical exploratory procedures which are important for tactile exploration. We discuss the technical implementation of PantoGuide and demonstrate its use in a set of applications that we closely co-designed with our blind co-author. The system showcases the potential for haptic and audio cues to support tactile graphics training with applications in design education and group collaboration. We conclude with a summary of challenges for increasing the inclusion of BVI designers.

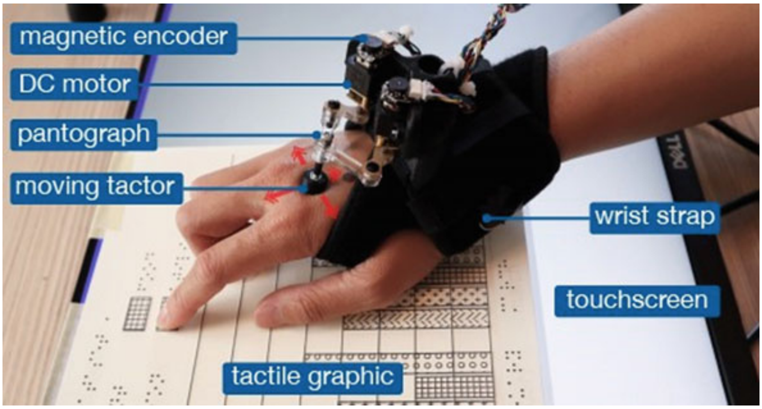


Fig. 1 PantoGuide is a system that provides shaptic and audio guidance cues to a user while exploring a tactile graphic. The device has a tactor in contact with the user that moves in a 2D plane (red arrows) to provide skin-stretch feedback

2 Background and Related Work

2.1 Diagrams to Support Design

Across domains, the use of shared visualizations and diagrams plays a key role in all stages of the design process. This is especially true in the early phases when ideas are vague. Diagrams and sketches (e.g., mind and empathy maps, personas, customer journeys, etc.), make ideas visible, thus helping to communicate ideas (Goldschmidt 1994), create a shared understanding, and enable quicker iterations and insights (Kernbach and Nabergoj 2018). Diagrams support the design process in three ways (Kernbach and Nabergoj 2016): (1) they capture knowledge through their structure and constraints ensuring users define a consistent problem or solution space; (2) they facilitate team discussion by offering representational guidance leading teams' conversations (Suthers 2001); and (3) they help in decision making by externalizing concepts and making new important patterns visible.

Torres et al. conducted a systematic review of different approaches to make diagrams accessible for blind users (Torres and Barwaldt 2019). Their review found that most approaches focused exclusively on either their perception or creation but less so on supporting both. Thus, 85% of reviewed systems relied on auditory feedback for perception and only 15.4% of them considered collaborative features important for design activities. Several systems have been proposed to address accessibility challenges faced by BVI users with specific types of diagrams (Pölzer et al. 2013; Regal et al. 2016; Schnelle-Walka et al. 2014). Schnelle-Walka et al. introduced CoME a system to facilitate easier brainstorming specifically using mind maps for both BVI and sighted users (Schnelle-Walka et al. 2014). Blind and sighted users had access to different digital interfaces with the feedback for blind users being solely auditory. The identified asynchronous review of sessions was a valuable experience for blind users in keeping up with changes taking place during a synchronous meeting. Regal et al. (Regal et al. 2016) investigated the use of *Talking Cards*, relying on a smartphone application and tangible near-field communication (NFC) cards to store information. The tactile cards facilitated clustering information spatially and easier retrieval and navigation. Building on prior design guidelines for accessible design collaboration, in our work, we explore how we can support access to spatial information in both synchronous and asynchronous scenarios.

2.2 Audio-Augmented Tactile Graphics

Audio-based augmented interfaces have been used to provide BVI users automated and spatially-encoded feedback from tactile graphics. The Tactile Graphics Helper (TGH) system (Fusco and Morash 2015) used a mounted camera and machine vision to enable touch responsive audio clarifications on pre-existing tactile graphics. Similarly, Baker et al. presented on a mobile phone-based system leveraging QR code labels to play audio annotations when elements of a tactile graphic were touched (Catherine M Baker et al. 2014). Others have utilized cameras to provide a larger workspace for audio-based feedback about real-world objects (Shen et al. 2013). Some researchers have utilized tactile sheets overlaid on additional hardware such as touch pads (Landau et al. 2006; Li et al. 2019; Melfi et al. 2020) and by using finger-mounted color sensors (Nasser et al. 2018). While audio is an effective tool for communicating details on demand and contextual annotations, it is less effective for guidance especially when trying to direct a user's attention to specific areas of interest (Wai and Brewster 2002; Wai et al. 2001). Audio is quite useful for providing details upon request; however, its use is limited in understanding graphics at a larger scale and also how to navigate an unfamiliar display. In this work, we explore how audio-based approaches can be complemented with guidance through haptic cues.

2.3 Haptic Guidance for Graphics Exploration

The haptic sense is often used for communication in a variety of contexts, as haptics can increase the information provided to the user (MacLean et al. 2017). Our haptic sense comprises both tactile or cutaneous feedback and kinesthetic feedback. Prior works have investigated haptic guidance strategies using both feedback approaches.

Vibrotactile feedback is the result of an actuator that vibrates against the skin. It has been shown that humans can detect differences in frequency and amplitude of the vibrations produced (Pongrac 2006). Several large-scale guidance tasks have used this technology with much success, such as a wearable suit with real-time tactile feedback at joints that served to improve motor learning (Lieberman and Breazeal 2007). In order to aid BVI individuals in locomotion tasks, researchers developed two wrist-worn vibrotactile devices that allow for aid in guidance while navigating outside (Scheggi et al. 2014). However, it has been shown that users prefer the back of the hand and to receive vibrotactile information when completing a precise task of moving a cursor on the screen (Oron-Gilad et al. 2007). Some have also done work exploring the use of asymmetric vibration to provide directional guidance (Rekimoto 2014).

Kinesthetic force feedback requires physical forces to be created and applied to the user. Mendez et al. created a device for sighted users that gave two-dimensional haptic guidance to users exploring volumetric datasets (Méndez et al. 2005). Others created a shape-changing device intended for helping pedestrians navigate in outdoor environments. It was found that this handheld device allowed participants to complete the task more quickly than with a vibrotactile device (Spiers and Dollar 2016).

An additional method used to provide haptic guidance cues is skin-stretch feedback. This type of feedback has been shown to be easily perceived for planar motion guidance on the fingertip (Norman et al. 2014). Others have utilized two different locations of skin stretch simultaneously, such as the thumb and index fingers, in order to combine translation and rotation cues that are easily perceived by users (Walker et al. 2019). Devices with additional degrees of freedom have been created to provide even more information for guidance on the forearm (Chinello et al. 2017) and wrist (Yem et al. 2015). Research on skin-stretch feedback approaches has revealed many examples of devices that provide precise directional guidance with limited actuation.

Overall, there are a variety of haptic technologies that have shown great promise for use in haptic guidance, but there not many in the BVI community. Some researchers have developed systems that incorporate haptic feedback specifically for consumption of graphics and data for BVI students. To explore graphs and charts, Abu et al. utilized a combination of audio feedback with haptic force feedback via a commercial haptic device to allow users to explore charts and data multimodally (Doush et al. 2010). Others explored the use of a kinesthetic device to allow instructors to author audio-kinetic graphics translating visual content from videos for their students, which gave a more guided presentation of content while increasing independence (Muehlbradt et al. 2018). Some researchers adopted magnetic tactile markers as a low-cost solution to render additional information on top of a static graphic (Suzuki et al. 2017).

While there are many types of haptic feedback, skin-stretch has been used in many devices that provide precise directional guidance with limited actuation (Chinello et al. 2017; Norman et al. 2014; Walker et al. 2019; Yem et al. 2015) and as such is ideal for this scenario.

3 Interaction Methods

We closely designed our system with co-author Gene Kim, an engineering college student who is blind and an expert tactile graphics reader, who uses tactile graphics regularly for work (at least once per week). Our co-designer reflected on his own tactile graphics exploration strategies in identifying scenarios where guidance from a teacher had been helpful. These were in agreement with much of the literature that surveys BVI students and teachers on their use of tactile graphics and teaching strategies (Muehlbradt et al. 2018; Penny Rosenblum et al. 2018). We used these tactile graphics and scenarios as a basis to explore the design space of audio-haptic guidance cues through skin-stretch feedback. After these discussions and review, we created a preliminary set of applications and delivered the PantoGuide prototype to our co-designer. Due to COVID-19, we worked remotely to set up the system, iterate through changes, and collect feedback.

3.1 Teaching Scenarios

From literature and our co-design discussions, we envision two different teaching contexts where guidance could be used: (1) synchronous and (2) asynchronous instruction. We explore these interactions in a teaching context and discuss their applicability for remote collaboration in future work.

In the ***synchronous scenario***, guidance could facilitate remote instruction or the instruction of multiple students together. A TVI could be teaching remotely and a guidance tool teleoperated by the teacher, to provide feedback, as the student explores the tactile graphic; therby mirroring a one-on-one in-person teaching scenario. This scenario could also apply when a teacher has multiple students to support and a guidance tool could serve the teacher in providing instruction to all students simultaneously.

In the ***asynchronous scenario***, the tactile graphic guidance could be pre-recorded and a student would be able to review the tactile graphic and annotations at any time and at the speed the student desires. In this scenario, the guidance could provide more self-directed learning opportunities for tactile learners to review material independently.

3.2 Guidance Strategies

Through reflections withour co-author about tactile graphics where teacher guidance was helpful for facilitating understanding, we identified two guidance strategies: (1) point-to-point and (2) continuous.

With ***point-to-point*** guidance, the device starts from an initial resting position—which is located in the center of its reachable workspace. From there, it can move in a direction to give guidance cues (such as the cardinal and intercardinal directions). This allows the user to be guided to and from landmarks around the graphic in order to better understand the content as well as guide the user to areas of particular interest.

With ***continuous*** guidance, the device's cues are constantly being displayed to the user without any need to reset to a resting position. This can be used in a variety of scenarios, such as creating trajectory shapes, aiding in line following, and informing tactile exploration of movement patterns.

4 Application Demonstrations

To demonstrate these different interactions, two applications were developed—both focused on an asynchronous learning scenario (due to current COVID-19 related restrictions on in-person user testing). We specifically looked at two types of graphics: (1) a bar chart and (2) a physics diagram. While here we focus on these specific forms of spatial data, the interactions we investigate can apply to diagrams in the design context.

The first tactile graphic demonstration utilizes a bar chart (Fig. 2a) and automatically guides the user through a systematic exploration of the graph using point-to-point feedback and audio. For example, when the user identifies the *y*-axis values, the system provides verbal instruction (e.g., "locate the *x*-axis") and provides haptic cues that stretch the skin in the desired direction.

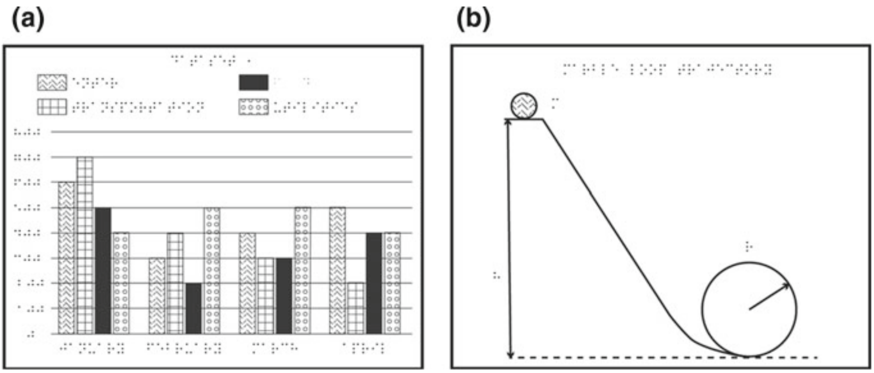


Fig. 2 We demonstrate the use of PantoGuide in two applications: **a** providing point-to-point directional guidance cues when a user explores different elements of a tactile bar chart, and **b** providing continuous guidance on the trajectory of a moving marble

The second application involves a tactile graphic of a marble rolling down an inclined plane and through a loop the loop (Fig. 2b). In this case, the haptic cues provide continuous feedback to trace the marble's trajectory on the user's hand. This feedback provides the user with an overview of the marble's motion as the user explores the graphic.

5 Technical Implementation

The haptic device consists of a pantograph with a tactor on the end effector, which comes into contact with the back of the user's hand to provide skin-stretch cues. The pantograph is a common mechanism for providing two-dimensional haptic feedback as it consists of a simple five-bar linkage, and it is forward and inverse kinematics are well documented (Campion 2005).

Two Pololu 6 V 100:1 micro-metal DC gearmotors are controlled by a Teensy LC microcontroller through a DRV8835 motor driver. Rear shaft-mounted magnetic encoders (12 CPR) allow for relative angle measurements of each motor. The angle of each motor is regulated using independent PID controllers. Inverse kinematics (Campion 2005) are used to determine the necessary motor angles to achieve the desired x–y positions. To remain within the reachable workspace, the total movement of the device is limited to a 14 mm square located 23.5 mm away from the base of the device.

The device connects over USB serial to an application developed with Unity 3D. The application is run on a multi-touch screen. A tactile graphic is placed on top of the touchscreen depending on the application. The user can interact with the applications through two gestures: (1) one finger double tap for reading audio labels, and (2) two finger double tap for selection. The selected gesture set allows the user to freely explore and remain in contact with the tactile graphic while minimizing unintentional input.

6 Design Recommendations

Based on our initial user feedback with our co-designer (who tested the device and the applications), we present some design recommendations for future design iterations.

Point-to-point directional cues can provide useful guidance
Point-to-point feedback worked well in guiding the user to systematically explore the different bar chart elements (Fig. 1b). This was the preferred use scenario for our co-designer. He expressed interest in use of point-to-point feedback for more complex mathematical graphics. For example, in a polynomial curve where a student might need to understand different points in the curve such as the local maximum and minimum.

Continuous feedback can give an overview of the shape, but needs enough space to be understandable
The limited workspace, 2 cm^2 area, was not large enough to successfully convey complex motion trajectories. In the second application, continuous haptic guidance was used to communicate the trajectory of the marble (Fig. 1c). Our co-designer commented that motion guidance had been helpful during in-person instruction with a teacher to get trajectory overviews. However, with the current system, it was hard to distinguish complex motions such as the loop (Fig. 1c).

Haptic and audio cues need to be tightly coordinated
The haptic guidance cues helped complement the audio, because they prevented overloading the auditory channel. Our co-designer commented on how important it was for the two feedback modalities to be tightly coordinated so they are not competing for attention. When the haptic cue was not coordinated with the audio, our co-designer had trouble focusing on either one.

Guidance cues need to be more perceivable under high mental load
Our co-designer described how the guidance cues were less perceivable when moving his hands compared to the static case. He tried different tactor variations: a square surface with twice the surface area (1.9 cm^2) and a pointed tactor. The larger surface area greatly improved the guidance cue's perceivability, likely due to the increased amount of skin being stretched.

7 Future Work and Remaining Challenges

Participatory design has been at the core of designing the system presented here. However, even within this process, it is apparent how several of our traditional rapid prototyping and design practices—both artifacts and collaborative tools—often fail to consider the different ways in which BVI users interact with and consume information. With collaboration being a central tenet in design activities, moving forward, we need to consider how both the tools and methodologies that we use enable collaboration with people with disabilities.

Toward contributing to a future of inclusive design practices, with PantoGuide, we have started investigating the use of haptic guidance to support instruction of tactile graphics. While here we explored two general types of spatial information, our ultimate goal is to apply these findings to the design domain to support both education and collaboration. Our two application demonstrations were grounded in an education context. In a collaboration context, we envision PantoGuide as helping to provide important contextual cues around deictic gestures used when referencing design artifacts, in particular with (/in the case of) diagrams such as mind maps and affinity diagrams. The point-to-point feedback would be suitable for providing context relevant to deictic gestures while the continuous feedback would be helpful for supporting content. Regardless of the system outcome, it is important to consider not only the effectiveness of the information being communicated but also the social factors (Shinohara et al. 2018). For example, understanding whether this type of system would affect or support the group dynamics with both sighted and blind, whether BVI users would feel their agency is impacted, whether the feedback is timely and socially acceptable, etc.

Another domain where we see a lot of opportunities and need to close the accessibility gap is in the design tools used to produce artifacts. We see the potential for a new set of tools that both educate on accessibility issues and afford inclusion. Design tools (e.g., Figma, Sketch, storyboards) and methodologies are at the foundation of the design journey. Designers and developers employ these tools at the early stages of development to iterate and test potential solutions with stakeholders. However, these are often incompatible with assistive technologies. For example, lo-fi interactive wireframes created through vector design tools such as Figma, only offer visual affordances for interaction. These lo-fi prototypes cannot be tested with BVI users that rely on screen readers or with users with limited mobility that rely on voice and keyboard input. As a result, accessibility is only considered at the end-stage, which has led to the myriad of accessibility band-aid solutions used today. To truly design with accessibility in mind from the beginning, it is crucial for designers and developers to have the right tools to be able to iterate and design with all stakeholders as early testers. Future areas of work could investigate ways design tools would enable even non-domain experts to iterate as much on accessible UIs as we do now with traditional UIs.

8 Technical Limitations

Several technical improvements need to be made in future iterations of PantoGuide. A main point of feedback in the discussion is the need for more perceivable cues especially when working on a task. Future improvements to the prototype will need to amplify the skin-stretch feedback by increasing the size of the workspace, creating a higher friction tactor, and/or using stronger motors that can deliver higher torque.

Authoring the guidance cues and feedback is another important aspect to address. This becomes even more challenging when designing multimodal cues distributed across audio, haptics, and vision. We will investigate authoring interfaces to annotate tactile graphics using both audio and haptic cues. Currently, we specifically program the different applications for each tactile graphic. An authoring interface could also allow a user to adjust the guidance level and cues depending on literacy skills or context.

9 Conclusion

Design activities are often inaccessible to BVI users. Addressing challenges that prevent inclusion are even more important in design, so that BVI users can be designers and advocates of solutions that address their own unique needs. In this work, we investigate a guidance system to support design education and collaboration. With PantoGuide, we demonstrate the functionality of a novel system that combines skin-stretch feedback on the hand's dorsum with audio feedback, as the user explores a tactile graphic overlaid on a touchscreen through several co-designed applications. We discuss feedback, challenges, and opportunities for improvement. Overall, the system shows promise toward creating tools that support tactile literacy both through synchronous and asynchronous learning.

References

Baker, C. M., Milne, L. R., Scofield, J., Bennett, C. L., & Ladner, R. E. (2014). Tactile graphics with a voice: Using qr codes to access text in tactile graphics. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility* (pp. 75–82).

Beal, C. R., & Penny Rosenblum, L. (2018). Evaluation of the effectiveness of a tablet computer application (app) in helping students with visual impairments solve mathematics problems. *Journal of visual impairment & blindness, 112*(1), 5–19.

Bennett, C. L., Stangl, A., Siu, A. F., & Miele, J. A. (2019). Making nonvisually: Lessons from the field. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS'19* (pp. 279–285). New York, NY, USA: Association for Computing Machinery.

Campion G. (2005). The pantograph mk-ii: A haptic instrument. In *The Synthesis of Three Dimensional Haptic Textures: Geometry, Control, and Psychophysics* (pp. 45–58). Springer.

Chinello, F., Pacchierotti, C., Bimbo, J., Tsagarakis, N. G., & Prattichizzo, D. (2017). Design and evaluation of a wearable skin stretch device for haptic guidance. *IEEE Robotics and Automation Letters, 3*(1), 524–531.

Doush, I. A., Pontelli, E., Son, T. C., Simon, D., & Ma, O. (2010). Multimodal presentation of two-dimensional charts: An investigation using open office xml and microsoft excel. *ACM Transactions on Accessible Computing (TACCESS), 3*(2), 1–50.

Eppler, M. J., & Kernbach, S. (2016). Dynagrams: Enhancing design thinking through dynamic diagrams. *Design Studies, 47*, 91–117.

Fusco, G., & Morash, V. S. (2015). The tactile graphics helper: Providing audio clarification for tactile graphics using machine vision. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (pp. 97–106).

Goldschmidt, G. (1994). On visual design thinking: The vis kids of architecture. *Design Studies, 15*(2), 158–174.

Gorlewicz, J. L., Tennison, J. L., Palani, H. P., & Giudice, N. A. (2018). The graphical access challenge for people with visual impairments: Positions and pathways forward. In *Interactive Multimedia-Multimedia Production and Digital Storytelling*. IntechOpen.

Kernbach S., & Nabergoj, A. S. (2018). Visual design thinking: Understanding the role of knowledge visualization in the design thinking process. In *2018 22nd International Conference Information Visualisation (IV)*, pages 362–367. IEEE, 2018.

Landau, S., Russell, M., & Erin, J. N. (2006). Using the talking tactile tablet as a testing accommodation. *RE: View, 38*(1), 7.

Li, J., Kim, S., Miele, J. A., Agrawala, M., & Follmer, S. (2019). Editing spatial layouts through tactile templates for people with visual impairments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–11).

Lieberman, J., & Breazeal, C. (2007). Tikl: Development of a wearable vibrotactile feedback suit for improved human motor learning. *IEEE Transactions on Robotics, 23*(5), 919–926.

MacLean, K. E., Schneider, O. S., & Seifi, H. (2017). Multisensory haptic interactions: Understanding the sense and designing for it. In *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations-Volume 1*, pp. 97–142.

Melfi, G., Müller, K., Schwarz, T., Jaworek, G., & Stiefelhagen, R. (2020). Understanding what you feel: A mobile audio-tactile system for graphics used at schools with students with visual impairment. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–12).

Méndez, E., Yoshida, S., Noma, H., Lindeman, R. W., Yanagida, Y., Masaki, S., & Hosaka, K. (2005). Haptic-assisted guidance system for navigating volumetric data sets. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics Conference* (pp. 531–534). IEEE.

Morash, V., & Mckerracher, A. (2014). The relationship between tactile graphics and mathematics for students with visual impairments. *Terra Haptica, 4*, 13–22.

Muehlbradt, A., Atreya, M., Guinness, D., & Kane, S. K. (2018). Exploring the design of audio-kinetic graphics for education. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (pp. 455–463).

Nasser, A., Zhu, K., & Rao, P. V. M. (2018). Colortact: A finger wearable audio-tactile device using customizable color tagging. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (pp. 178–181).

Norman, S. L., Doxon, A. J., Gleeson, B. T., & Provancher, W. R. (2014). Planar hand motion guidance using fingertip skin-stretch feedback. *IEEE transactions on haptics, 7*(2), 121–130.

Oron-Gilad, T., Downs, J. L., Gilson, R. D., & Hancock, P. A. (2007). Vibrotactile guidance cues for target acquisition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 37*(5), 993–1004.

Penny Rosenblum, L., Cheng, L., & Beal, C. R. (2018). Teachers of students with visual impairments share experiences and advice for supporting students in understanding graphics. *Journal of visual impairment & blindness, 112*(5), 475–487.

Penny Rosenblum, L., Ristvey, J., & Hospitál, L. (2019). Supporting elementary school students with visual impairments in science classes. *Journal of Visual Impairment & Blindness, 113*(1), 81–88.

Penny Rosenblum, L., Cheng, L., Zebehazy, K., Emerson, R. W., & Beal, C. R. (2020). Teachers' descriptions of mathematics graphics for students with visual impairments: A preliminary investigation. *Journal of Visual Impairment & Blindness, 114*(3), 231–236.

Pölzer, S., Schnelle-Walka, D., Pöll, D., Heumader, P., & Miesenberger, K. (2013). Making brainstorming meetings accessible for blind users. In *AAATE Conference, 2013*.

Pongrac, H. (2006). Vibrotactile perception: Differential effects of frequency, amplitude, and acceleration. In *2006 IEEE International Workshop on Haptic Audio Visual Environments and Their Applications (HAVE 2006)* (pp. 54–59). IEEE.

Regal, G., Mattheiss, E., Sellitsch, D., & Tscheligi, M. (2016). Talkingcards: Using tactile nfc cards for accessible brainstorming. In *Proceedings of the 7th Augmented Human International Conference 2016* (pp. 1–7).

Rekimoto, J. (2014). Traxion: A tactile interaction device with virtual force sensation. In *ACM SIGGRAPH 2014 Emerging Technologies* (pp. 1–1).

Scheggi, S., Talarico, A., & Prattichizzo, D. (2014). A remote guidance system for blind and visually impaired people via vibrotactile haptic feedback. In *22nd Mediterranean Conference on Control and Automation*, pages 20–23. IEEE, 2014.

Schnelle-Walka, D., Alavi, A., Ostie, P., Mühlhäuser, M., & Kunz, A. (2014). A mind map for brainstorming sessions with blind and sighted persons. In *International Conference on Computers for Handicapped Persons* (pp. 214–219). Springer.

Shen, H., Edwards, O., Miele, J., & Coughlan, J. M. (2013). Camio: A 3d computer vision system enabling audio/haptic interaction with physical objects by blind users. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 1–2).

Shinohara, K., Wobbrock, J. O., & Pratt, W. (2018). Incorporating social factors in accessible design. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 149–160).

Spiers, A. J., & Dollar, A. M. (2016). Outdoor pedestrian navigation assistance with a shapechanging haptic interface and comparison with a vibrotactile device. In *2016 IEEE Haptics Symposium (HAPTICS)* (pp. 34–40). IEEE.

Suthers, D. D. (2001). Towards a systematic study of representational guidance for collaborative learing discourse. *Journal of Universal Computer Science, 7*(3), 254–277.

Suzuki, R., Stangl, A., Gross, M. D., & Yeh, T. Fluxmarker: Enhancing tactile graphics with dynamic tactile markers. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 190–199, 2017.

Torres M. J. R., & Barwaldt, R. (2019). Approaches for diagrams accessibility for blind people: A systematic review. In *2019 IEEE Frontiers in Education Conference (FIE)* (pp. 1–7). IEEE.

Walker, J. M., Zemiti, N., Poignet, P., & Okamura, A. M. (2019). Holdable haptic device for 4-dof motion guidance. In *2019 IEEE World Haptics Conference (WHC)* (pp. 109–114). IEEE.

Yem, V., Otsuki, M., & Kuzuoka, H. (2015). Development of wearable outer-covering haptic display using ball effector for hand motion guidance. In *Haptic Interaction* (pp. 85–89). Springer.

Yu, W., & Brewster, S. (2002). Comparing two haptic interfaces for multimodal graph rendering. In *Proceedings 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. HAPTICS 2002* (pp. 3–9). IEEE.

Yu, W., Ramloll, R., Brewster, S., & Ridel, B. (2001). Exploring computer generated line graphs through virtual touch. In *Proceedings of the Sixth International Symposium on Signal Processing and its Applications (Cat. No. 01EX467)* (Vol. 1, pp. 72–75). IEEE.

Zebehazy, K. T., & Wilton, A. P. (2014a). Quality, importance, and instruction: The perspectives of teachers of students with visual impairments on graphics use by students. *Journal of visual impairment & blindness, 108*(1), 5–16.

Zebehazy, K. T., & Wilton, A. P. (2014b). Charting success: The experience of teachers of students with visual impairments in promoting student use of graphics. *Journal of Visual Impairment & Blindness, 108*(4), 263–274.

Capturing Collaboration with Interaction Dynamics Notation

Mark Roman Miller, Will Gutzman, Jeremy N. Bailenson, Ade Mabogunje, and Neeraj Sonalkar

Abstract Interaction Dynamics Notation (IDN), introduced by Sonalkar et al. in (Int J Des Creat Innov 1(2), 93–108, 2013), is a notation system for team collaboration. In the years since publication, it has garnered much interest and follow-up work. In this chapter, we consolidate much of the known work performed using IDN. We begin with an introduction to IDN, first abstractly through its description, origin, and symbol set, then concretely through short portions of team discussion. Then, we review research that references IDN, first detailing what is known about and through IDN, and second how IDN has been perceived and how it has inspired other work. We also report descriptive statistics of the symbols produced during collaboration, revealing patterns that help characterize the sequence of IDN symbols. Finally, we discuss two directions for future work.

1 What is IDN?

1.1 Interaction Dynamics Notation

Interaction Dynamics Notation (or IDN as it is abbreviated) is a symbolic system for annotating team interactions that highlight the development of conversational

M. R. Miller (✉) · W. Gutzman · J. N. Bailenson
Virtual Human Interaction Lab, Department of Communication, Stanford University, McClatchy Hall, Room 411, Stanford, CA 94305-2050, USA
e-mail: mrmillr@stanford.edu

J. N. Bailenson
e-mail: bailenso@stanford.edu

A. Mabogunje · N. Sonalkar
Center for Design Research, Stanford University, Building 560, 424 Panama Mall, Stanford, CA 94305, USA
e-mail: ade@stanford.edu

N. Sonalkar
e-mail: sonalkar@stanford.edu

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_10

outcomes through interpersonal force dynamics. The notation surfaces how individual contributions are supported, blocked, built upon, questioned, or otherwise developed through each individual on the team responding to the previous person, when the team is collaborating to deliver product concepts, make design decisions or develop problem-solving strategies. IDN converts team conversations into a sequence of categories (or symbols when visualized) that can be further analyzed quantitatively to understand how individuals perform on a team and contribute to the results that are relevant to its success. It thus forms the basis for a quantitative analysis of the quality of teamwork (Sonalkar et al. 2016a, b, Developing Instrumentation).

The Interactions Dynamics Notation was developed as a system to analyze the co-creation of product concepts in design teams (Sonalkar et al. 2013). It is based on the Force Dynamics framework from Cognitive Semiotics and the principles of improvisational behavior in improvisational theater and music. The Force Dynamics framework (Brandt, 2004; Talmy, 1988) posits that there are forces in a narrative that act upon the agents in the narrative and that can be visualized. IDN adapts the Force Dynamics framework to visualize the forces in conversational interactions. The symbol categories of Move, Block, Overcoming, and Deflection are derived from Force Dynamics. The symbol categories of Yes-And, Support, Question, and Humor are derived from principles used by theater or music performers to practice and train for improvisational teamwork. These principles have been observed to aid creative collaboration in a team (Schinko-Fischli, 2018). The rest of the symbol categories in IDN were added to account for behaviors seen in design team data, which had an observed impact on the concepts being developed. The entire symbol category set of IDN is listed in Table 1.

1.2 IDN in Context

"Design is a social activity" (Sonalkar et al. 2016a, b, "Visualizing…" p.1). Every step of the design process is marked by humans interacting and communicating with each other. Some examples of IDN in context may illustrate and bring greater clarity to both its definition and applications. Examples here are from work that is further described in the Methods section.

In these studies, the most frequently used symbols were Move, Question, Support, and Yes-And. Listed below is a table depicting a typical IDN interaction which showcases these symbols in action.

Time	Speaker	Transcription	IDN symbol
02:58	C	I was going to say, wasn't that the always-on millennials, wasn't that the target audience for the company in general? And they said 'A' was like more of a health and fashion choice?	Question
03:12	B	What I remember, and I could be wrong, the company wants to hit this target–	Move

(continued)

(continued)

Time	Speaker	Transcription	IDN symbol
03:18	C	Yes! Right	Support
03:18	B	The always-on millennial. 'A' aligned very well with that group. But then 'B' was not so much	Yes-And
03:26	C	It was like with the urban audience or something–	Yes-And
03:27	B	Yeah, that's it. Good job	Support

In this interaction, participants were given a scenario in which they had to decide to invest in product A or B. A was a tattoo which measures various physiological functions including heart rate and temperature, while B was a necklace that monitors one's mood and emotions. The three participants each received slightly different briefings. In many teams, time was spent getting on the same page about background information. Because of the different briefings, some teams disagreed about the study scenario. Phrases like "Well that's not what I read on my sheet" were fairly frequent. Arguments were never intense, as participants quickly realized that they had received different briefings. However, there was sometimes a qualitative sense of conflict while teams discussed the background information upon which the study rested.

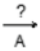
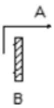
The transcript also shows a section of the conversation in which participants B and C were establishing a common understanding of the tattoo and necklace. Here the IDN symbols indicate that the conversation is generally cooperative and supportive. The participants give and take space, listening to and interacting with what their partner has said. This is shown by the presence of Support and Yes-And utterances, which necessarily imply active listening.

A less frequent but nonetheless important IDN symbol is a Block. A Block is the opposite of a Support. When a person posits an idea and someone else responds negatively, it is coded as a Block. Below is a table showing an example.

Time	Speaker	IDN	Transcription
09:24	A	Move	A nice VR system would be great like you could do a meeting, VR work in the virtual reality environment, so I think VR is a must in the self-driving car
09:44	C	Block	Or it's just– I feel like a projected screen would be good, it doesn't necessarily have to be VR, but some type of console and screen to be able to work because some people don't feel good wearing VR devices
10:01	A	Overcoming	Yeah, I'm assuming the VR has kinda matured, where it's no longer an issue

The task in question was to design a self-driving car with an emphasis on what the driver could do now that their hands were free. The presence of a Block indicates there is some disagreement in the conversation. Reading the transcript will reveal that Participant A believes incorporating virtual reality into the design is a necessity.

Table 1 IDN symbol set (Jablokow et al. 2019a, b)

Symbol	Name	Description	Example
	Move	A Move indicates that a participant has made an expression that moves the interaction forward in a given direction	A: I need to buy Legos (at) home. Think about how therapeutic it would be
	Question	A Question indicates an expression that elicits a Move	A: Where should we start?
	Yes-And	Yes-And accepts the content of the previous Move and adds on to it	A: What about… if we made a toy that incorporates girls and boys playing together… C: I think that's a good point to have some sort of educational point in it
	Support	Support indicates that the speaker understands and/or agrees with the previous Move	C: Safe and entertaining (bending forward to write) B: Safe and entertaining, yes
	Block	Block indicates an obstruction to the content of the previous Move	B: Maybe have something which looks like a computer, but you can just type your name or do a simple math… C: Er, but I don't know, I mean, considering the age segment we are targeting 3–7 years
	Block-support	Block-support indicates an acceptance of a Block by another person	A: But that's also, I think that's already done C: Yeah, it's already there B: Ok
	Overcoming	Overcoming a Block indicates a speaker was able to overcome the Block and persist on course of the original Move	C: Er, but I don't know, I mean, considering the age segment we are targeting 3–7 years B: So, 7 years they go to school, they would learn A, B, C, right?
	Deflection	A speaker can deflect the Block with a Move that presents an alternative direction for the interaction	B: So, when you say we need to divide the age group, but you cannot have like 3, 4, 5 A: No, no of course not, but I mean you might have a few different (concepts)
	Humor	Humor indicates instances of shared laughter in teams	A: I don't know I probably would have swallowed but… (All of them laugh)

(continued)

Table 1 (continued)

Symbol	Name	Description	Example
	Ignored	Ignored indicates that a person was heard but not responded to by the rest of the team	A: We could build Lego forts and have little people in them B (looks at A, and then turns to C): I was thinking we could do a Lego zoo
	Silence	Silence is a state in which none of the participants speak as they are engaged in individual-level activities	
	Ambiguous	Ambiguous is used when it is not clear what a person said	A: Shall we finalize on Lego sandbox? X: (inaudible)

Participant B disagrees, stating that a simple screen will do. The final utterance is A's response, coded as Overcoming because he is arguing against the Block and therefore supporting his original utterance. After a Block, Overcoming, Deflection, or SupportForBlock are often coded.

While a Block means that participants are disagreeing with each other, there is another combination of utterances that could indicate participants are not listening. When a Move is followed by another Move, and not a Support, Question, or Yes-And, this can suggest that participants are not responding directly to the previous speaker. The table below shows one such example.

Time	Speaker	IDN	Transcription
00:33	A	Move	I was just thinking super broadly about having something where it's sort of like a projected screen and your form of response is not like using your hands, but like nodding, more of that visual input. The way I think about it is like, communicating, you have your input, and then you have your output, so it's sort of like your input can be audio, so maybe some kind of voice telling you things based on maybe like tweets–
01:28	B	Move	Okay, so–
01:29	A	Move	Or, how we–
01:31	B	Move	Well, I was thinking of cool technology but just like, I feel like I have to think about who are the people using these self-driving cars, and even though they're self-driving, do they have someone in the driver seat, or are all the seats centered and there's a computer taking you from point A to point B and you don't have to face the road

The design task here was to ideate features for a self-driving car. In reading the two longer utterances in the above table, one can see that they are mostly unrelated. Participant A first talks about various ways of interfacing with the car. Then, Participant B cuts off Participant A. Participant A tries to keep speaking, but then cedes the floor to Participant B, who goes on to talk about the target consumer as well as whether or not the car will have someone in the driver seat. This was not coded as a Block because Participant B did not comment on Participant A's statement. Instead, he pushed what he wanted to say to the forefront of the conversation. A Block necessitates listening and responding to another's utterance in order to argue against it. While no Blocks occurred in this segment, there was minimal acknowledgement. When using IDN to analyze a team, one must look out for less obvious instances of potentially problematic behaviors in teamwork.

2 Related Work

In order to understand IDN better in both its value as a theory and its role in research, we reviewed the articles that cite the original IDN paper by Sonalkar et al. (2013). At the time of writing, Google Scholar provided a total of 43 research articles that reference Sonalkar et al. (2013). Upon investigation, one paper was listed twice, so a total of 42 articles were reviewed.

In reviewing the articles, one distinction becomes readily apparent. Some works engage with IDN categories as a way of describing and annotating behavior and use the results of IDN coding to draw conclusions about the teams, behaviors, and interactions studied. Other works refer to the conclusions drawn or assumptions made by IDN-based research. A primary difference between the two is how much of IDN is abstracted away. If a researcher codes ten hours of conversations and shows that ideation occurs around particular IDN symbols, there is very little abstraction. In comparison, if IDN is standing in as a reference example of team interaction notations, we learn less about IDN's links with other constructs of interest and more about how it is perceived by other researchers.

Within the first class, there are a total of 10 articles. These kinds of papers are valuable because they develop a more nuanced and more empirically grounded theory of behavior through the lens of IDN.

Within the second class, there are then 32 papers. While these papers do not use IDN to more deeply understand human behavior, it is still insightful to see how the research community refers to IDN.

What features of IDN are mentioned in related work sections? Patterns that appear in these references can inform future directions of work.

3 Use of IDN

In Sonalkar et al. (2014a, b), (2016a, b), IDN is used to study the interactions between professional design instructors and industrial design graduate students. In addition to IDN, another variable of interest was *professional vision* (Goodman 1994), a system of seeing and interpreting, specific to a profession, and applied to events in domains of interest to those groups.

The primary element in the conversations was Move, occurring about 45–55% of the time across sessions. Episodes of professional vision were likely to start with a Move or Question (85%). This finding is intuitive, as those symbols mark larger changes in topic. Most episodes (80%) were continued by the student through Moves, Questions, and Supports. The remaining 20% were expert monologues. This corroborates the well-established fact from learning science that most episodes of learning are interactions.

Sonalkar et al. (2016a, b Developing) iterate on the symbol set for IDN coding and consider some questions involving the coding process. The shift in symbol set came about by investigating the reliability of IDN coding between raters. After deciding the most effective measure of reliability was weighted Levenstein edit distance, work was done to increase accuracy while retaining the value of categorizing.

In Sonalkar (2016a, b Diagnostics), the primary method of analysis was correlation between frequency of IDN sequences and outcome scores, namely novelty, and utility of the concepts. Instances of Humor had a positive correlation ($r = 0.55$) with utility of the team's idea, and Yes-Ands had a negative correlation ($r = -0.83$) with novelty of the idea. Considering the ubiquity of the Yes-And as a signal for improvisation and creativity within the design teamwork literature, this finding does not match previous work. This result might point to the difference between the use of Yes-And for generating ideas, i.e., creativity as behavior, and its impact on the quality of ideas, i.e., creativity as judgment of content. The third sequence that was reported was a "dialectic episode," which consists of more than two consecutive Block and Overcoming responses indicating that participants are engaged in an argumentative dialectic. The count of dialectic episodes correlated positively ($r = 0.9$) with utility of the submitted idea.

Sonalkar et al. (2017 Design Whodunit) reported an analysis of four teams of three individuals each investigating the sequence of IDN symbols leading up to an idea. Some utterances were coded as "idea expressions" if they contained a verbally expressed change to the original material, analysis, or concept.

For each team, a path was generated to predict the symbol sequences that lead to an idea. For two of the teams, there was no path that predicted ideation at a rate of more than 50% accuracy. However, for two teams, there were paths. For one team, the pattern was "Silence or Question" into any symbol into anything other than "Support, Humor, Silence, Question, or Move." For the other team, the pattern was "Yes-And" into "Move" into "Yes-And." It is unclear how many sequences followed these patterns, and the degree to which these findings may have been expected by chance.

Koenig and Lim (2018) use a machine learning algorithm called a Conditional Random Field to automatically classify transcripts of utterances into IDN symbols. Based upon hand-selected text features, accuracy was at 79.0%. This result is better than a baseline in which Move is always predicted, which has an accuracy of 53%, and also better than a simple logistic regression, which achieves 69% accuracy. This demonstrates some signal between symbols is visible in the words being said, even without prosody and nonverbals.

Jablokow et al. (2018) give an effective measure of the possible inter-rater reliability by noting their weighted Levenshtein's ratio, the ratio of sequence length to maximum common subsequence (Sonalkar et al. 2016a, b, Developing Instrumentation…) was 0.80. In their study consisting of 14 teams with 3 participants each, they found that of all IDN symbols, only deflection was moderately correlated ($r = 0.75$) with the total number of idea occurrences ($p < 0.05$). In sequences associated with ideas, Yes-And occurred almost twice as often as the rest of the IDN symbols (14% versus 8%). In sequences associated with unique ideas, Move and Question occurred more often than Yes-And.

Jablokow et al. (2019a, b) report null results with all analyses involving IDN. In a study of seven teams, they found no links between either team or individual IDN symbol counts and team or individual KAI scores.

Endrejat et al. (2019) created an utterance coding system called "Analyzing Idea Finding Interactions" (AIFI). In this work, they explicitly mention IDN as the source of one symbol in their coding scheme, "blocking." There are also symbols named "support" and "humor" in the scheme. The scheme is coded into two processes, ideas, and team spirit, with facilitating and inhibiting symbols to each. There are also neutral symbols, like logistics or turn-taking.

In this study consisting of 20 teams of 4, idea facilitating behaviors were correlated positively with perceived effectiveness ($r = 0.34$; $p = 0.03$) and more generated ideas ($r = 0.31$; $p = 0.05$). Team spirit facilitating interactions, such as support and humor, also correlated positively ($r = 0.56$; $p < 0.001$) with team members' perceived effectiveness ratings, which came as a surprise to the paper's authors. There were also correlations between idea facilitating and team spirit facilitating ($r = 0.54$; $p < 0.001$), and also between idea facilitation and idea inhibitive behaviors ($r = 0.28$; $p = 0.08$).

Menold et al. (2020) investigated the effect of a human or computer turn-taking facilitator. Both the human and the computer system were to alert a participant if the participant had not participated in the conversation for a period of time. Symbols related to the facilitation were added to the default IDN categories.

When the IDN symbol sequences were fed into a Hidden Markov Model (HMM), the result was a primarily diagonal transition matrix, due to the high temporal resolution (0.2 s) compared to turn length. Therefore, there were many symbols of the same utterance in sequence, and high accuracy can be achieved merely by predicting the same symbol as seen in the previous step. Because the HMM was limited to only 4, there was still some information provided by which symbols compress in which symbols, but it is unclear whether there is structure between the two conditions (computer system and human facilitation).

4 References to IDN

It is insightful to catalog the ways in which IDN is referenced outside of the use of the symbol set itself. In some cases, IDN becomes an example of a class of papers or justifies a paper's assumptions. How researchers summarize the work can inform what aspects of the research are most memorable and useful to the community.

In reviewing previous work, we settled upon eight different categories. These categories are not exhaustive nor were decided upon a priori. Rather, these categories provide useful insight into the role IDN has played in research following its publication.

The first three sections come from the name Interaction Dynamics Notation itself. First, IDN is a *notation*. Like letters representing sounds, or music notes representing pitches, notations condense rich but transient information into a stricter but more stable format. The title of the original article (Sonalkar et al. 2013) calls IDN a "visual representation", echoed in Cagan et al. (2013) and Feng (2013). Mabogunje et al. (2016) refer to IDN as a "notational system" consisting of "twelve symbols for notating design interactions." Adly Taha et al. (2019) refer to IDN as a "coding scheme … to analyze protocol data", and compare it to another visual system called Linkography (Goldschmidt 2014). They highlight the value of a stricter and more stable format as more conducive for computers to work with, opening up the possibility of automatic coding.

Cash and Štorga (2015) contrast their own network-based representation against "notation-based approaches," of which IDN is an example. In their work, connections are not limited to the alphabet of moves but rather can link based upon "various elements in the design process, moves, activities, ideas, objects, or artefacts, across multiple levels of granularity." This point highlights the trade-off due to the degree to which data is structured.

Second, the notation describes *dynamics*. Dynamics, in contrast to statics, involves change over time, and therefore change ordered within time. This aspect of IDN is described in some works as "sequential" (Paulsen et al. 2013), "time-scale" or "critical moments" (Georgiev and Taura 2014), "temporality" (Luck 2014), and "moment-to-moment" (Mabogunje et al. 2016).

Paulsen et al. (2013) and Georgiev and Taura (2014) both referred to IDN while discussing future work investigating the sequential features of their own notation system. Luck (2014) refers to IDN as research investigating "temporality in creative interactions" motivating their own study of "how aesthetic qualities of the built [architectural] form were brought into conversation." Dinar et al. (2015) refers to IDN under the section heading of "Team Dynamics" and places IDN alongside ethnographic studies of industry teams in the 1980s. Mabogunje et al. (2016) summarizes IDN's new insights enabled by "track[ing] the interaction between designers on a moment-to-moment basis."

Thirdly, the dynamics are *interactions*. It is an important assertion in the concept of IDN that concept generation is an interaction between team members. Both Luck (2014) and Dinar et al. (2015) refer to IDN not only as dynamic, as mentioned above, but also team-based. Wulvik et al. (2017) describe IDN as "captur[ing] the interpersonal interactions." Martinec et al. (2018, 2019a, b, c, 2020) and Horvat et al. (2020) use the work to highlight that "creative conceptual design tasks such as idea generation or concept selection, are often performed exclusively as team activities" (Martinec et al. 2019a, b, Model Info Proc).

Beyond the three qualities that compose the name Interaction Dynamics Notation, there are qualities of the interactions that are of interest to researchers. Two references use the original IDN study as a description of concept generation behavior. McInnis et al. (2018) refer to IDN in order to demonstrate problem-talk—how a group defines and decides a problem—"use proposals to elicit questions and constraints." Lugnet et al. (2020) point out that the interactions involve facial expressions and nonverbal behavior, which is not commonly studied by designers.

Two references pull out specific symbols from the IDN symbol set. Camburn et al. (2017) echo the "Yes-And" technique originally borrowed from improvisational theatre and highlight its role in the practice of design. Alexander et al. (2019) base some symbols such as Blocking in their design notation upon the blocking and overcoming symbols in IDN.

Finally, there are references that provide no direct insight into the research community's perception of IDN. In five works (Bracken et al. 2019; Jablokow 2019a, b; Sonalkar et al. (n.d); Henderson et al. 2020; Sonalkar et al. 2020), the paper is cited as part of the methodology, indicating the data was collected, but no results are reported. One work (Ju et al. 2016) was a review of work performed at the Center for Design Research at Stanford, the primary affiliation for authors of the original IDN paper. Two papers (Sonalkar 2014a, b; Menold and Jablokow 2019) refer to other aspects of the original paper, in particular, KAI theory and C-K theory.

5 Method

In the interest of space, we are reporting only preliminary results of a study. More substantial work is under review.

In this study, we analyzed the work of nine teams of three design thinkers who were part of a larger group that had been asked to complete four design tasks in virtual reality. The tasks varied in both virtual environment (garage or conference room) and task (concept generation and decision making.)

The content of this study is both virtual and physical. The apparatus consists of the physical space, the hardware interfacing with the physical world, the software running during the experiment, the virtual content displayed to the participants while in VR, and the design prompts used as study conditions.

The physical space for these experiments consisted of two adjacent rooms in a campus building associated with the design department. The larger room was 5 m by 3 m and the smaller was 4 m by 3 m. In order to have walking space (about 2 m by 3 m) for each of the three participants, the larger room was separated into two VR spaces by a blackout curtain. Within each participant's space, there was also a physical table and chair. Between sessions, participants sat at the table to fill out the study questionnaires.

The virtual reality headset in use was the HTC Vive headset and controllers. These headsets were connected to one VR-enabled laptop computer for each of the three participants. The software used for rendering and networking the multi-user

VR space was *High Fidelity*, a consumer-facing product that was available at the time of the experiment. The first-person VR audio and video were captured using OBS Studio (OBS Studio Contributors 2020), an open-source software for screen recording.

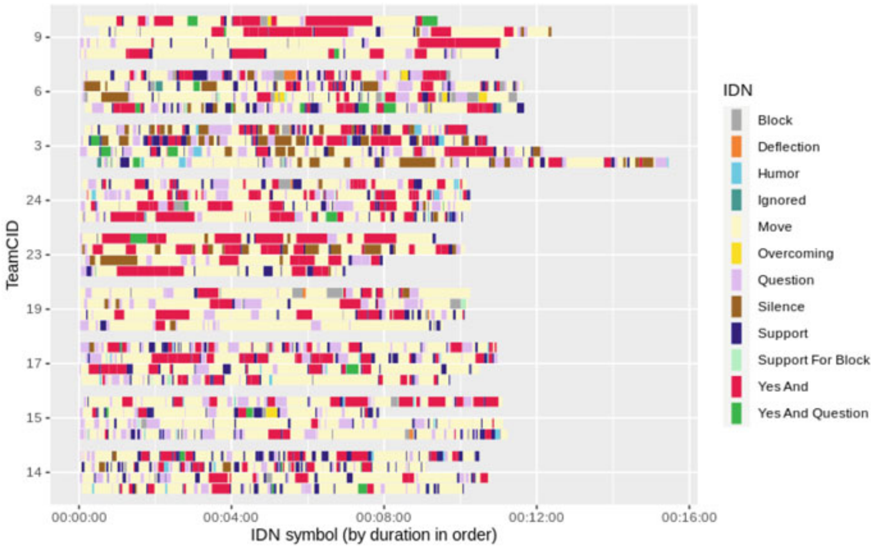
There were three virtual locations the participants experienced: a garage, a conference room, and a "practice room." The "practice room" was not designed to resemble a space encountered in the physical world but rather encourage familiarity with VR as a medium. Participants performed both concept generation tasks and decision-making tasks in the conference room and garage.

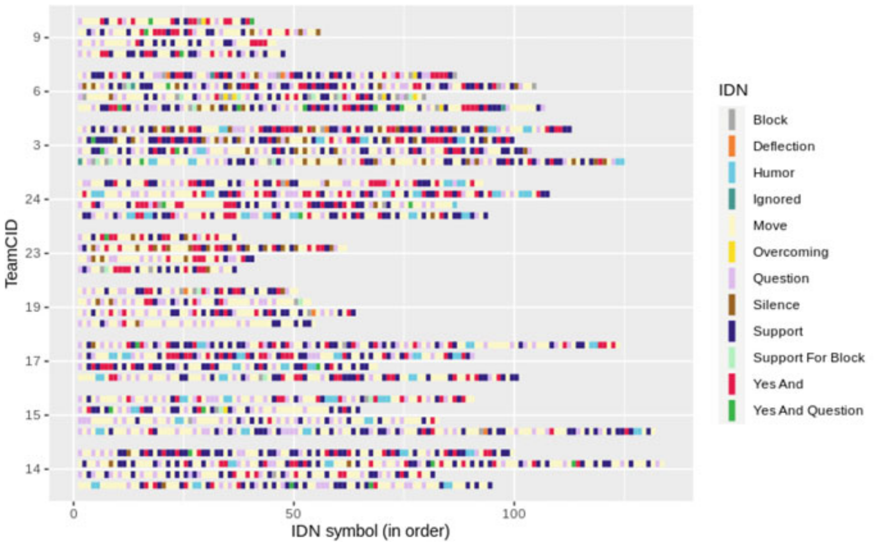
Verbal interaction information was recorded from first-person video and audio data. A trained coder annotated the combined recording using IDN notation. The output of this annotation process was a set of speaking turns (and silences) with the corresponding start time of the turn, the speaker of the turn, and the IDN symbol of the turn.

6 Descriptive Statistics

In order to continue diving deep into IDN, we show figures that have a lot of information but are not easily digestible. This is because we do not use the figures to support a single conclusion but wish to leave the opportunity to draw multiple conclusions based on the IDN categories. Some visual features are translated into statistical observations, but we keep these to a minimum.

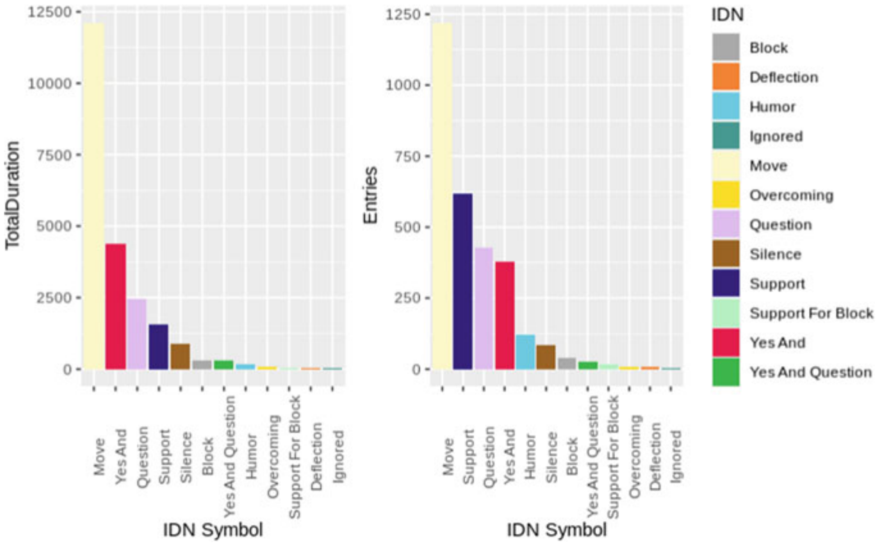
In these figures, IDN symbols are annotated by colors while the horizontal axis is time. Four of these bars, one representing each trial, are grouped vertically within teams. In the top figure, the width of a bar is proportional to the duration of the utterance, and in the bottom figure, the width of the bar is constant.





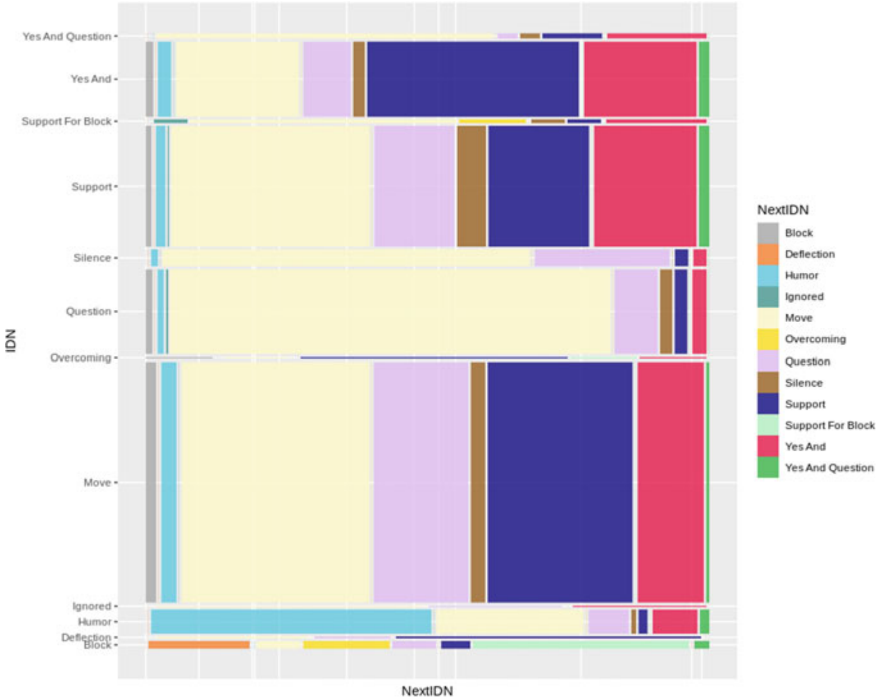
While total time is consistent due to experimental procedure, total symbols are not because utterances could vary in length.

The top graph, with width representing duration, is visually dominated by tan (Move) and red (Yes-And), but the bottom graph has noticeably more blue (Support). This is easily visible when looking at total duration and count by symbol.



Both distributions resemble a power law, either as a Yule-Simon or Zipf distribution, which often describes other sets of language or human classification data.

After learning the distribution of symbols, it is natural to consider the distribution of pairs of adjacent symbols, which indicates how symbols follow each other. These patterns are expressed especially well in a mosaic plot. In this mosaic plot, each row refers to a symbol, and the rectangles within each row refer to the distributions of each symbol that follows the row's assigned symbol. The height of each row is a prior probability, the width of each rectangle is the conditional probability, and the area, as the product of the two, is the joint probability.



In this figure, there are a handful of visual features that can be translated into statistical observations. The Question row in the middle of the box has a much wider light beige bar (representing Move) than other rows. From this, we can conclude that Moves follow Questions more often than expected. This observation is intuitive because it is likely for answers to questions to be coded as moves, as they provide information.

It is also visible that humor (near the bottom of the figure) has an exceptionally wide dark-blue rectangle (which is also Humor). This is because participants are likely to laugh at the same time, but one is still coded to follow the other. About half are the first pair, and about half are the second in the pair, followed by the next symbol.

7 Future Work

We take some space to detail a vision of two directions for the future. In the first, we envision the natural end of the design observatory, nicknamed the Ultimate Design Observatory. In the second, we focus on IDN specifically and the challenges that remain in automating the IDN coding process.

7.1 Ultimate Design Observatory

In his book, The Sciences of the Artificial (1969), Simon described design as follows: "Everyone designs who devises courses of action aimed at changing existing situations into preferred ones." This bias towards situated action and change in situations are additional primary foci that distinguish a design observatory from other observatories, whose primary foci has been observation and measurement. Thus bringing these foci on the external environment together along with new technologies in neuroscience such as the functional near-infrared spectroscopy (fNIRS) to observe the internal environment, the ultimate design observatory will enable us to observe and measure the behavior and action of designers and design teams within an environment, the process of change in the conditions of the external environment as well as the environment itself, the process of change in the internal environment, the brain and central nervous system, the responses of the designer, design team, and the environment to novel and unanticipated events, and the adaptations and non-adaptations of designers and teams to these novel conditions.

Coincident with our observations and measurement, the observatory, along with the scenarios enacted therein, also serves as a social teaching simulator to train designers and design teams in new and better ways of responding to novel situations. Given that the environment can be a virtual one, and some members of the team can be virtual avatars, it becomes clear that the number of combinations can be quite large. We see the ultimate design observatory as an ellipsoid with three foci. At the molecular level, the discovery of place cells and grid cells in the entorhinal cortex by May-Brit and Edvard Moser (Moser et al. 2015), the research by Tversky (2019), and earlier work we reported on the effect of location on synchrony between design team members (Mabogunje et al. 2020), shows a basic relationship between animals/humans and their physical environment. At the social level, the description of IDN described in this paper shows a variety of ways of interacting at the human–human level. The third level is a little harder to describe than the first two, nevertheless our work is beginning to give us some insight into what it might be. We do not know, whether to call it the evolutionary level, the temporal level, or the consciousness level, for it is a level that with virtual reality and historical evidence we can more easily and readily go back in time and space as a way to reimagine new ways of being.

The psychologist and neuro-anthropologist, Donald (2000) best describes what we can do in this ellipsoid of an observatory when he states:

> Conscious capacity is the key evolutionary feature of the human mind. It provides our connection with culture. At the same time, it is also the mediator for acquiring and assembling all our complex symbolic skills. But its clinching achievement is an ability to generate a virtual infinity of skills. One culture invents sailing, another throat singing. One culture insists on a capacious oral memory to remember entire pharmacopoeias, whereas another demands great dexterity with external symbols to manage an electronic universe. We are asked to handle now broadswords, now biplanes, now remote microsurgical hands. Here we are asked to manage the politics of a tribal village with a memory that reaches back 10 generations, and there we live in a global culture with an infinity of instant information but no

> collective memory. All these specialized skills take years to acquire, under close conscious supervision and with deliberate instruction. Pedagogy and interaction are both essential to the process, because these skill complexes are always the product of more than one conscious mind. Specialized skills result from deep cognitive interactions between brains and constitute the cognitive core of any given human culture.

Thus, the ultimate design observatory enables us to observe and design culture.

7.2 Automated IDN Coding

IDN at its core is a coding scheme that is applied to a conversation either in a verbal form as captured on video or in a written transcript in order to convert the conversation into a sequence of symbol categories. An attempt has been made to automate this assignment of symbol categories to a written transcript using a deep learning model (Koenig and Lim 2018). Thinking of an ideal automated system for converting a conversation into IDN-based insights, the following elements need to be automated.

1. Converting speech to text: A key challenge that remains for speech-to-text conversion for IDN is the unstructured nature of design meetings where language is less structured due to interruptions, elisions, and other anomalies. This challenge could be overcome by taking a high-precision speech-to-text system and personalizing it to the specific team that is being coded for IDN.
2. Identifying speakers in conversation: Speaker diarization systems are needed to identify which speaker said a specific utterance in a conversation. VR technology, video conferencing, and spatialized microphones make diarization much easier to deal with since each individual is speaking into a separate microphone.
3. Converting text to IDN categories: The deep learning model that was developed by Koenig and Lim (2018) is reasonably accurate in converting text to IDN categories. It needs further refinement to be ready for use in a practical setting.
4. Using vocal pitch and tone information to aid IDN categorization: A fair portion of the IDN signal is carried over voice. The modulation of voice when a question is asked, or a disapproval is expressed is an important input source for IDN categorization. Automated pitch and tonal information systems need to be developed that could take an individual's baseline voice information and recognize in a meeting specific modulations that could signal questions, sarcasm, frustration, disapproval, excitement, or humor.
5. Using visual signals such as gestures and facial expressions to aid IDN categorization: Computer vision algorithms could be used to detect facial expression signals and gestures that could be combined with voice information and text analysis to understand an utterance and assign an IDN categorization.
6. Detecting content information in transcript: IDN analysis in order to result in useful insights needs to be combined with an understanding of the content of the conversation. This could be the generation of product concepts, development of strategy, or making of decisions in the conversation being analyzed. Such

content information could be gleaned from a transcript with the aid of NLP algorithms that are aided by the context information about the key goals and key results of the meeting.

7. Visual dashboard of statistical analyses on IDN data: The IDN data needs to be analyzed with a number of statistical methods and visualized to show the relevant insights. The automated system needs to have a set of predetermined methods that can be applied to the IDN data of a meeting and a visual dashboard that shows the outcome of these methods.

8 Conclusion

IDN has continued to help to uncover patterns in design teamwork and human interaction. While future work remains for many kinds of questions, including questions of automation and a vision of its future use, there is also much work that has been done so far—linking both individual and groups of symbols with qualities like novelty, quantity, and effectiveness of design teams. IDN has also inspired other types of notations and has helped re-frame concept generation as a team activity. All together we continue to be hopeful in IDN's potential for the future.

References

Adly Taha, F. M., Adly Taha, R. M., West, K., Fazelpour, M., Herrmann, J. W., & Polvinale, M. A. (2019). An automated approach to recording and analyzing design activities using a graphical user interface. *Proceedings of the ASME Design Engineering Technical Conference, 2B–2019*, 1–11. https://doi.org/10.1115/DETC2019-98226

Alexander, E., Eppler, M. J., & Bresciani, S. (2019).Visual replay methodology: A mixed methods approach for group discussion analysis. *Journal of Mixed Methods Research, 13*(1), 33–51. https://doi.org/10.1177/1558689816664479

Arlitt, R. M., Nix, A. A., Sen, C., & Stone, R. B. (2016). Discovery of mental metadata used for analogy formation in function-based design. *Journal of Mechanical Design, Transactions of the ASME, 138*(10), 1–11. https://doi.org/10.1115/1.4034090

Björklund, T. A., Hannukainen, P., & Manninen, T. (2018). Measuring the impact of design, service design and design thinking in organizations on different maturity levels. *ServDes2018 - Service Design Proof of Concept*, (June 2018), 500–511. Retrieved from http://www.ep.liu.se/ecp/article.asp?issue=150&article=040&volume=

Bracken, J., Jablokow, K., Glavin, F. X., Henderson, D., & Erdman, A. M. (2019). *DETC2019-97704 a pilot study of student design teams.* 1–10.

Brandt, P. A. (2004). Spaces, domains, and meaning. *Essays in cognitive semiotics*, *4*.

Cagan, J., Dinar, M., Shah, J. J., Leifer, L., Linsey, J., Smith, S. M., & Vargas-Hernandez, N. (2013). Empirical studies of design thinking: Past, present, future. In *Proceedings of the ASME Design Engineering Technical Conference*, 5 https://doi.org/10.1115/DETC2013-13302

Camburn, B. A., Auernhammer, J. M., Sng, K. H. E., Mignone, P. J., Arlitt, R. M., Perez, K. B., … Wood, K. L. (2017). Design innovation: A study of integrated practice. *Proceedings of the ASME Design Engineering Technical Conference*, *7*, 1–10. https://doi.org/10.1115/DETC201768382.

Cash, P., & Štorga, M. (2015). Multifaceted assessment of ideation: Using networks to link ideation and design activity. *Journal of Engineering Design, 26*(10–12), 391–415. https://doi.org/10.1080/09544828.2015.1070813

Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J., Smith, S. M., & Hernandez, N. V. (2015). Empirical studies of designer thinking: Past, present, and future. *Journal of Mechanical Design, Transactions of the ASME, 137*(2), 1–13. https://doi.org/10.1115/1.4029025

Dolata, M., Uebernickel, F., & Schwabe, G. (2017). The power of words: Towards a methodology for progress monitoring in design thinking projects. *Proceedings Der 13. Internationale Tagung Wirtschaftsinformatik (WI)*, (February), 1156–1170. Retrieved from http://www.wi2017.ch/de/proceedings

Donald, M. (2000). The central role of culture in cognitive evolution: A reflection on the myth of the "isolated mind.". In L. Nucci, G. B. Saxe, & E. Turiel (Eds.), *Culture, thought, and development*. Psychology Press.

Endrejat, P., Meinecke, A., & Kauffeld, S. (2019). It all starts with a good idea: A new coding system for analyzing idea finding interactions (AIFI). *Proceedings of the 52nd Hawaii International Conference on System Sciences*, *6*, 305–314. https://doi.org/10.24251/hicss.2019.038

Feng, B. (2013). *Design of a therapeutic light device using Kano's model and a decision-making method based on HOQ.*

Georgiev, G., & Taura, T. (2014). Polysemy in design review conversations. *Design Thinking Research Symposium*, (2003), 1–19. Retrieved from http://docs.lib.purdue.edu/dtrs/2014/Identity/2

Goldschmidt, G. (2014). *Linkography: Unfolding the design process.*

Henderson, D., Booth, T., Jablokow, K., & Sonalkar, N. (2020). *Best fits and dark horses: Can design teams tell the difference?* 1–11. https://doi.org/10.1115/detc2020-22589

Horvat, N., Škec, S., Perišić, M. M., & Bojčetić, N. (2020). Relating problem-solving approach to turn-taking in communication of design teams. *Tehnicki Vjesnik*, *27*(3), 703–710. https://doi.org/10.17559/TV-20180911161642

Jablokow, K. W., Sonalkar, N., Avdeev, I., Thompson, B. D., Megahed, M. M., & Pachpute, P. S. (2018). Exploring the dynamic interactions and cognitive characteristics of NSF innovation corps (I-Corps) teams. *ASEE Annual Conference and Exposition, Conference Proceedings*, *2018-June*. https://doi.org/10.18260/1-2--30497

Jablokow, K. W., Sonalkar, N., Edelman, J., Mabogunje, A., & Leifer, L. (2019a). Investigating the influence of designers' cognitive characteristics and interaction behaviors in design concept generation. *Journal of Mechanical Design, Transactions of the ASME*, *141*(9). https://doi.org/10.1115/1.4043316

Jablokow, K. W., Vora, A., Henderson, D. A., Bracken, J., Sonalkar, N., & Harris, S. (2019b). Beyond Likert scales: Exploring designers' perceptions through visual reflection activities. *ASEE Annual Conference and Exposition, Conference Proceedings.* https://doi.org/10.18260/1-2--32150

Ju, W., Shluzas, L. A., & Leifer, L. (2016). *People with a paradigm: The center for design research's contributions to practice*. 209–222.

Koenig, C., & Lim, A. (2018). *IDN dialogue act classification with conditional random field and recurrent neural network*.

Luck, R. (2014). Seeing architecture in action: Designing, evoking, and depicting space and form in embodied interaction. *International Journal of Design Creativity and Innovation, 2*(3), 165–181. https://doi.org/10.1080/21650349.2013.875488

Lugnet, J., Ericson, Å., & Wenngren, J. (2020). Innovation supports for small-scale development in rural regions: A create, build, test and learn approach. *International Journal of Product Development, 24*(1), 30–42. https://doi.org/10.1504/IJPD.2020.106446

Mabogunje, A., Sonalkar, N., & Leifer, L. (2016). Design thinking: A new foundational science for engineering. *International Journal of Engineering Education, 32*(3), 1540–1556.

Mabogunje A., Sonalkar N., Miller, M. & Bailenson, J. (2021) Design team performance: Context, measurement, and the prospective impact of social virtual reality, to appear in Meinel C. & Leifer L. (Eds.), *Design thinking research.* Springer, Cham. (in press)

Martinec, T. (2019). A model of information processing and interactions in teams developing processing and interactions in teams developing technical systems. University of Zagreb.
Martinec, T., Horvat, N., Škec, S., & Štorga, M. (2018). Verbal engagement in teams solving a conceptual design task. *Proceedings of International Design Conference, DESIGN*, *5*, 2075–2086. https://doi.org/10.21278/idc.2018.0540
Martinec, T., Škec, S., Horvat, N., & Štorga, M. (2019a). A state-transition model of team conceptual design activity. *Research in Engineering Design, 30*(1), 103–132. https://doi.org/10.1007/s00163-018-00305-1
Martinec, T., Škec, S., Šklebar, J., & Štorga, M. (2019b). Applying engineering design ontology for content analysis of team conceptual design activity. *Proceedings of the International Conference on Engineering Design, ICED*, *2019-Augus*(AUGUST), 2467–2476. https://doi.org/10.1017/dsi.2019.253
Martinec, T., Škec, S., Perišíc, M. M., & Štorga, M. (2020). Revisiting problem-solution co-evolution in the context of team conceptual design activity. *Applied Sciences (Switzerland)*, *10*(18). https://doi.org/10.3390/APP10186303
McInnis, B., Xu, X., & Dow, S. P. (2018). How features of a civic design competition influence the collective understanding of a problem. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW). https://doi.org/10.1145/3274389
Menold, J., & Jablokow, K. (2019). Exploring the effects of cognitive style diversity and self-efficacy beliefs on final design attributes in student design teams. *Design Studies, 60*, 71–102. https://doi.org/10.1016/j.destud.2018.08.001
Menold, J., Starkey, E., & Mccomb, C. (2020). *Analyzing the characteristics of cognitive-assistant-facilitated ideation groups*. 1–11.
Moser, M., Rowland, D. C., & Moser, E. I. (2015). *Place cells, grid cells, and memory*. 1–15.
OBS Studio Contributors. (2020). Open Broadcaster Software. Retrieved January 28, 2020, from. https://obsproject.com/
Paulsen, H., Klonek, F., Meinnecke, A., Schneider, K., Liskin, O., & Kauffeld, S. (n.d.). Driving and hindering forces in group discussions: Analyzing change and sustain talk in a software engineering project.
Schinko-Fischli, S. (2018). *Applied improvisation for coaches and leaders: A practical guide for creative collaboration*. Routledge.
Simon, H. A. (1969). *The sciences of the artificial*. MIT Press.
Sonalkar, N., Mabogunje, A., Miller, M., Bailenson, J., & Leifer, L. (n.d.). *Augmenting learning of design teamork using immersive virtual reality*. 1–13.
Sonalkar, N., Mabogunje, A., & Leifer, L. (2013). Developing a visual representation to characterize moment-to-moment concept generation in design teams. *International Journal of Design Creativity and Innovation, 1*(2), 93–108.
Sonalkar, N., Jung, M., Mabogunje, A., & Leifer, L. (2014a). A structure for design theory. *An Anthology of Theories and Models of Design*, 67–81. https://doi.org/10.1007/978-1-4471-6338-1_3.
Sonalkar, N., Mabogunje, A., & Leifer, L. (2014b). Analyzing the display of professional knowledge through interpersonal interactions in design reviews. *DTRS 10: Design Thinking Research Symposium 2014*, (1992), 1–18.
Sonalkar, N., Mabogunje, A., Pai, G., Krishnan, A., & Roth, B. (2015). Diagnostics for design thinking teams. *Design Thinking Research: Making Design Thinking Foundational*, 1–290. https://doi.org/10.1007/978-3-319-19641-1.
Sonalkar, N., Mabogunje, A., Hoster, H., & Roth, B. (2016a). Developing instrumentation for design thinking team performance. *Design Thinking Research*. https://doi.org/10.1007/978-3-319-40382-3_15
Sonalkar, N., Mabogunje, A., Leifer, L., & Roth, B. (2016b). Visualising professional vision interactions in design reviews. *CoDesign, 12*(1–2), 73–92. https://doi.org/10.1080/15710882.2015.1135245

Sonalkar, N., Jablokow, K., Edelman, J., Mabogunje, A., & Leifer, L. (2017). Detc2017–68239 design whodunit: The relationship between individual characteristics. 1–10.

Sonalkar, N., Jahanikia, S., Xie, H., Geniesse, C., Ayub, R., Beaty, R., & Saggar, M. (2020). Mining the role of design reflection and associated brain dynamics in creativity. https://doi.org/10.1007/978-3-030-28960-7_10

Sosa, R. (2019). Accretion theory of ideation: Evaluation regimes for ideation stages. *Design Science*, *5*(Yang 2009), 1–33. https://doi.org/10.1017/dsj.2019.22

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science, 12*(1), 49–100.

Tversky, B. (2019). *Mind in motion: how action shape thought*. Basic Books.

Wulvik, A., Jensen, M. B., & Steinert, M. (2017). Temporal static visualisation of transcripts for pre-analysis of video material: Identifying modes of information sharing. *Analysing Design Thinking: Studies of Cross-Cultural Co-Creation*, pp. 271–287. https://doi.org/10.1201/9781315208169.

PretoVids: A New Approach to Digital Prototyping

Jonathan Antonio Edelman, Joaquin Santuber, and Babajide Owoyele

1 Introduction—What This Is About

This chapter introduces PretoVids, a digital and practical, research-based method for structured, evidence-based prototyping. This new prototyping technique is genuinely relevant for showcasing digital media's role in design-as-performance in concept development (Edelman et al. 2020). The PretoVids Method results from experimenting on digital Design during Covid-19 and further observing the effect of the media and artifacts in high-performance design teams and object mediated interaction patterns at work. Preliminary research has indicated that traditional prototyping methods do not provide an agile heuristic for digital product development teams. Furthermore, a critical examination of the prototyping approaches reveals uncharted territory regarding prototyping as an agile product development methodology. Dedicated to translational research and development funded by the Hasso Plattner Design Thinking Research Program (HPDTRP), the Research to Impact group has conducted action research with design thinking coaches and industry experts to develop the PretoVids method. In the context of new product development, PretoVids reduces many risks of building full digital projects, saves time, and affords easier value proposition communication and sharing and learning. In 5–7 steps, design teams can perform an iterative process to develop new product concepts without spending a penny on software development or writing a single code line. One advantage of the PretoVids approach over traditional prototyping is that team members explicitly

J. A. Edelman · J. Santuber · B. Owoyele (✉)
Hasso Plattner Institute for Digital Engineering, Campus Griebnitzsee, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany
e-mail: Babajide.Owoyele@hpi.de

J. A. Edelman
e-mail: jonathan.edelman@hpi.de

J. Santuber
e-mail: Joaquin.Santuber@hpi.de

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_11

communicate product ideas in a constrained environment, which aids in sensemaking and a crisp understanding of the touchpoints, usability, scenarios, form, and features, well as the core functionality of the product/service or system. Building on previous work on performative patterns (Edelman et al. 2020), the steps of PretoVids can be instantiated as roles or parts that the designer or design team members can perform. Our research question is, "***How can new digital product development teams prototype new digital products without writing a line of code or spending a dime on software development to learn what customers love?***"

2 BIG IDEA: PretoVids Developed for Low Cost, Agile New Product Development Without Engineering and Writing Code

PretoVids are short animated videos that enable early-stage product-service-system development teams to concentrate on user experience and value creation before writing a code line. PretoVids are used in a structured sequence comprising basic digital interface visualization (sketching), voice-over narration, and—most importantly—video editing. For the past five years, we have developed them in academic courses and professional workshops, and in consulting with over 24 companies, PretoVids are used to explore interactions amongst networks of users in the context of new digital products.

In today's world, businesses are continually looking for new opportunities to create new value by generating more profits, reducing risks, lowering costs, and increasing social impact (Carayannis et al. 2015; Dunne 2018; Yan 2018). Firms adopt many innovative methods to develop new products and services and help open up how companies achieve their primary goal: value creation (Breuer and Lüdeke-Freund 2017; Brown and Katz 2009; Dahlander and Gann 2010). Design-led innovation, Design-driven innovation, Human-centered design are all labels for methods companies approach value creation for end-users (Dell'Era et al. 2018; Fleischmann 2019; Goey et al. 2017; Leifer et al. 2014).

3 Problem/Opportunity: Customer Engagement and Feedback When You Cannot Meet in Person

We began using PretoVids as tools for agile and flexible early-stage concept development. However, the current situation with COVID19 required our team to consider using PretoVids as a medium for communication and as a way to promote customer engagement and feedback when face-to-face interactions with customers and other stakeholders are not possible.

4 Impact: Customer Engagement at a Distance; Insights

1. PretoVids helps *solicit* customer input about scenarios, functions, and features that they have not considered.
2. PretoVids helps companies incrementally transform their ideas into a product that customers will want and will buy, leading to more profit and company value.
3. Building PretoVids leads to having a tangible blueprint for product development; it also increases the company product portfolio and can be crucial to patentable assets.

5 Related Work and Theoretical Background

Prototyping and the boundary objects resulting from the method is at the very heart of what makes design a catalyst for innovation, transformation, and change (Brown, 2008; Buchanan 2019; Curedale 2019; Brown and Wyatt 2010; Buchanan et al. 2001; Liedtka 2018, 2020). Beyond aesthetics and form, prototyping in design helps *communication* (Andriole 1994; Capata and Tatti 2020; Gerber 2010), as, in recent times, increasing attention is given to design prototypes by the industry and business organizations because of their generative role in developing products (Durao et al. 2018; Baldassarre et al. 2020). Prototyping makes design a tangible and hands-on approach to practicing innovation management, a crucial element of the creative mindset, expanding teams' dynamic capability, and is relevant for inter-organizational strategic development (Marks and Chase 2019; Newman et al. 2015; Shalpegin et al. 2018). Some have suggested that prototyping is the most crucial part of designing because of its impact on *team creativity* (Dow et al. 2011; Kim and Hinds 2016). Based on a bibliometric analysis of the web of science articles on design thinking alone, prototyping is the subject of 18 articles (3%) of the literature on design. See Fig. 1. Figure 2 is a word cloud of author keywords showing how critical prototyping is to design thinking practice and research.

Fig. 1 **a** Key Themes in design thinking,18 articles on prototyping, **b** word cloud (Web of Science 2020) (https://www.webofscience.com/wos/woscc/basic-search)

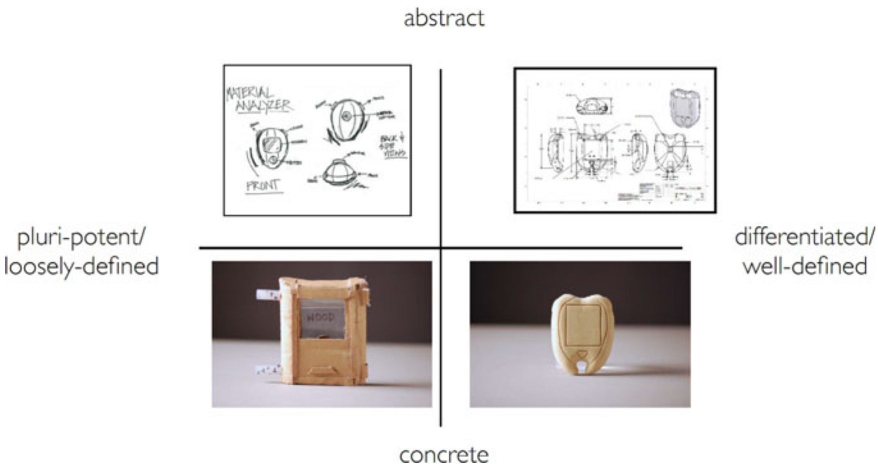


Fig. 2 Media models framework (Edelman 2011)

5.1 *Gaps/Opportunity Identified in Design Thinking Prototyping Literature*

The gaps we identified from this existing literature are related to *how* to perform prototyping in the digital transformation paradigm and in light of the effects of the COVID-19 pandemic during the year 2020.

The literature on prototyping highlights its value as a phase in the design process and its benefits in learning and fostering tangible experiences in design teams (Dow et al. 2009; Efeoglu et al. 2014). Firms look to leverage human-centered design for integrating diverse individual perspectives into a shared understanding and boundary concepts (Dong et al. 2013). Such shared knowledge later informs iteration of product features and service offerings (Caprari et al. 2018; Oehlberg et al. 2012).

Prototypes are seen as innovation assets that carry the knowledge of user preferences (Cockton et al. 2016; Haldane et al. 2019). Managing requirements modeling and prototyping is risky, and complications can lead to unmanageable projects (Gomez 2015; McShane 2018). Requirements modeling and prototyping processes must be fast, powerful, cost-effective, sane, and objective, and prototypes should always be cost-effective and always improve specifications (Andriole 1994; Canedo and Da Costa 2018; Carroll and Richardson 2016; Levy and Hadar 2018). Some work exists (Vetterli et al. 2013; O'Driscoll 2016; Martins et al. 2019) suggesting generic steps as a process for prototyping, e.g., obtain initial requirements; model requirements; constraints identification; first according to importance; Design; design evaluation; specification; interaction modeling; and validating requirements (Andriole 1994).

In recent work, the characteristics of prototyping and its dynamics were investigated by Dow and colleagues. It was concluded that "prototyping leads to better design results, more divergence, and increased self-efficacy." (Dow et al. 2011) Here the authors take an experimental yet scientific approach to understand and untangle how prototyping practices affect learning, motivation, communication, and design outcomes. Iterating and idea refining are advantages of prototyping, which helps designers learn (Beckman and Barry 2007; Bharathi and Pande 2019; Borge et al. 2020). Dow and co-authors further suggest that "groups who produce and share multiple prototypes report a significantly higher increase in rapport, and also exchange more verbal information, share more features, and generally reach a better consensus" (Dow et al. 2011).

Today, one relevant question with the COVID-19 situation is how teams can effectively perform iterative prototyping for learning when dispersed. To build on Dow's work on team Prototyping dynamics in the digital era can thus improve our understanding of how sharing multiple designs improves concept exploration, group rapport, and value creation.

Further empirical research in the context of *Product management* prototypes (Beltagui 2018; Laptev and Shaytan 2019; Seidel 2007) revealed the role of prototyping as part of "*a repertoire of vital universal competencies of product managers* involved in innovative product development and working in *a digital and collaborative environment*". (Laptev and Shaytan 2019). For the authors, design thinking project-based learning combined with a digital rapid prototyping technique has proven to be a practical tool for developing these competencies (Steinbeck 2011; Voinea 2019). They have also shown themselves to be new but better products over those created using traditional product development methods (Meinel et al. 2020). These studies further show that prototyping is a crucial "universal competency" through which firms can leverage agile approaches for innovative product development based on digital rapid team-based techniques (Pereira and Russo 2018).

In the specific context of *Software Engineering* and agile approaches (Gurusamy et al. 2016; Mesquida et al. 2017; O'Driscoll 2016), we find PretoVids can augment work on prototyping software (Eickhoff et al. 2018; Corral and Fronza 2018). Software and digital artifacts are also tools for organizational and social change, rather than just end products or solutions (Penzenstadler 2020). Furthermore, in product development contexts, buy-in from diverse actors is crucial to company growth and success. (Newman et al. 2015) Software engineering teams nowadays need an *agile approach for handling uncertainties in the software development process.* Prototyping is crucial in software development, and so addressing the spatial challenges teams face when engaging in developing software products is valuable (Bas and Guillo 2015; Brown and Wyatt 2010).

This paper further offers the opportunity for a *product development heuristic* for digitally operationalizing such an agile process in the context of Covid-19, where local and global teams are dispersed and mostly engaged in remote work (Wenzel et al. 2016).

In H*ealthcare design and innovation*, we see gaps that can be filled through prototyping approaches such as PretoVids, following work by Hopkins and colleagues, as well as evidence in the research by (Aakhus and Harrison, 2016; Aaronson et al. 2020; Abookire et al. 2020; Altman et al. 2018). Their case studies provide a framework using a multidisciplinary approach to digital health design, using "*rapid* digital prototyping" to build products, services, and systems. Informed by the health-centered design process, our idea of PretoVids' relevance to the healthcare domain draws on our teaching practice in the Digital Health Design Lab (DHDL) at the Digital Health center in the Hasso Plattner Institute. We are now using PretoVids as an agile exploratory approach to dive into patient experience improvement in themes such as Fetal Alcohol Syndrome (Mukherjee et al. 2006), N-of-1 Trials (Lillie et al. 2011), hypertension (Morassi Sasso et al. 2020) and the broader emerging digital health field. A review of the literature mentioned above, combined with our action research (practice), informs the opportunity to augment current digital health design approaches with prototyping heuristics and methods such as PretoVids. Such prototyping methods could better contribute to a systematic approach to *digital product development* in healthcare following work by (Hopkins et al. 2016; Succi et al. 2018; Vechakul et al. 2015; Woods et al. 2017), as well as engaging teams for exploring customer/patient engagement in healthcare (Woods et al. 2018; van de Grift and Kroeze 2016).

At the Hasso Plattner Institute, the intention behind our research in previous years was to create a rigorous set of materials, tools, and methods for training practitioners for increased satisfaction and performance (Edelman et al. 2020). Since the original research and creation of material, tools, and methods appeared, we have come out with new work showcasing the translated research insights and new Design Thinking knowledge into actionable frameworks and materials for teaching and learning Design Thinking.

After subsequently running workshops and training sessions[1] to introduce and refine our teaching materials, our emphasis continues to be on the performative act of designing or design-as performance. A broad and indexed digest of research that will serve as material for creating training is forthcoming. Through engaging researchers and coaches at the HPI, Stanford, and the Digital Health Center, we ran workshops with research materials, thereby transforming them into preliminary frameworks, tools, and methods to emphasize Digital Engineering and Digital Transformation training materials. We developed nine refined, research-based educational packages. The set effectively communicates relevant theoretical insights into actionable frameworks, focused skills, team drills, and relevant scenarios for free play that support improved and measurable team performance. They are:

- MEDGI (Edelman et al. 2020)
- Generative Design Questions and Deep Reasoning Questions: two kinds of questions that designers use to frame effective inquiry (Eris, 2003)
- Solicitations: designers create models that solicit appropriate phase actions (Rietveld et al. 2018)
- Noun Phrases: designers create new language (Mabogunje 1997)
- Dimensions of Engagement: a system approach to generative product service architecture (Edelman et al. 2012)

[1] (Stanford University, Politecnico di Milano, Hasso Plattner Institute, Royal College of Art, and S.R.H. Hochschule).

- Disruption-Integration: the primary algorithm (Edelman et al. 2012; Menning et al. 2017)
- Enactment: acting out in semi-imaginary worlds (Carlgren et al. 2016; Dong et al. 2013; Edelman 2011)
- Marking: designers enlist a shorthand sketch for enacting interaction (Kirsh 2011; Muntanyola-Saura and Kirsh 2010)
- Remapping: transposing touchpoints onto different form factors for new usability and use-cases (Edelman et al. 2020)
- Four Forces of Change: Aristotle's Four Causes in the service of Design (Edelman et al. 2020)
- Metaphor: using metaphor to leverage high impact opportunities (Casakin 2006; Choi and Kim 2017; Cienki and Müller 2012; Edelman 2011; Hey et al. 2008)
- Media Models: the media that designers enlist have cognitive affordances (Edelman 2011).

5.2 Media Models and Systems, Why not Wireframes, Use-Case Scenarios

People engaged in innovation and designing are situated in a material world—with tools, surfaces/ paper, pen/or screens, mice, and keyboards. These material objects scaffold specific behavior at various times and spaces. We refer to these material objects as media. Media carries and—most importantly—display information in diverse ways, allowing different possibilities for actions. The Media Models Framework (Fig. 2) is fundamental in accounting for media characteristics that afford diverse kinds of questions, behaviors, and solicitations (Edelman et al., 2020). The media model framework is grounded in the paradigm of extended and distributed cognition (Tversky 2003, 2019). This perspective posits that design teams as a unit are made up of individuals performing "thinking" as a loop that engages the brain, the body, and the media/tools they work. Some media is useful for optimizing and manufacturing, while others are effectively refining physical details or thinking about concepts. Media may also be useful for acting out experiences (Kirsh 2010; Tversky 2019) beyond sketching (Buxton 2007).

According to their refinement and abstraction level, the media model's framework was developed in design thinking research and sorts prototypes relatively well. Designers often work with rough prototypes (left side of the matrix)—this helps to communicate and consider core design ideas. By contrast, refined prototypes (right side of the matrix) suggest almost finalized solutions; instead, they stimulate conversations about and consideration fine-grained design details.

Our research group has been exploring an important question: "***Is it possible to use this analytic framework to translate models to afford generative team behaviors?***" Understanding how the media used influences our work as designers and affects teams' innovation outcomes is central to our research agenda. Numerous cases in academia (our research and taught courses) and our practice-oriented engagement with industry as well as with the HPI Academy has shown this not only to be possible but robustly practical. This approach is essential to understanding internal team dynamics and the testers, clients, users, and eventually managing and effectively communicating the work.

In making PretoVids, teams can now operationalize the Media model framework systematically and iteratively. PretoVids provide a rapid and agile approach to support digital product development teams in their missions to embrace the emergent ambiguity and uncertainty of complex value propositions and systems-level concepts while communicating the solutions' core offerings to diverse stakeholders. The advantage of framing the PretoVids as a crucial instantiation of the Media Model Framework is to account for how this kind of digital-friendly prototyping reduces cognitive loading (Sepp et al. 2019) through externalizing memory and processing in phase-appropriate media (Jablokow et al. 2019; Edelman et al. 2020; Kirsh 2000).

6 The Pretovids

The PretoVids method is the outcome of several experiences in different contexts—education and industry—and across different countries through years of research and practice.

7 The Journey to PretoVids: Stanford ME310, R.C.A., POLIMI., D.H.D.L., R2I, and SAP

The development of PretoVids occurred over the span of ten years, on two continents and more than four countries. Our first interaction with companies wanting videos as a template for new product development happened in the context of Stanford Engineering's capstone course ME310, run by Professor Larry Leifer. In ME310, sponsoring companies present a problem or opportunity to student teams, who then spend the academic year creating a working mechatronic prototype to hand off to the company. After the hand-off, the company will often move the prototype into a development cycle. In 2010, one sponsoring company proposed that instead of a working prototype, they would prefer to receive several videos with a dummy object being used in different use-case scenarios. The company said that they already had engineers. They already had industrial designers; what they needed was new, fleshed-out concepts to bring to their development teams—ideas were not enough. The ME310 teaching team rejected the request, as one of the course's primary purposes was to teach master's students how to do a full cycle of product engineering and development.

Nonetheless, the sponsored team did produce several short videos as they felt it would be a valuable exercise in exploring their new concepts for “object-interactions.”. What is notable about the project sponsors’ request and the team videos is that they favor depicting the context in which the new product would exist rather than solely focusing on the idea or the object. We call this integration of object and experience the ‘object-interactions.’

The next contact we had with the enlistment of this kind of video in the development process was at the Royal College of Art in London. In his course, Object Mediated Interactions, master designer and teacher Durrell Bishop would instruct his students to make short videos that explored and depicted concepts for his design students’ potential projects. While final projects would involve hardware and code, the videos were made in advance of engineering to focus on what the object was, how it worked, and just as importantly, what constituted user interactions with the object. Perhaps the name of the course says it all: Object Mediated Interactions. By putting the ‘interactions’ on the same level as the ‘object,’ Bishop signaled to his students that they were creating more than hardware, and one of the simplest ways of doing this was by creating small videos. In his professional practice, Bishop had pioneered this technique over twenty years before as agile concept development and communication tool, with his animation of the ‘Marble Answering Machine’ (Dragicevic and Jansen, 2012; Bishop, 1992), which illustrated different scenarios for interactions with a novel concept for an answering machine.

We started using the video prototypes in the Global Innovation Design program in 2016 at the Royal College of Art and Imperial College, London, for several reasons. One reason was following Bishop’s lead in allowing fast iteration, communication, and feedback on new interactions. Before the enlistment of PretoVids, we found that students were struggling with developing their "ideas” in their heads rather than exploring new object-interactions. The teaching team found it challenging to visualize and critique student work when presented as ideas in words. The PretoVids allowed the students’ peers and the teaching team to see the object-interactions being brought to light. Other reasons for enlisting PretoVids included language barriers that faced the community in the context of a radically multicultural student body and because students would spend several semesters abroad and were often unable to present their work in person.

At Politecnico di Milano in the spring of 2020, an international teaching team of an introductory masters’ level design course found themselves in a difficult situation. One foundation of the course is user-testing, gathering feedback from users, and making changes based on that feedback. Unfortunately, the region was under lockdown because of the COVID19 situation. Face-to-face user research and testing was impossible. We considered different strategies and landed on having the students create a series of videos that would give the user the experience of their interventions. Initially, we had considered beginning with a sketch video, similar to an animatic used in the film industry to animate storyboards, and ending with a more polished video using well-rendered wireframes. However, some companies pushed back on the final deliverable, as they found the animatic style a better fit for their organizational culture. In some cases, this meant that the student teams who had already completed the wireframe version had to go back and create a new animatic version for their final deliverable.

During an extensive new product development program with SAP and their partners, the HPI Academy realized the most sophisticated iteration of the PretoVids method to date. Ten professional design coaches led ten industry partners to develop new products. What emerged through these iterations was a more rigorous structure to the method, both in terms of a professional development schedule with milestones and deliverables. In terms of a tighter structure for the PretoVids itself, we used a theatrical three-act structure for the central part of the PretoVids, with an explanatory prologue and epilogue as bookends to the three acts. With the SAP project, we could see how partners interacted with customers in using the PretoVids to get new insights from customers about what kind of product they would want and want to buy.

Finally, we introduced the PretoVids method in its mature form to our Digital Health Design Lab classes in digital health design at the Hasso Plattner Institute's Digital Health Center. In this context, master's students work individually and on teams, using PretoVids to explore and get feedback about digital interventions in the health field. PretoVids serve as final deliverables, which provide a lasting communication piece that goes far beyond having many ideas written on sticky notes. They concentrate on far more than a brilliant idea or object: PretoVids give the viewer experience of the object-interaction. They can see themselves using and benefiting from the object-interaction and can imagine new use-case scenarios and benefits that the digital health design team may not have considered.

8 The PretoVids in Action

Developing new products is risky. The optimal choice to start developing new products would be to use the least time as possible. In turn, companies save money, people's effort, and other resources. A risk management perspective informs using PretoVids in building products efficiently as this affords learning through prototyping at almost no cost. However, without the PretoVids approach, high cost is invested in developing a high-end product that customers may not buy. This business decision would represent a significant loss that could have been avoided (Savoia 2019). This background informs why we coined the term PretoVids.[2] Pretotyping entails performing an Agile development approach to get evidence before making a significant investment of time, money other resources. A similar approach is evident in the entrepreneur-oriented book "The Right It" (Savoia 2019). One of the main objectives of the pretotypes is to generate data with a low financial commitment. Thus, entrepreneurs can make evidence-based decisions regarding the development of a new business opportunity. In designing and new product development, the same is achieved with PretoVids: low-cost digital media that is "open" and can generate

[2] "Pre" meaning before.

first-hand data to make informed decisions regarding the design process. (Drechsler and Natter 2012).

- Informed by the principles shown in Fig. 3, PretoVids are grounded in the idea of leveraging systems and digital approaches for generating evidence in agile development contexts. By systems approaches, we draw on the literature level for effectively intervening in a system, e.g., the *Leverage Points* model by Donella Meadows (Meadows 2008) and the *Dimensions of Engagement* (Edelman 2011; Meadows and Wright 2008). We also draw on the literature on the challenges and opportunities of the digital transformation paradigm: openness, affordance, and generativity of digital technology (Autio et al. 2018; Nambisan et al. 2017). From these explicitly digital transformation lenses (Nambisan et al. 2019), we ground the PretoVids and Pretotypes approach in designing apps, platforms, or infrastructure that people (diverse users in the network) will find meaningful (Verganti 2008; Verganti et al. 2020). PretoVids helps to get evidence before innovation stakeholders, who make a significant investment of time, money, or people into the concept or product. The evidence has three different focuses: First, the PretoVids is a *Pretotype in video form* and helps to solicit input from customers about scenarios, functions, and features that actors in the innovation team may not have considered
- Second, the PretoVids functions as *an iterative external instantiation* that allows design and innovation teams to incrementally transform their concept into a product that customers will want and will want to buy
- Third, the PretoVids serves as *a tangible, shareable asset for iterative product development* and the marketing of concepts generated during the new product development session. The PretoVids can be *easily handed over* to either U.I. or UX designers or professional video specialist to make it higher fidelity as required.

9 Component of the PretoVids

9.1 Principles

Human-Centered Design Principles for Iteratively Building PretoVids

Human-centered

a. Making PretoVids is grounded in the fundamental notion that "**people and their situation are primary drivers of change**."
b. Using PretoVids showcases how-to "**create *experiences*, not *ideas*.**"
c. The team or individual creating PretoVids should "**put themselves and their organizations in the driver's seat**."
d. The PretoVids designer or team should "**see and experience from the user's point of view**".

Do-to-know

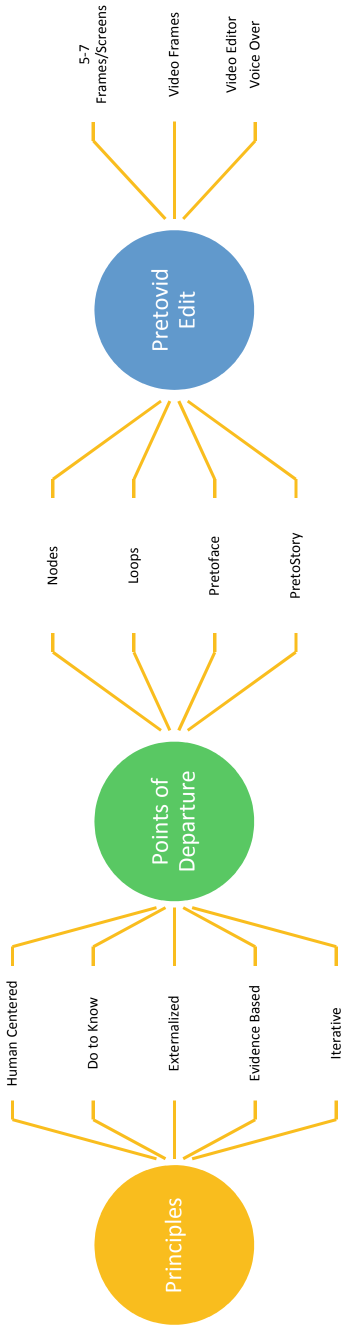


Fig. 3 PretoVids components: principles, points of departure, PretoVids edits

Human systems are extraordinarily complex

a. difficult to predict
b. made even more difficult to predict due to recent technology.

Evidence-based

a. based on observations
b. tested in the world
c. subject to reduced risks in new product development and introduction
d. tested with customers and partners.

Iterative

a. proposal
b. intervention
c. assessment
d. repeat.

Externalized

a. shared & "open" instantiations.

10 Building the PretoVids Step-By-Step

In this section, we introduce the method. For teaching/training purposes, we start with a straightforward concept to try out the first steps to create the PretoVids. To illustrate, we use the "Whole lot of Love" concept to get participants familiar with the technique we are using (Fig. 4).

Nevertheless, digital product development can be more complicated than a "Whole lot of Love," thus, we provide the designers with a very explicit guideline on starting with the PretoVids. In the following paragraphs, we present the reader with the step-by-step on creating your first PretoVids.

Fig. 4 A visual example of the PretoVids process, step 1, creating a PretoStory (courtesy Anne Victoria Talbot)

11 Point of Departure

Every PretoVids starts with a user. The point of departure for digital product development is a user-input–output system (see Fig. X). A user has a problem and needs to achieve something (output). The opportunity is the input data that the user can provide—there is always some sort of data to trigger the process (Fig. 5).

11.1 PretoStory

Once the problem and the user-input–output system are identified, we need to start working on a very simple storyline. In 7 frames, we present the new product concept by showing the value proposition and explaining how it works in three steps. By explaining "how it works," we add digital capabilities to the user-input–output system with a processing element. The processing power could become a machine learning model or a data analysis module, among other possibilities. Test your story in different ways before moving on to the next step (Fig. 6; Table 1).

PretoFace V0

Now that we have our system running—*user-input-processing output system* (UIPO System onwards)—we can envision specific use case scenarios that can benefit from the original individual user's problem. Put in simple words, with this system in place

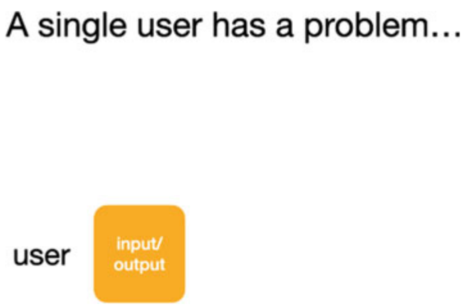


Fig. 5 The user-input–output system diagram to understand problems as the starting point

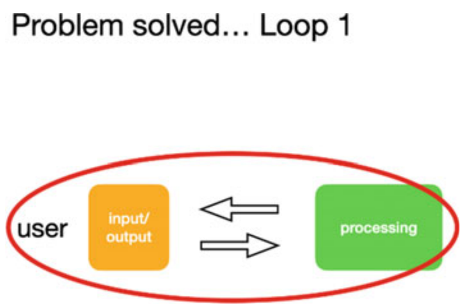


Fig. 6 The user-input–output point of departure system is extended by a digitally-enabled processing power—a machine learning model, a data visualization tool, or other

Table 1 The PretoStory structure will help you get the story right in a structured and straightforward way

Frames	Description	Example
Name of the product	Give it a name and a simple logo (sketch something)	"A Whole lot of Love."
Value proposition	Big idea/problem opportunity/impact	"Say I love you 1000 times a day!"
How it works 1 (input)	Basic function of the platform	Get 1000 people
How it works 2 (processing)	How is your digitally enabled solution helping the user to achieve their goal?	Push button to send Valentines instantly
How it works 3 (output)	What is the user getting back?	Make 1000 people happy
Value proposition	Big idea/problem opportunity/impact	"Say I love you 1000 times a day!"
Name of the product	Close the PretoStory with the same image you started with	"A Whole lot of Love."

within an organization, who else and how will they benefit from it? (See Fig. X). How can a user in "Management" also benefit from the "Whole lot of Love" UIPO System? For example, someone (/an employee) in the HR department.

Following the same structure, elaborate in up to 7 seven sketches, one use-case scenario from a new user's perspective. Think of screenshots of your new digital product. What will the user see in every screenshot? (Fig. 7; Table 2).

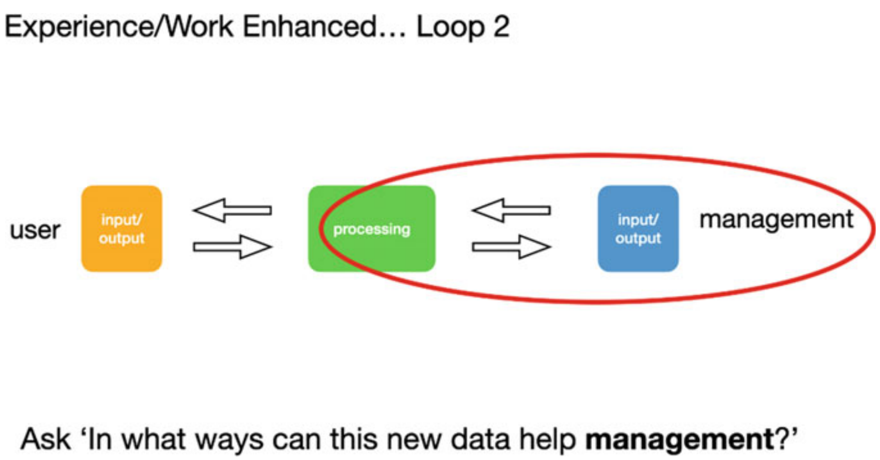


Fig. 7 Envisioning new use-case scenarios within an organization that could benefit from the original UIPO system

Table 2 The PretoFace V0 structure and first use-case scenario try

Frames	Description	Example
Name of the product	Give it a name and a simple logo (sketch something)	"Whole lot of Love."
Use case scenario intro	What can the user do?	HR: Track employee connections across departments to flourish collaboration
Interface 1 (input)	What does the user see on the screen? Sketch buttons, input text boxes, or other actionable elements for the user	Generate graph network of sent "Whole lot of Love." Button: "View connections."
Interface 2 (processing)	What is happening in the backend?	Generating "Whole lot of Love" Social Network across Accounting and Legal Department (48% out 100%)
Interface 3 (output)	What does the user see as an output to this specific use-case?	Graph with all "Whole lot of Love" sent across Acc and Legal
Use case scenario outro	Reinforce your message on what the user can do	Understand where collaboration is more likely to flourish in your company with "Whole lot of Love."
Name of the product	Close the PretoFace with the same image you started with	"Whole lot of Love."

PretoFace V1

We explored in the previous step, envisioning use-case scenarios within the organization that would benefit from the original user-input-processing-output system (UIPO System). Using the same logic, we can now tap into the same UIPO and envision a second and third use-case scenario. How can this UIPO System create value for a user in the C-Suite?

Let us move one step further, and in 24 frames, dig deeper into the three use case scenarios you envision for your new digital product. Again, sketch them as a sequence of screenshots of your future new digital product (Fig. 8 and Fig. 9; Table 3).

12 PretoVids Edit

With all the building blocks in place, bring all the frames together in a basic video-editor, record a voice-over narration, and export it as a video. To enhance the video, you can add appropriate background music (Fig. 10).

Apart from the structured guidelines on creating a PretoVids, there are concepts from systems design that are especially relevant when designing digital products. A digital product or service is a complex—yet not complicated—network with data

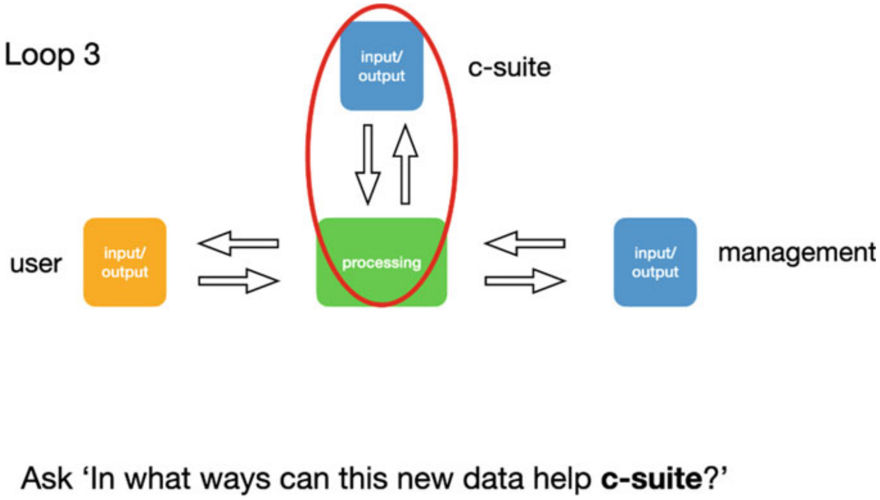


Fig. 8 Use-case scenario 2

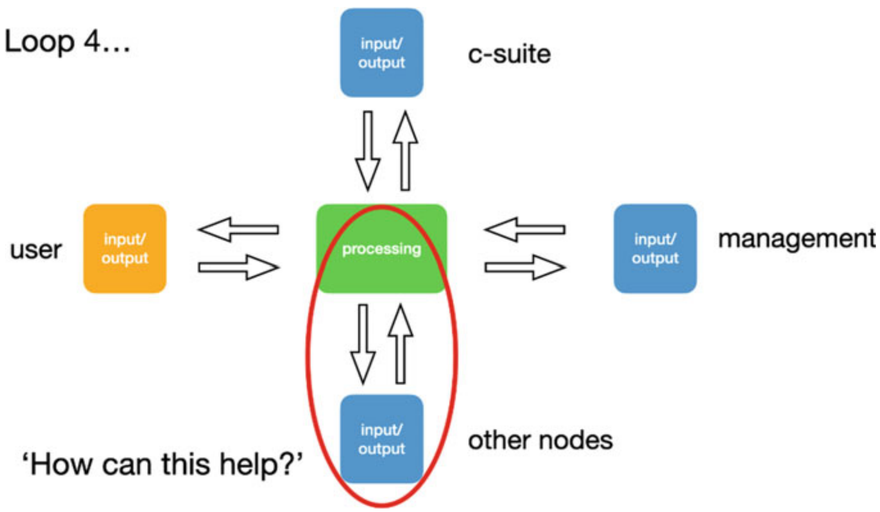


Fig. 9 Use-case scenario 3

input and output. This network is made of nodes (centers or poles) and links (relationships) that connect the different nodes. A specific user group usually represents a node; in the previous example, the "Management" or "C-Suite" are considered a node. Each node is connected to the system, and eventually, other nodes benefit. These nodes represent user-object interactions, which we call use-case scenarios. Nodes are connected to the system via feedback loops—data input, processing, and output. Nodes feed and are fed by the system.

Table 3 The PretoFace V1 structure, the last step, and the building blocks for the PretoVids

Section	Frames	Description
Product Core & Context	Title	In one frame: the name of your product, simple logo, and the subhead "The smart tool …"
	Big Idea	What is this about?
	Problem	Put numbers on it. How much does this problem cost?
	Opportunity	Your solution in 3 lines
	Impact	Are you saving costs? Reducing risks? Increasing profits? Strengthening company values?
	Value Equation	Our product saves ___________€ a year with "name of your product", "your subhead"
Product Description	How it works 1	Show the key interface 1
	How it works 2	Show the key interface 2
	How it works 3	Show the key interface 3
Use-Case Scenario A	Use-case scenario A intro	What can the user in "Management" do?
	PretoFace A1	What does the user see on the screen? What is the user input? Sketch buttons, input text boxes, or other actionable elements for the user
	PretoFace A2	What is happening in the backend?
	PretoFace A3	What does the user see as an output to this specific use-case?
Use-Case Scenario B	Use-case scenario B intro	What can the user in the "C-Suite" do?
	PretoFace B1	What does the user see on the screen? What is the user input? Sketch buttons, input text boxes, or other actionable elements for the user
	PretoFace B2	What is happening in the backend?
	PretoFace B3	What does the user see as an output to this specific use-case?
Use-Case Scenario C	Use-case scenario C intro	What can other users (nodes do?
	PretoFace C1	What does the user see on the screen? What is the user input? Sketch buttons, input text boxes, or other actionable elements for the user
	PretoFace C2	What is happening in the backend?
	PretoFace C3	What does the user see as an output to this specific use-case?

(continued)

Table 3 (continued)

Section	Frames	Description
Product Core	Value Equation	Our product saves ___________€ a year with "name of your product", "your subhead"
	Big Idea	What is this about? What does the product do?
	Title	Name, logo, and subhead

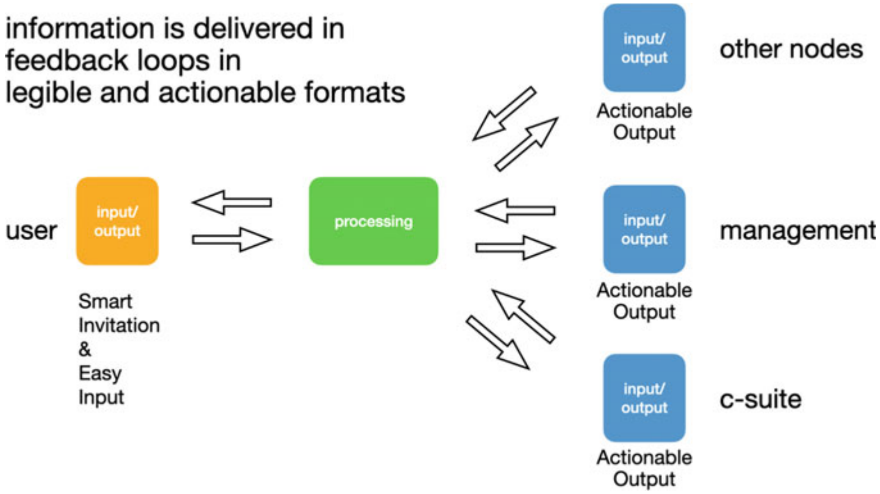


Fig. 10 Schematic for integrating user input with processing nodes with feedback loops

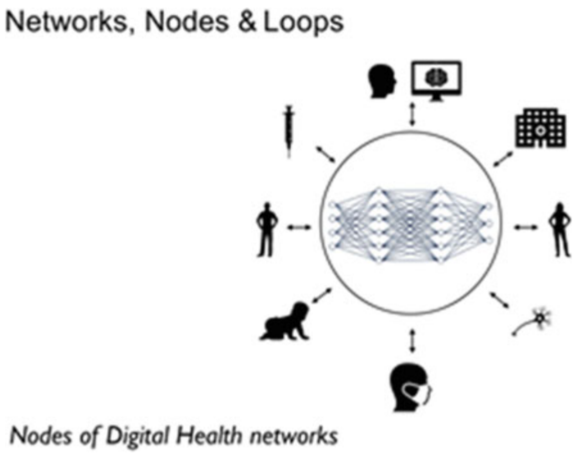
The PretoVids step-by-step guideline takes the designers and teams to account for human-centered networks when designing digital products or services.

For example, in digital health, a network can be developed around the patient's medical records. The nodes could be patients, their families, nurses, medical staff, facility management, public office, health insurance, research centers, etc. (see Fig. 11). Each node is linked to the UIPO System through a specific interface specially designed for the use-case. Each node (i.e., the patient) can interact with the network, providing data inputs and benefiting from generating data. These are the loops that keep the network running.

13 Discussion

This work's contribution is the introduction of PretoVids, a novel approach to digital prototyping and prototyping the digital. As a prototyping form, PretoVids are at the

Fig. 11 A visual example of Digital Health network nodes and loops. Different user groups benefit from a UIPO System in different use-case scenarios



heart of the new product development process. This is even more relevant when it comes to digital products and services. The material constraints of prototyping challenge designing teams to explore the complexity of networked products and services which may have multiple users while contributing to and benefiting from the same network yet with radically different use-case scenarios. In this sense, PretoVids are an essential element of the repertoire of skills that successful designers employ.

Building on the Media Models (Edelman 2011), the media in use while designing serves to communicate and scaffolds certain behaviors in the team. In this sense, PretoVids guide the team to explicitly design for networks, to think in terms of interconnected nodes and feedback loops. Thus, the PretoVids serves as a scaffolding mechanism for the design team to envision complex digital systems in simple ways. Understanding—and sketching out—a basic system, its core function, and how it works opens up multiple possibilities to create value within an organization. Furthermore, the media employed—PretoStory and PretoFaces—pushes the design team to envision possible use-case scenarios and externalizes them using simple sketches. Such a process guides a digital team to reflect on how use-case scenarios crystalize into user interfaces.

In the quest to generate rich data to make better decisions in the early stage of the design process, innovation units tend to avoid early prototypes and move forward. Often, they lack a "thing" to show to customers or potential users that could provide valuable feedback. Using the PretoVids step-by-step structure, teams can be confident that they are creating a "thing" to show and engage customers. In this sense, teams can run demos and interviews using the PretoVids to kick-off the sessions and provide the customer with feedback. This is of particular relevance when engaging with customers online or via video-conferencing platforms because that interaction lacks context; thus, the PretoVids provide cues to the customer to know what is expected. In the same sense, the chance of losing interest in an online session is high, yet watching a short, animated video keeps people's attention and solicits a reaction from them. The PretoStories, PretoFaces, and PretoVids become tangible assets within the organization beyond the design team or project. In this sense, PretoVids can be easily handed-off to other groups within the company (engineering, marketing, UI/UX). This serves as a communication vehicle across departments and disciplines and facilitates alignment in the often bumpy cross-department collaboration.

Moreover, the PretoVids are innovation drivers within organizations, playing a similar role as a boundary learning object and communication driver. However, in the sense that they can be easily shared, sent, modified, and extended, PretoVids have a significant advantage because they leverage digital generativity and openness. This gives them a much more versatile existence. In the context of the COVID-19 pandemic during the year 2020, this versatility proved to be key to the success of the innovation programs run in digital environments. This adds value to designing teams during a crisis or exceptional situations and, in general, when distributed teams engage in prototyping remotely.

Traditionally, video formats were expensive and reserved for experts that knew how to use sophisticated video editing software. Luckily, nowadays, any person with a smartphone—roughly one-third of the world's population by the end of 2020 (Statista 2020)—can easily capture and edit a video in a matter of seconds or minutes. And thus, the power of digital technologies is channeled into a generative engine that people can access.

14 Closing Words

The PretoVids, a new approach to digital prototyping and prototyping the digital, is framed within the ongoing effort to bridge the research to practice gap. This effort lies at the core of our activities and is a continuum with the performative patterns training packages and material. With the PretoVids, we explored, conceptualized, theorized, and tested a new form to create value from the same building principles. In this sense, we are not far from the systems design approach proposed above. Furthermore, we are creating a network of research-based training material, performative patterns that can create value in different situations, depending on the design team's specific use-case. Performative patterns can be seen as a network of nodes in which designers can find endless possibilities for designing. Now PretoVids joins this repertoire of highly effective tools for both beginners and experienced designers as well.

The authors would like to acknowledge and thank the contributions of our colleagues and collaborators. Thanks to Anne Victoria Talbot for her 'Whole lot of Love' and her work in performative patterns. Thank you, Fabrizio Pierandre at Politecnico di Milano and PACO, for his significant and long-time contribution to our work in Performative Patterns and PretoVids. Fabrizio, you are a dream collaborator. Thank you to the Hasso Plattner Foundation for the great support, Professor Christoph Meinel for making our work possible, and to the HPI Digital Health Center community, particularly Professor Erwin Böttinger, Professor Christoph Lippert, and Dr. Stefan Konigorski, for your support, encouragement, and enthusiasm about our work. Thank you to Stanford's Center for Design Research and Professor Larry

Leifer, and Professor Dave Beach, and Senior Lecturer Craig Milroy of Stanford's Product Realization Lab. Thank you to our students at Politecnico Milano and at the Digital Health Design Lab: sometimes the eggs teach the chickens. Thank you to the visionary leadership at SAP for their trust, support and support and support, and courage to innovate. Thank you to the coaching team at the HPI Academy, to the amazing Katrin Lütkemöller Shaw for helping us turn new theory into professional workshop-ready teaching material. Thank you to Dr. Sharon Nemeth for language support. And finally, thank you, Dr. Timm Krohn, who has encouraged and supported our research from the beginning, and his commitment to bringing research insights to a wider professional audience.

Appendix

Reflection from working with the PretoVids during the COVID-19 pandemic with a Software company, excerpt extracted from an interview (SAP 2020):

> "We are very much part of SAP's approach to driving innovation, which is to build an innovation ecosystem. There are several elements to the one that is changing the nature of the relationship between the partner and the partner-business manager to become co-creators of new products. The approach is to inspire and support new knowledge facilitation to create not just product relationships that are important but also a culture of innovation in their organizations. SAP is creating a culture of innovation with three elements. The first one is space–that is the space around you, and it needs to support innovation. Working in cubicles or working behind closed doors generally does not support innovation, and you can see it at this facility. What they have done is amazing. It is a beautiful, really well-considered space that is flexible for innovation. The second element involves people's behaviors. It is not just about ideas, but what you do in these spaces. This element deals with teaching people ways of working in these spaces that they may not be accustomed to and want them to get comfortable with. The third element is what kind of stories we are telling. We can ask ourselves: are we a company that is satisfied doing the norm, or are we an innovative company that takes chances and risks and then test them and blazes new trails to places nobody has ever been. Because innovation is really about going places, nobody has been. Three other elements stand out when creating innovation. The first is having a good business model. The second is understanding human factors because innovation has to do with bringing technology to people. The third is, of course, new technology. Those three elements come together in PretoVids, and innovation is born because it is a place where human factors, business factors, and technical factors come

together. PretoVids brings employees closer to job satisfaction in the knowledge that they can engage in jobs not only for our end-users. For the people working in companies who are doing the innovation, these can be for the home, workplace, or medical field. An innovative culture is an exciting culture to be in. Not only are you making the world a better place, but you are also creating new value, and you are going places nobody has ever gone before. We hope PretoVids can augment current approaches to inspire excitement in people working on their projects. They learn and discover new things, work with people in their field and craft, as a digital or virtual team, this new solution. Innovation creates all kinds of value, including financial, personal, and material value. For us, PretoVids lie at the center of design in the digital era, i.e., it augments how product teams make sense, meaning, and change in a way that rewards all actors. With the approach designers (new and experienced) with a feeling of satisfaction in e creating new experiences and value for people in the world, and engaging with technical people to reimagine all sorts of new ways that objects and interactions can."

References

Andriole, S. J. (1994). Fast, cheap requirements prototype, or else! *IEEE Software, 11*(2), 85–87.

Baldassarre, B., Konietzko, J., Brown, P., Calabretta, G., Bocken, N., Karpen, I. O., & Hultink, E. J. (2020). Addressing the design-implementation gap of sustainable business models by prototyping: A tool for planning and executing small-scale pilots. *Journal of Cleaner Production, 255.*

Beckman, S. L., & Barry, M. (2007). Innovation as a learning process: Embedding design thinking. *California Management Review, 50*(1), 25+.

Beltagui, A. (2018). A design-thinking perspective on capability development: The case of new product development for a service business model. *International Journal of Operations & Production Management, 38*(4), 1041–1060.

Bharathi, V. S., & Pande, M. (2019). Why design thinking in IT business management program? An exploratory study on students' design thinking learning experience. *International Journal of Information and Communication Technology Education, 15*(4), 76–93.

Borge, M., Toprani, D., Yan, S., & Xia, Y. (2020). Embedded design: Engaging students as active participants in the learning of human-centered design practices. *Computer Science Education, 30*(1), 47–71.

Breuer, H., & Lüdeke-Freund, F. (2017). Values-based network and business model innovation. *International Journal of Innovation Management, 21*(03), 1750028.

Brown, T. (2008). Design thinking. *Harvard Business Review, 86*(6), 84+.

Brown, T., & Katz, B. (2009). *Change by design: How design thinking transforms organizations and inspires innovation/Tim Brown; with Barry Katz* (1st ed.). . Hsrper Collins.

Brown, T., & Wyatt, J. (2010). Design thinking for social innovation (SSIR). Available at: https://ssir.org/articles/entry/design_thinking_for_social_innovation. Accessed December 22, 2020.

Buchanan, R. (2019). Surroundings and environments in fourth order design. *Design Issues, 35*(1), 4–22.

Buchanan, R., Doordan, D., & Margolin, V. (2001). The development of design thinking in our time—Introduction. *Design Issues, 17*(4), 1–2.
Buxton, B. (2007). Sketching user experiences: Getting the design right and the right design.
Canedo, E. D., & Da Costa, R. P. (2018). The use of design thinking in agile software requirements survey: a case study.
Capata, R., & Tatti, F. (2020). Designing, prototyping, assembling and costs analysis of a gas turbine hybrid vehicle. *Energies, 13*(18).
Caprari, E., Porsius, J. T., D'Olivo, P., Bloem, R. M., Vehmeijer, S. B. W., Stolk, N., & Melles, M. (2018). Dynamics of an orthopaedic team: Insights to improve teamwork through a design thinking approach. *Work-A Journal of Prevention Assessment & Rehabilitation, 61*(1), 21–39.
Carayannis, E. G., Sindakis, S., & Walter, C. (2015). Business model innovation as lever of organizational sustainability. *The Journal of Technology Transfer, 40*(1), 85–104.
Carlgren, L., Rauth, I., & Elmquist, M. (2016). Framing design thinking: The concept in idea and enactment. *Creativity and Innovation Management, 25*(1), 38–57.
Carroll, N., & Richardson, I. (2016). Aligning healthcare innovation and software requirements through design thinking.
Casakin, H. (2006). Assessing the use of metaphors in the design process. *Environment and Planning B: Planning and Design, 33*, 253–268.
Choi, H. H., & Kim, M. J. (2017). The effects of analogical and metaphorical reasoning on design thinking. *Thinking Skills and Creativity, 23*, 29–41.
Cienki, A., & Müller, C. (2012). Metaphor, gesture, and thought. In R. W. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 483–501). Cambridge University Press, Cambridge.
Cockton, G., Larusdottir, M., Gregory, P., & Cajander, A. (2016). Integrating user-centred design in agile development. In Cockton, G., Larusdottir, M., Gregory, P., & Cajander, A. (Eds.), *Integrating user-centred design in agile development*, *human/computer interaction a series of monographs, edited volumes and texts* (pp. 1–46).
Curedale, R. (2019). *Design thinking: Process et methods* (5th ed.). . Design Community College Inc.
Dahlander, L., & Gann, D. M. (2010). How open is innovation? *Research Policy, 39*(6), 699–709.
de Goey, H., Hilletofth, P., & Eriksson, L. (2017). Design-driven innovation: Making meaning for whom? *The Design Journal, 20*(sup1), S479–S491.
Dell'Era, C., Altuna, N., & Verganti, R. (2018). Designing radical innovations of meanings for society: Envisioning new scenarios for smart mobility. *Creativity and Innovation Management, 27*(4), 387–400.
Dong, A., Kleinsmann, M. S., & Deken, F. (2013). Investigating design cognition in the construction and enactment of team mental models. *Design Studies, 34*(1), 1–33.
Dow, S. P., Heddleston, K., & Klemmer, S. R. (2009). The efficacy of prototyping under time constraints. In *Proceedings of the seventh ACM conference on Creativity and cognition*, ACM Digital Library.
Dow, S., Fortuna, J., Schwartz, D., Altringer, B., Schwartz, D., & Klemmer, S. (2011). Prototyping dynamics. In *29th Annual CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York.
Drechsler, W., & Natter, M. (2012). Understanding a firm's openness decisions in innovation. *Journal of Business Research, 65*(3), 438–445.
Dunne, D. (2018). Implementing design thinking in organizations: an exploratory study. *Journal of Organization Design, 7*(1), 16.
Durao, L. F. C. S., Kelly, K., Nakano, D. N., Zancul, E., & McGinn, C. L. (2018). Divergent prototyping effect on the final design solution: The role of "Dark Horse" prototype in innovation projects.
Edelman, J. (2011). Understanding radical breaks. Media and behavior in small teams engaged in redesign scenarios. Doctoral Dissertation, Stanford University, 2011.

Edelman, J., Agarwal, A., Paterson, C., Mark, S., & Leifer, L. (2012). Understanding radical breaks. In H. Plattner, C. Meinel, & L. Leifer (Eds.), *Design thinking research: Studying co-creation in practice, understanding innovation* (Vol. 37, pp. 31–51). Springer.

Edelman, J. A., Owoyele, B., Santuber, J., Talbot, A. V., Unger, K., & von Lewinski, K. (2020). Accessing highly effective performative patterns. In: Meinel, C. & Leifer, L. (Eds.), *Design thinking research, understanding innovation* (pp. 15–33). Springer International Publishing, Cham.

Efeoglu, A., Moller, C., & Serie, M. (2014). Solution prototyping with design thinking—Social media for SAP store: A case study.

Eickhoff, F. L., McGrath, M. L., Mayer, C., Bieswanger, A., & Wojciak, P. A. (2018). Large-scale application of IBM Design Thinking and Agile development for IBM z14. *IBM Journal of Research and Development, 62*(2-3).

Eris, O. (2003). *Effective inquiry for innovative engineering design*. Springer.

Fleischmann, K. (2019). Design-led innovation and circular economy practices in regional Queensland. *Local Economy, 34*(4), 382–402.

Gerber, E. (2010). Prototyping practice in context the psychological experience in a high tech firm. *Journal of Design Studies.*

Gomez, A. (2015). Risk, uncertainty, and the dimensions of technological rationality. In Gonzalez, W. J. (Ed.), *New perspectives on technology, values, and ethics: Theoretical and practical*, *Boston Studies in the Philosophy and History of Science* (Vol. 315, pp. 105–121).

Gurusamy, K., Srinivasaraghavan, N., & Adikari, S. (2016). An integrated framework for design thinking and agile methods for digital transformation.

Haldane, V., Koh, J. J. K., Srivastava, A., Teo, K. W. Q., Tan, Y. G., Cheng, R. X., Yap, Y. C., Ong, P. S., van Dam, R. M., Foo, J. M., Mueller-Riemenschneider, F., Koh, G. C. H., Foong, P. S., Pere, P., & Legido-Quigley, H. (2019). User preferences and persona design for an mhealth intervention to support adherence to cardiovascular disease medication in Singapore: A multi-method study. *JMIR mhealth and uhealth, 7*(5).

Hey, J., Linsey, J., Agogino, A. M., & Wood, K. L. (2008). Analogies and metaphors in creative design. *International Journal of Engineering Education, 24*(2), 283–294.

Hopkins, I. G., Dunn, K., Bourgeois, F., Rogers, J., & Chiang, V. W. (2016). From development to implementation-A smartphone and email-based discharge follow-up program for pediatric patients after hospital discharge. *Healthcare-The Journal of Delivery Science and Innovation, 4*(2), 109−115.

Jablokow, K. W., Sonalkar, N., Edelman, J., Mabogunje, A., & Leifer, L. (2019). Investigating the influence of designers' cognitive characteristics and interaction behaviors in design concept generation. *Journal of Mechanical Design, 141*(9), 377.

Kim, H. H., & Hinds, P. (2016). Iterative prototyping for team creativity: A cross-cultural examination. *Academy of Management Proceedings, 1*, 12335.

Kirsh, D. (2000). A few thoughts on cognitive overload. Intellectica. *Revue de l'Association pour la Recherche Cognitive, 30*(1), 19−51.

Kirsh, D. (2010). Thinking with external representations. *AI & Society, 25*(4), 441–454.

Kirsh, D. (2011). How marking in dance constitutes thinking in the body. *Versus: Quaderni di studi semiotici*, pp. 183–214.

Laptev, G., & Shaytan, D. (2019). Product managers for the digital and collaborative age.

Leifer, L., Plattner, H., & Meinel, C. (Eds.) (2014). *Design thinking research: Building innovation eco-systems, understanding innovation*. Springer International Publishing; Imprint: Springer, Cham.

Levy, M., & Hadar, I. (2018). Multidisciplinary requirements engineering for addressing social-oriented concerns.

Liedtka, J. (2018). Why design thinking. *Harvard Business Review, 96*(5), 72–79.

Liedtka, J. (2020). Putting technology in its place: Design thinking's social technology at work. *California Management Review, 62*(2), SI, 53–83.

Lillie, E. O., Patay, B., Diamant, J., Issell, B., Topol, E. J., & Schork, N. J. (2011). The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Personalized Medicine, 8*(2), 161−173.
Mabogunje, A. (1997). *Noun phrases as surrogates for measuring early phases of the mechanical design process. In Proceedings of the 9th International Conference on Design Theory and Methodology.*
Marks, J., & Chase, C. C. (2019). Impact of a prototyping intervention on middle school students' iterative practices and reactions to failure. *Journal of Engineering Education, 108*(4), 547–573.
Martins, H. F., de Oliveira Junior, A. C., Canedo, E. D., Dias Kosloski, R. A., Paldes, R. A., & Oliveira, E. C. (2019). Design thinking: Challenges for software requirements elicitation. *Information, 10*(12).
McShane, M. (2018). Enterprise risk management: History and a design science proposal. *Journal of Risk Finance, 19*(2), 137–153.
Meadows, D. (2008). *Thinking in systems: A primer*.
Meadows, D., & Wright, D. (2008). *Thinking in systems: A primer*. In Meadows, D. H. & Wright, D. (Eds.), White River Junction, VT: Chelsea Green Pub.
Meinel, M., Eismann, T. T., Baccarella, V., Christian, F., & S.K. and Voigt, K.-I. . (2020). Does applying design thinking result in better new product concepts than a traditional innovation approach? An experimental comparison study. *European Management Journal, 38*(4), 661–671.
Menning, A., Ewald, B., Nicolai, C., & Weinberg, U. (2017). *"… and not building on that": The relation of low coherence and creativity in design conversations* (pp. 195–213). Springer International Publishing.
Mesquida, A. L., Karac, J., Jovanovic, M., & Mas, A. (2017). A game toolbox for process improvement in agile teams.
Morassi Sasso, A., Datta, S., Jeitler, M., Steckhan, N., Kessler, C. S., Michalsen, A., Arnrich, B., & Böttinger, E. (2020). HYPE: Predicting blood pressure from photoplethysmograms in a hypertensive population. In M. Michalowski, R. Moskovitch (Eds.), Artificial intelligence in medicine: 18th International Conference on Artificial Intelligence in Medicine, AIME 2020, Minneapolis, MN, USA, August 25-28, 2020, Proceedings/Martin Michalowski, Robert Moskovitch (eds.), LNCS sublibrary, SL 7, Artificial intelligence (Vol. 12299, pp. 325–335) Cham: Springer.
Mukherjee, R. A. S., Hollins, S., & Turk, J. (2006). Fetal alcohol spectrum disorder: An overview. *Journal of the Royal Society of Medicine, 99*(6), 298−302.
Muntanyola-Saura, D., & Kirsh, D. (2010). Marking as physical thinking: A cognitive ethnography of dance.
Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital innovation management: Reinventing innovation management research in a digital world.
Nambisan, S., Wright, M, & Feldman, M. (2019). The digital transformation of innovation and entrepreneurship: Progress, challenges and key themes. *Research Policy, 48*(8), 103773.
Newman, P., Ferrario, M. A., Simm, W., Forshaw, S., Friday, A., & Whittle, J. (2015). The role of design thinking and physical prototyping in social software engineering.
O'Driscoll, K. (2016). The agile data modelling & design thinking approach to information system requirements analysis. *Journal of Decision Systems, 25*(1), SI, 632–638.
Oehlberg, L., Simm, K., Jones, J., Agogino, A., & Hartmann, B. (2012). Showing is sharing. In 2012, D. I. S. C. (Ed.), *Proceedings of the Designing Interactive Systems Conference*, New York, NY: ACM.
Penzenstadler, B. (2020). When does design help thinking, and when does design thinking help? *IEEE Software, 37*(2), 6--9, March-April 2020, https://doi.org/10.1109/MS.2019.2958263.
Pereira, J. C., & Russo, Rosaria de F. S. M. (2018). Design thinking integrated in agile software development: A systematic literature review.
Rietveld, E., Denys, D., & van Westen, M. (2018). Ecological-enactive cognition as engaging with a field of relevant affordances. In A. Newen, L. de Bruin, & S. Gallagher (Eds.), *The Oxford handbook of 4E cognition, Oxford library of psychology* (1st ed., pp. 40–70). Oxford University Press.

SAP. (2020). Inside Track 2020. In Review: Top Ten, 10x10 New Product Development Sprint, available at: https://hpi.de/fileadmin/user_upload/upload_academy/SAP_Inside_Track_2020__2_-_Page_22-25.pdf Accessed 22 Oct 2020.
Savoia, A. (2019). *The right it: Why so many ideas fail and how to make sure yours succeed/Alberto Savoia*, New York, NY: HarperOne.
Seidel, V. P. (2007). Concept shifting and the radical product development process. *Journal of Product Innovation Management, 24*(6), 522–533.
Sepp, S., Howard, S. J., Tindall-Ford, S., Agostinho, S., & Paas, F. (2019). Cognitive load theory and human movement: Towards an integrated model of working memory. *Educational Psychology Review, 31*(2), 293−317.
Shalpegin, T., Sommer, S., & Wan, Z. (2018). Collaborative prototyping of alternative designs under a target costing scheme. *Production and Operations Management, 27*(3), 496–515.
Statista (2020). Smartphone users 2020. Statista. Number of smartphone users worldwide from 2016 to 2021(in billions). Available at: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/. Accessed December 23, 2020.
Steinbeck, R. (2011). Building creative competence in globally distributed courses through design thinking. *Comunicar, 19*(37), 27–34.
Succi, M. D., Uppot, R. N., Gee, M. S., McLoud, T. C., & Brink, J. A. (2018). Medically engineered solutions in health care: a technology incubator and design- thinking curriculum for radiology trainees. *Journal of the American College of Radiology, 15*(6), 892−896.
Tversky, B. (2003). Structures of mental spaces. *Environment and Behavior, 35*(1), 66–80.
Tversky, B. (2019). Mind in motion. In Dow, S., Maher, M. L., Kerne, A., & Latulipe, C. (Eds.), *C&C'19: Proceedings of the 2019 Creativity and Cognition June 23–26, 2019, San Diego, CA, USA, 23 06 2019 26 06 2019, San Diego CA USA* (p. 107). New York, NY: ACM.
Vechakul, J., Shrimali, B. P., & Sandhu, J. S. (2015). Human-centered design as an approach for place-based innovation in public health: A case study from Oakland, California. *Maternal and Child Health Journal, 19*(12), 2552−2559.
Verganti, R. (2008). Design, meanings, and radical innovation: A metamodel and a research agenda. *Journal of Product Innovation Management, 25*(5), 436−456.
Verganti, R., Vendraminelli, L., & Iansiti, M. (2020). Innovation and design in the age of artificial intelligence. *Journal of Product Innovation Management, 37*(3), 212−227.
Vetterli, C., Brenner, W., Uebernickel, F., & Petrie, C. (2013). From Palaces to Yurts why requirements engineering needs design thinking. *IEEE Internet Computing, 17*(2), 91–94.
Voinea, M. (2019). The development of future competences—A challenge for the educational system. *Revista Romaneasca Pentru Educatie Multidimensionala, 11*(4), 1, 328–336.
Wenzel, M., Gericke, L., Thiele, C., & Meinel, C. (2016). Globalized design thinking: Bridging the gap between analog and digital for browser-based remote collaboration. In H. Plattner, C. Meinel, L. Leifer (Eds.), Design Thinking Research: Making Design Thinking Foundational, Understanding innovation, pp. 15−33.
Woods, L., Cummings, E., Duff, J., & Walker, K. (2017). Design thinking for mHealth application co-design to support heart failure self-management.
Woods, L., Cummings, E., Duff, J., & Walker, K. (2018). Partnering in digital health design: Engaging the multidisciplinary team in a needs analysis.
Yan, M.-R. (2018). Improving entrepreneurial knowledge and business innovations by simulation-based strategic decision support system. *Knowledge Management Research & Practice, 16*(2), 173–182.

Drawventure: Teaching Design Sketching Through Gameplay

Catherine Mullings, Camille Utterback, and Michael Bernstein

Abstract How can we make design sketching a more engaging, bite-sized practice in our daily schedules? Learning design sketching typically involves repeated practice of fundamentals. This approach can be discouraging and demotivating. We want to enable more self-directed learning and thereby democratize the design sketching education. In this project, we paper-prototyped and user-tested an application called Drawventure that reframes sketching lessons as micro-challenges, where player sketches become objects that populate an increasingly engaging toy world.

1 Introduction

How can we increase the number of people who engage in design thinking in their everyday lives? Visual communication, and in particular design sketching, is a core skill of design thinking that many find challenging to learn (Williford, 2017). At Stanford, design sketching is important enough to learn that the Product Design major devotes an entire required course to it: ME 110. As with learning a musical instrument, learning design sketching involves the repeated practice of fundamentals via mastery learning (Block & Burns, 1976). Outside of formal curriculums, however, this mastery learning approach can make progress feel distant, leading many interested learners to become discouraged and demotivated (Williford et al., 2017).

Our goal is to enable more self-directed learning and thereby democratize design sketching education and visual communication. Achieving this goal requires that

C. Mullings · C. Utterback · M. Bernstein (✉)
Stanford University, 353 Jane Stanford Way, Stanford, CA 94305, USA
e-mail: msb@cs.stanford.edu

C. Mullings
e-mail: cmulling@stanford.edu

C. Utterback
e-mail: camilleu@stanford.edu

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_12

we make learning design sketching more enjoyable, and, in particular, that we overcome the lack of visible short-term progress that discourages self-taught learners. We introduce an approach that recasts the mastery learning practice of design sketching into a series of activities that populate a virtual world. Activities begin simple, for example, sketching circles to create marbles and boxes to create toy blocks, each of which begins to flesh out an imaginary toy world (Fig. 1). Lessons are embedded in this population mechanic, and exercises are micro-challenges whose results become artifacts added to the world. This approach is manifested in Drawventure, an application for learning design sketching that we are constructing and will deploy publicly. Drawventure structures the lessons, the challenges, and the imaginary toy world, and captures learners' sketches through uploaded pictures from their sketchbooks.

Mastery learning requires both repetition and feedback (Block & Burns, 1976). We hypothesize that this approach of embedding a design sketching curriculum into short tasks toward the population of a toy world will increase the learners' sense of progress, and through this, their commitment and resulting skills.

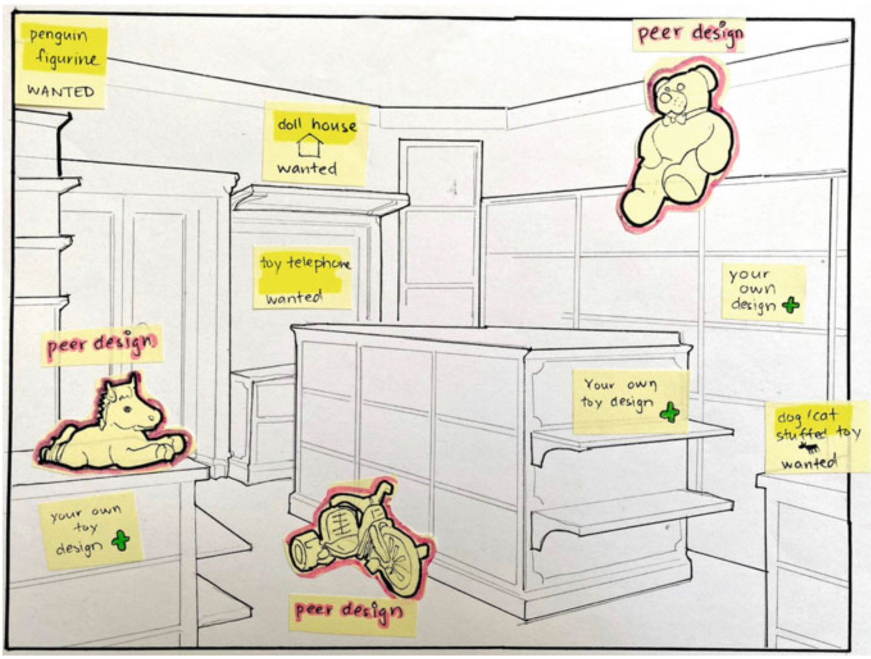


Fig. 1 An early-stage paper prototype of a toy store that needs to be stocked with items in Drawventure's imaginary toy world. In Drawventure, learners are presented challenges to design and sketch various toys (e.g., a dollhouse or stuffed animal) that exercise form giving fundamentals. Their sketch will populate the toy store. Additionally, learners can view other peer design sketches to see the techniques and processes that the peers used to achieve their final version

Our goal in prototyping Drawventure was to create the first thirty minutes to hour long experience of the game. This chapter captures our progress in developing the game's narrative, visual design, and core game mechanic.

2 Game Design

The initial setting for the game is a toy store (Fig. 1). Players practice basic shapes such as cubes and cylinders. These parts can then be joined to create simple toys (Fig. 2) and populate the store. Through this setting, we can teach the most basic building blocks of design sketching.

The core mechanic of the game, or the "core loop," is the actions that the player will perform the most often. Successful games require a strong core loop that motivates the player to continue. In our case, this reinforces self-directed learning.

The core mechanism behind Drawventure is to learn design sketching through many micro-activities, each of which provides a direct sense of progress to the learner. To achieve this, we set a design constraint that learners will use Drawventure for no more than 30 min in a given session. Drawventure then presents the learner with a toy world where you design and draw objects to unlock skills and missions as you venture into the various themed areas of the world.

Drawventure begins by teaching learners core sketching techniques and primitive forms like cubes and ellipses. As a part of the game, these lessons and practice exercises are framed as unlocking tools or skills. The tools allow learners to interact with the toy world. Over time with more skills in their toolbelt, learners can then accomplish missions, advancing and exploring farther in the toy world. Missions are micro-challenges, whereby Drawventure will guide learners on how to use their tools and combine primitive forms to gradually create more complex and contextually meaningful shapes, like a car or a helmet. Overall, as learners master skills, Drawventure transitions from tasks that are more technique-focused (e.g., drawing

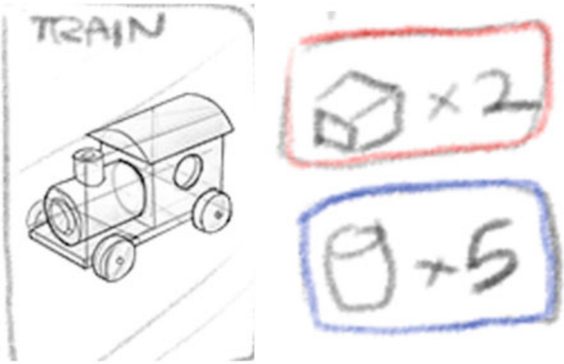


Fig. 2 Learners are presented with a toy object to draw (left), in this case a train. To draw the train, the learner must draw two cubes and five cylinders (right)

an open ring case) to more concept-focused (e.g., generating and sketching a creative concept for a teddy bear).

The high-level flow is as follows (Fig. 3). First, the learner opens the Drawventure application and navigates to a part of the Drawventure world. Here, Drawventure gives learners a prompt or mission ("Design and draw a toy for the toy store") to complete a design sketching task with paper and pen. Then learners use their phone's camera to photograph their drawing and upload it to Drawventure, where it enters Drawventure's peer correction workflow. After learners receive feedback, they will make any necessary revisions, repeating the redo-submit-feedback loop as needed. When the final version is ready and passed by the peer correction flow, the original mission will be marked as completed and the learner will advance in the game.

As we prototype and develop the core Drawventure mechanics, we also hope to supplement solo activities and incorporate a collaborative component. We draw upon the fact that collaboration remains a key part of traditional design sketching classes and even professional design work. In Drawventure, collaboration will manifest itself in three ways: giving and receiving peer feedback; seeing and comparing multiple solutions to a design prompt; and finally working together to tackle a design problem.

As the drawer completes tasks, their sketches become a part of the world. For example, the drawer might finish the practice sketches for the cube. Then, the output

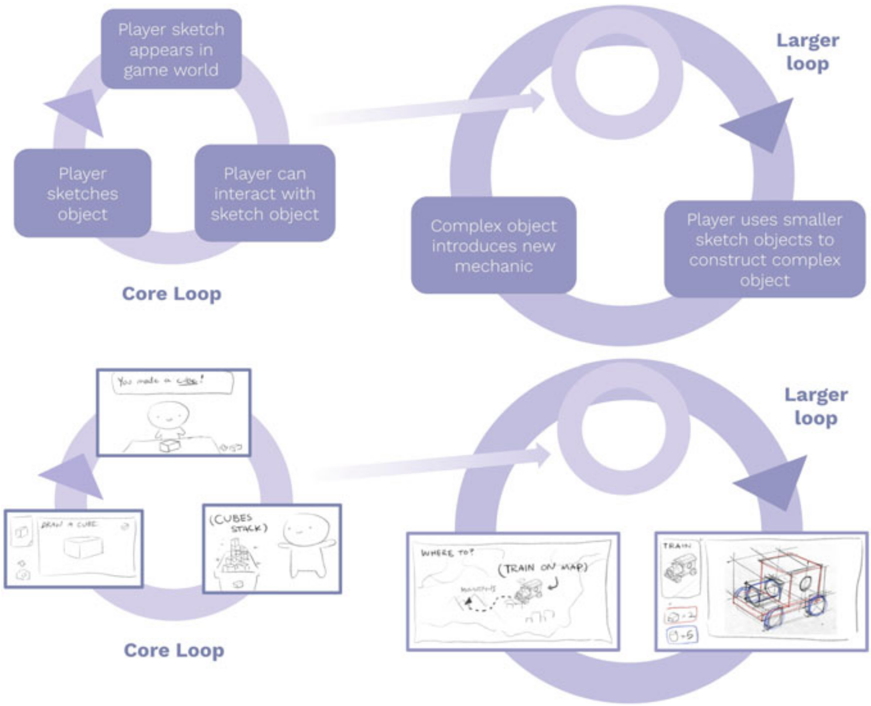


Fig. 3 The core game mechanic of Drawventure

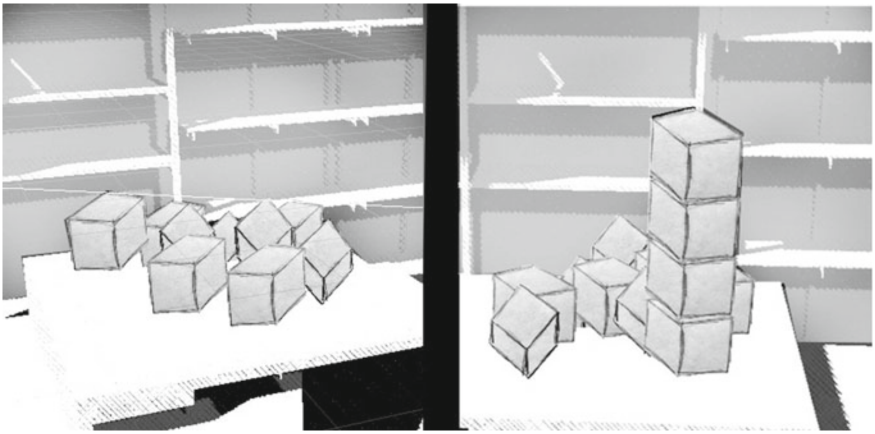


Fig. 4 Implementation of cubes (left) which can later be stacked by the player (right)

from that activity—the cube—is populated into the world, with cubes littering the table of the toy store (Fig. 4, left). The cubes become interactive: as the user taps them, the cubes stack (Fig. 4, right).

Our goal was to develop storyboards of how a user would engage with Drawventure in 30-min sequences per day. We developed a series of storyboards to explore alternative concepts (Fig. 5). These storyboards generally involved the following sequence of events:

1. The user views lessons that teach them a tool, such as a specific drawing technique (e.g., a hinge).
2. The user engages in a practice task that evaluates that technique in isolation.
3. After learning several techniques, the user is given a summative task that combines techniques in novel ways. For example, the task might require combining ellipses in perspective with boxes in perspective to create a simple toy train.
4. At the end of the week, the output of this task (e.g., the train) is populated into the virtual world.

Given this sequence, our next step is to translate this storyboard into a first week set of lessons and tasks.

3 Visual Development

In terms of visual design and aesthetic, we decided on a hybrid 3D/2D world (Fig. 4). The idea is that the user is in a 3D world and their drawings populate the world as 2D objects. So, as the user continues to participate in Drawventure, the space becomes more and more populated with people and objects.

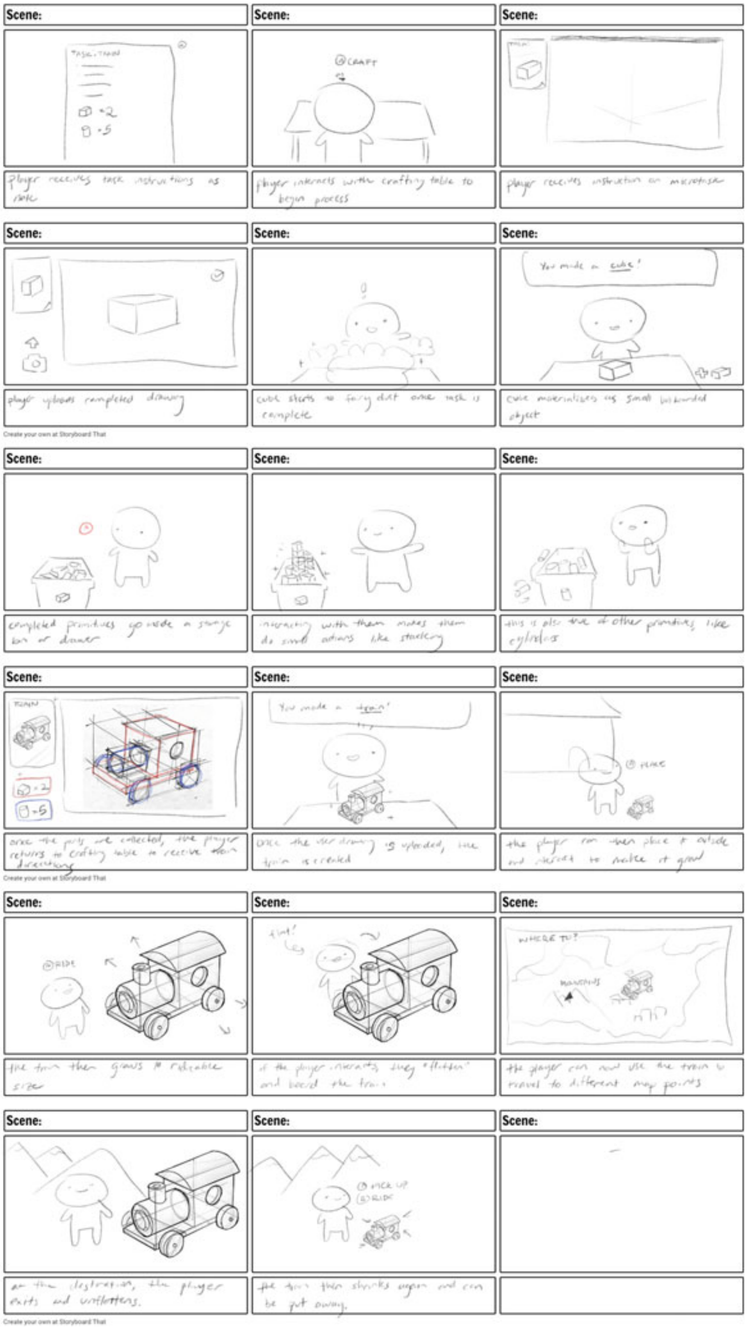


Fig. 5 Storyboards of the full 30 min to hour long experience in the game

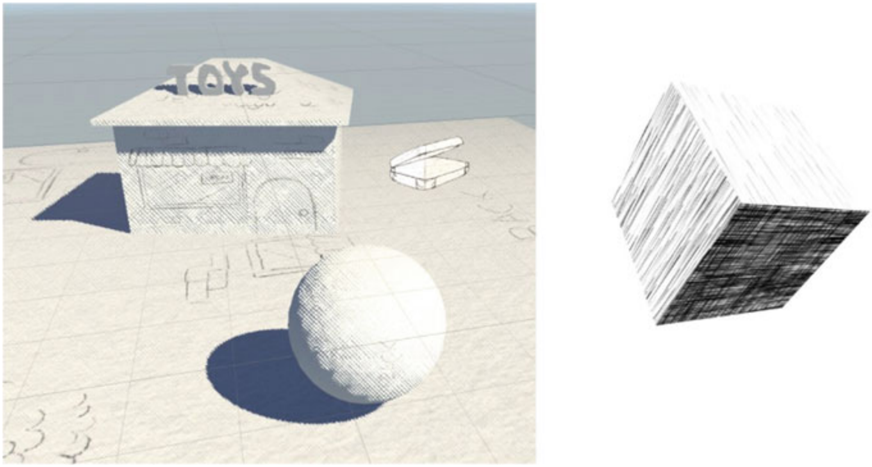


Fig. 6 Implementation of various shaders to establish visual style

Making a convincing 3D/2D aesthetic is challenging. In particular, 2D drawings have a “sketchy” quality that needs to blend in with the surrounding environment. This 3D/2D world would need to have a style with a hand-drawn feel.

We addressed this issue by exploring shaders. Shaders are effects that are applied to 3D objects in the world. Our goal is to make these 3D objects also have a “hand-drawn” feel. To avoid player-created assets looking out of place, a style must be found that spotlights and showcases player-created assets in a unique way. To do this, we implemented several shaders to give 3D modeled assets, such as buildings and furniture, a more hand-drawn feel. Each shader applies a crosshatching to the 3D objects in the scene (Fig. 6).

A second challenge is to populate the sketched 2D objects into this 3D world in such a way that they feel natural. The 2D objects are flat and planar, so they have no natural depth in the 3D scene. We took inspiration from recent video games such as Paper Mario that explore this tension. In particular, we focused on a technique known as “billboarding,” which fixes the 2D object always facing toward the camera (Fig. 7). *No matter where the camera moves around the scene, the sketched objects appear to remain flat against the camera as if rotating alongside the camera.*

4 Discussion

Drawventure takes inspiration from successes such as Duolingo, which identified that skills can be learned in short daily sprints through viscerally enjoyable experiences. Unlike Duolingo, we require peer feedback, because sketching cannot be automatically graded. This introduces more of a delay into the system, requiring that users wait potentially several hours or overnight before they can move on. For this

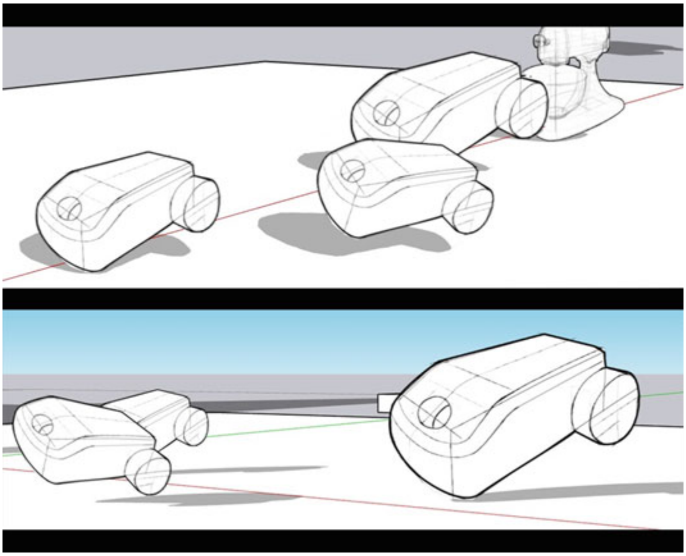


Fig. 7 Implementation of billboarding effect. Both views (top, bottom) have the drawing face toward the camera

reason, we are designing Drawventure around an assumption of "go sleep on it" after completing each task, with feedback available the next day.

Unlike other computationally driven learning systems, Drawventure has an opportunity to lean into its visual medium. This is the reason that we focus on virtual worlds such as physical toy stores and towns. Our hope is that the experience of populating a world with drawings can be much like that of a child populating an imaginary world with toys. In addition, our goal is to convey the user's drawings in a way that makes them feel proud of what they've created, rather than potentially embarrassed.

There are many drawing skills for which Drawventure does not yet provide solutions. It is mostly focused on design sketching, which is generally oriented around drawing lifelike objects based on shapes. It cannot support broader concepts of art, such as the choice of what to create. These require ongoing thought—we could even ask if they are a good fit for an approach like Drawventure.

5 Conclusion

In this chapter, we presented our iterative progress on the creation of Drawventure, a system that engages learners in lessons in design sketching. We are developing the system to proceed through a series of 30-min lessons and tasks, each of which iteratively populates a virtual world with the user's results.

References

Block, J. H., & Burns, R. B. (1976). Mastery learning. *Review of Research in Education, 4*, 3–49.

Williford, B. (2017). Sketchtivity: Improving creativity by learning sketching with an intelligent tutoring system. In: *Proceedings of the 2017 ACM SIGCHI conference on creativity and cognition* (pp. 477–483).

Williford, B., et al. (2017). Zensketch: A sketch-based game for improving line work. In: *Extended abstracts publication of the annual symposium on computer-human interaction in play* (pp. 591–598).

Reflective Design Practice

Adam Royalty

Abstract This chapter describes a procedure for promoting deeper student reflection in design thinking courses. Reflective Design Practice is a sequence of tools that allows students to develop more metacognitive awareness of their growing design abilities. It simultaneously serves as an assessment tool and feedback mechanism for instructors. Unlike existing methods of measurement, Reflective Design Practice evaluates authentic student work over the entire duration of a course. By adhering to three key design principles, instructors can easily adapt Reflective Design Practice to their own contexts.

1 Introduction

Design thinking is a methodology that leverages creativity to solve complex problems (Cross, 2006; Kelley & Kelley 2013). In academic settings spanning kindergarten through post-secondary education, instructors often teach design thinking using project-based learning (Carroll et al 2010; Goldman & Kabayadondo, 2016). This chapter focuses on design thinking in higher education where instructors tend to teach interdisciplinary groups of students how to address challenges through a design process (Kelley & Kelley 2013; von Thienen et al., 2017). One of the primary motivations is to teach basic principles of design to non-designers (Royalty, 2018). Besides equipping students with valuable design tools, these courses can enhance students' creative capacity (Bott et al., 2014; Klenitz et al., 2014; Saggar et al. 2017). The ultimate goal is to prepare students to transfer design thinking beyond the classroom and develop creative solutions to novel problems they encounter in their lives (Bransford & Schwartz, 1999).

Two interrelated components of design thinking education that instructors employ to support the goal of transfer are *feedback* and *assessment*. Instructors capture and measure student work to assess student progress. Feedback mechanisms return the

A. Royalty (✉)
Hasso Plattner Institute of Design (D.School), Building 550, 416 Escondido Mall, Stanford, CA 94305-3086, USA
e-mail: adam@dschool.stanford.edu

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_13

information back to students so they can improve their work. This chapter outlines a flexible procedure, *Reflective Design Practice* (RDP) that allows instructors to assess authentic design work and provide feedback to students learning design thinking. Although there are myriad forms of creative assessment and feedback, very few of them accommodate a progression of creative work over the duration of a course. This is important because learning design thinking is more than developing a set of skills; it includes absorbing a new, more creative approach to problem-solving (Royalty et al., 2012). We believe that repeated practice of design instills creative problem-solving. This unfolds as a progression—one that needs to be captured as completely as possible. RDP was built to dynamically assess a sequence of authentic design work and provide both formative and summative feedback to students.

Assessing creative work has a long history. Many of the assessment goals can be sorted into two broad categories: creative potential and creative achievement. Creative potential measures highlight constructs like personality traits, cognitive abilities, and even neurological structures that people who are considered highly creative tend to excel at. The Torrance Test (Torrance, 1972) and Guilford's Alternative Uses Task (Guilford, 1967) are among the most common divergent thinking assessments. The NEO-Five Factor Inventory (McCrae & Costa, 1997), the Adjective Check List (Domino, 1994), and the Creative Personality Scale (Kaufman & Baer, 2004) are examples of assessing individual personality traits associated with creativity. It is important to note that none of the items assess produced creative work.

Creative achievement measures, by contrast, assess outputs. The most widely used approaches are the Creative Achievement Questionnaire (CAQ; Carson et al., 2005) and the Consensual Assessment Technique (CAT; Amabile, 1982). The CAQ assesses an individual's lifetime achievement across a number of domains. It is a way to identify both people who have contributed to a field and those who have profoundly shaped a field. Meanwhile, the CAT is a technique for evaluating a single creative product. Essentially, a group of experts evaluate the merits of an individual creative output. Although both of these measures can be used to evaluate creative design output, neither is calibrated to assess design work produced throughout an entire academic course.

While design thinking assessments do not have as rich a history as creative assessments, plenty exist. Saggar et al. developed a modified Pictionary task that allowed them to look at neural correlates of design thinking (Saggar et al., 2015). The Design-based Creative Agency Scale measures 11 key design thinking dispositions. There has also been initial work on assessing students' ability to follow a design thinking process (Hawthorne et al., 2016), as well "mindshifts" students identify while learning design thinking (Goldman et al., 2012). However, as with the creativity assessments listed above, none of them measure the progression of student work over the duration of a course.

Feedback is central to design thinking. Prototyping, a core component of a design thinking process relies on feedback (Chin et al., 2019; Cravens, 2014; Dow et al., 2010; Marks & Chase, 2019). Although instructors provide student feedback in various ways (Dym et al., 2005), much of this happens in real-time where instructors observe a team's work process and guide them through modeling or direct instruction.

This feedback focuses primarily on the "style" of working. In other words, instructors want to see students practicing creative behaviors that lead to strong design work (Royalty et al., 2015). Students also receive feedback on submitted assignments. It is still the case that nearly all forms of feedback respond to single working sessions or individual assignments. Intentional feedback on a student's course-long progression is rarely provided. This is in large part because it is difficult to facilitate.

We use reflection as the mechanism to bridge assessment and feedback while capturing the progression of student work throughout a course. Donald Schön (Schön, 1987) laid out a conceptual framework highlighting *reflection-in-action* (applying learning from previously considered reflections to a problem as one is working) and *reflection-on-action* (reflecting on a situation after the event has passed). Additionally, reflection can help learners identify underlying concepts, which can lead to increased transfer (Brown, 1978; Flavell, 1979). RDP was initially developed as a way to help students reflect on their individual design growth (Royalty et al., 2018). Understanding how their own personal design practice evolved within a broader context prepares students to better transfer design thinking into novel environments (Greeno et al., 1993). What follows is an outline of how to implement RDP, what to do with the data it generates, and how to modify it to suit different learning goals.

RDP lasts the duration of a design thinking course or program and has four components: a pre-assessment, weekly reflections, an interview, and a post-assessment (Fig. 1). The pre and post-assessments allow instructors to detect student growth over time. The weekly reflections generate a sequence of examples where students couple output with their metacognitive reflection. Finally, the interview serves as a time for a student and instructor to explore patterns and themes across the sequence of weekly reflections. These components work together to create an overall assessment of student work and to provide a feedback mechanism. There is significant space to modify the activities and prompts within each component—which we will cover in our modifications section—however, any RDP process should adhere to the following design principles:

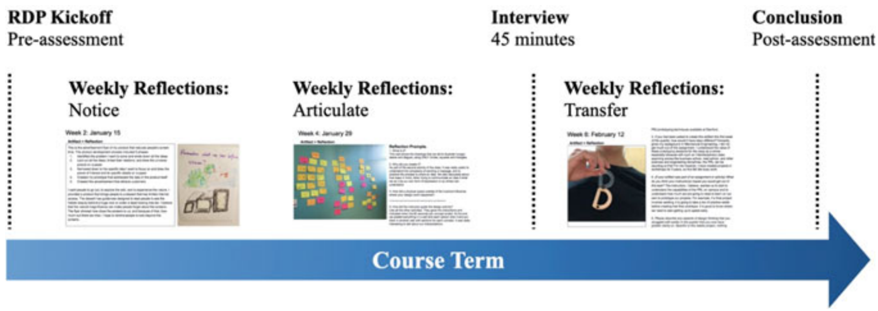


Fig. 1 The RDP timeline

Reflections are grounded in authentic, tangible work—students are asked to capture and comment on specific work they've completed, be it physical, digital, or experiential.

Prompts provoke metacognitive reflection and draw connections with the external environment—students should focus on their personal development while paying attention to the influence of the physical space and people they interact with.

Reflections capture a progression of work throughout the course or program—all components create a narrative of how a student's personal design practice evolves during the entirety of their learning experience.

Typically, RDP constitutes Schön's reflection on action in that it occurs after the work is finished. Additionally, RDP does not conflict with most existing reflection exercises. Its goal is to help students process a larger progression, and it does not necessarily help students internalize specific design tools or processes. RDP can complement more task-specific reflection exercises.

2 Reflective Design Practice Example

RDP was developed as a part of a multi-year research study at the Hasso Plattner Institute of Design at Stanford University (a.k.a. Stanford d.school). The goal was to understand how academic environments influence design thinking (Royalty et al., 2018, 2019, 2020). The format was a 1-unit independent study that any student enrolled in a Stanford d.school course could sign up for. The independent study ran in parallel to the design course a student took. We chose to offer RDP in parallel to a course because we wanted to draw students from multiple d.school courses. However, an identical RDP was successfully embedded as part of a semester-long design course through the Columbia Entrepreneurship Design Studio at Columbia University. Although the timelines in these two cases were the same, the major difference was who was providing feedback through RDP. In the case of the Columbia Design Studio, feedback was given by the design course instructor. While in the Stanford d.school case, feedback came from an instructor who was teaching the 1-unit independent study, not the instructor of the design courses where the students were enrolled. However, in both cases, the RDP administrator had over 10 years of design thinking experience teaching.

2.1 RDP Kickoff and Pre-assessments

The initial meeting lasted 30 min. We gave students a series of self-assessment items including the Design-based Creative Agency Scale (Royalty et al., 2014), a subset of the Innovation Self-efficacy Scale (Chen et al., 2018), and a Creative

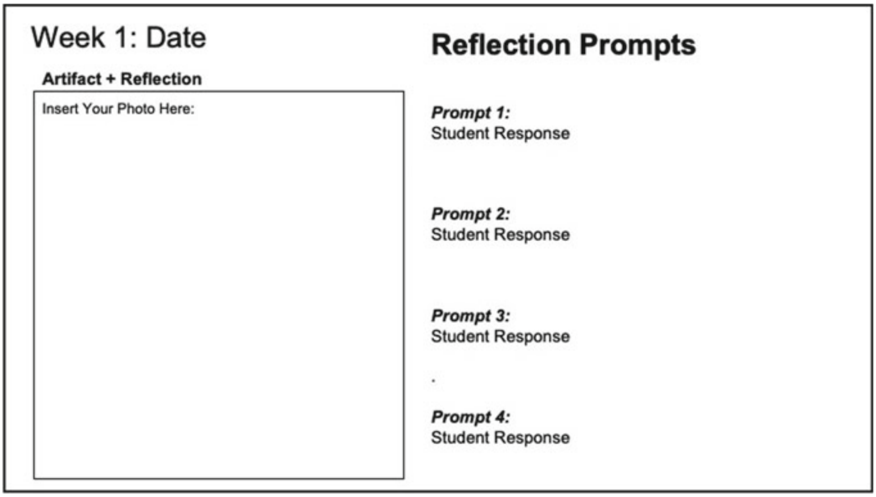


Fig. 2 The RDP weekly reflection template

Growth Mindset Scale (Royalty & Roth, 2016) modified from Carol Dweck's Original Growth Mindset Scale (Dweck, 2000). We chose these items because they measure student attitudes and beliefs toward design, creativity, and innovation. In the modifications section, we cover some alternative measures.

After the assessments, we shared the RDP timeline and introduced students to the weekly reflection tool by walking them through the template (Fig. 2). Finally, we responded to any clarification questions the students had.

2.2 Weekly Reflections

The weekly reflection is the primary RDP tool. We ask students to take a photo of one or more artifacts that represent something they created during their design course that week. During the kickoff we emphasize that an artifact can include something physically constructed, a drawing, a digital creation, or even something more abstract like an interview protocol. The point is to have students identify a piece of work they made for their course. For each image they capture, we asked students to respond to three to five questions related to the artifact (e.g., *What did you make? What aspects of the physical environment influenced the creation of this artifact?).* The image and corresponding reflections are added to a single slide of an online slide show tool that only the student and the instructor can view. We pre-populate the slides so each student has a single slide template for each week of the term. We chose this format because it is easy for students to add images and text. Likewise, instructors have the option of modifying the prompts along the way. In addition, students and instructors can quickly navigate between different weeks.

Students complete the weekly reflection outside of class. Most of the students report taking around 10 min to fill out the template. We sent a reminder email each Friday. Occasionally, students inserted their own slide so that they could do two reflections in a given week.

Table 1 highlights some of the prompts we used throughout the duration of the course. In this example, the weekly reflection prompts fell into three themes that progressed during the term: *notice*, *articulate*, and *transfer*. *Notice* prompts focused on helping the students be aware of the learning environment and their initial reactions to design thinking. *Articulate* prompts encouraged students to think more deeply about how the physical space and people they interact with impact their ability to learn design. Finally, *transfer* prompts helped students reflect on ways in which they may apply design thinking outside of their courses. Early in the term, we included prompts exclusively from the *notice* theme. As their course progressed, we added more *articulate* prompts. We did not ask *transfer* prompts until the last few weeks of the course.

Table 1 RDP weekly reflection prompts

Theme	Prompt
Notice	What did you make? Why did you make it?
Notice	Think about the location where you made the artifact. What do you notice about the physical environment that may have contributed to the creation of this artifact?
Notice	What three aspects of the context around you most influenced the creation of this artifact? These could include the environment, team, instructors, etc.?
Notice	Reflect on the types of behaviors you noticed in yourself while creating this artifact. Describe 1 or 2 behaviors that were comfortable/familiar and 1 or 2 behaviors that were uncomfortable/unfamiliar
Articulate	List 2 or 3 ways in which the work in your d.school course is different from the work you do in one of your other courses or in your program
Articulate	If this artifact was made as part of a team, describe any key aspects of the collaboration that contributed to the creation of this artifact
Articulate	Reflect on the types of behaviors you noticed in yourself while creating this artifact. Describe 1 or 2 behaviors that were comfortable/familiar and 1 or 2 behaviors that were uncomfortable/unfamiliar
Articulate	If your artifact was part of an assignment or activity, what do you think your instructor(s) hoped you would get out of this task?
Articulate	If you had been asked to create this artifact in the first week of the quarter, how would it have been different?
Transfer	Capture something you made using design thinking that was not part of your coursework. How did the location where you made it support and/or not support creative behaviors? Why do you think this was the case?
Transfer	Capture something you made using design thinking that was not part of your coursework. What creative behavior did you identify as the most uncomfortable for you personally?
Transfer	How would you design a way to continue to engage in a behavior that you find uncomfortable?

The data the weekly reflections yield are short answers. The length can vary. Sometimes the prompts don't trigger a deep reflection and the students write a few short sentences. Other times, the prompts elicit multiple paragraph responses.

By connecting the prompts with an artifact, students tend to link their metacognition with something observable. We saw a range of sentiments including excitement, confusion, as well as areas of struggle and growth. The weekly reflection tool grounds these sentiments in the process of using design to create an output. This helps both the student and instructor identify how specific exercises or activities impact student perspectives on design. This can drive both assessment and feedback. Here are some example responses from the weekly reflections:

Think about the location where you made the artifact. What do you notice about the physical environment that may have contributed to the creation of this artifact?

The openness of the room allowed my teammate and I to roll a high table away from the others and to feel like we had our own private conversation. The presence of movable whiteboards in the room created a sense of protective walls there, in case some additional privacy was needed, even though the space was still open.

List 2 or 3 ways in which the work in your d.school course is different from the work you do in one of your other courses or in your program.

Designing directly for another user is fabulous, and very different from writing a paper from a mythical CEO to be graded by a professor. Limited time windows in class to create an assignment or activity drove us to come to conclusions without endless analysis or late-night extra-tired research.

Reflect on the types of behaviors you noticed in yourself while creating this artifact. Describe 1 or 2 behaviors that were comfortable/familiar and 1 or 2 behaviors that were uncomfortable/unfamiliar.

I found it particularly challenging to create a stakeholder map without knowing what the problem statement was at the start and without having spoken to the partner organization that we were assigned to work with for the project. It felt like groping around in the dark! The d.school instructors were adamant about not giving us information that looked ahead as part of the exercise. Coming from the business school, I'm starting to get cognitive dissonance at the completely different schools of thought in approaching problems. :)

The first example reveals students intentionally modifying their learning environment to create a sense of privacy. The manipulation of physical space is a key part of learning design thinking (Doorley & Witthoft, 2012; Royalty, 2018). Instructors rarely observe students interacting with the learning environment beyond class. This is a unique opportunity for assessment as students typically have more agency to choose how they work outside the in-class exercises scaffolded by their instructors. Moreover, by linking actions with thoughts, we see the student's intent—creating space for a private interaction. Through this weekly reflection, the instructor has evidence of how and why the student manipulates the physical space.

The second example indicates that the student is engaging with two core tenants of design thinking: being user-driven and having a bias toward action. The student also reinforces these ideas by comparing them to working styles in other courses.

In the third example, the student appears to be struggling with the open-endedness of the exercise. This gives the instructor an opportunity to provide feedback to the student. It also presents valuable feedback for the instructor. Discomfort is often part of learning design thinking (Royalty et al., 2015), however, now the instructor has evidence detailing the degree of discomfort a student is feeling. This gives the instructor the option to adjust their teaching practice to the students' needs as the course unfolds.

2.3 Interviews

We conducted 45-min semi-structured interviews with each student two to three weeks before the end of the course. Students completed weekly reflections focusing on *notice* and *articulate* but not the *transfer* themed prompts. However, the interview provided an opportunity to reflect across the majority of their design course. We wanted them to do this broader retrospective reflection before shifting to transfer.

Because we were particularly interested in understanding the role of the academic context on learning design thinking, our interview protocol had three main topics. The first was how the environment affected the creation of each artifact. Next, we investigated any contrasts students noticed between the artifacts they created in their design course compared to artifacts created in non-design courses. From there, we asked students to reflect on how the way of working in a design course might contrast from the way of working in non-design courses. Finally, we asked them which artifacts were more or less comfortable to create. This helped them think about which areas of design thinking they might naturally gravitate toward or avoid. At the end of each interview, we left time for students to share any general comments or ideas that came up during the course of the interview.

For each question, we asked the student to respond by drawing on specific examples from their series of reflections. Specifically, students looked through their online slides and identified one or more weekly reflection—often expanding on it—to answer the question. The interviewer asked follow-up questions based on the student's response.

The interviews were recorded with the student's consent and we shared the recording of the interview with the student. Table 2 contains the protocol we used.

2.4 Concluding Meeting and Post-assessment

At the end of the term, we met with students for an hour-long session. Students completed a post-assessment for the three scales we gave them at the beginning of

Table 2 RDP interview questions

Topic	Question
Environment	Look back at the artifacts you have created so far. For one of them, tell me more about the aspects of the environment that supported the creation of this artifact
Environment	Were there any aspects of the environment that were barriers to the creation of this artifact? Why did they act as barriers?
Contrast	Find an artifact that you believe represents the largest contrast between the types of things you create in a d.school course and the types of things you create in your non-d.school courses. How is this artifact different from what you might normally create in another (non-d.school) course?
Contrast	The next few questions are about the ***style*** of work you engage in at the d.school and in your other courses. When we say style, we mean *the way* in which you have worked. This can include how you worked with others, how you made artifacts, how you thought about a challenge, how you learned, etc. How did the d.school style of working enable (or not) the creation of one of your artifacts? (pick one)
Contrast	How does the d.school style of working differ from the style of working in non-d.school courses?
Contrast	What is the relative value of each type of course (d.school and non-d.school)?
Comfort	Which artifact was created using a style or way of working that felt comfortable or familiar to you? Why?
Comfort	Which artifact was created using a style or way of working that felt foreign or uncomfortable? Why?

the term. We asked students to share any personal insights they had while completing the *transfer*-themed weekly reflections. Unlike the interviews, the reflections in the concluding meeting were not structured by topic and happened as part of a class-wide discussion. The final exercise in this RDP example was for students to identify aspects of an environment (inside and outside of the academic context) that support their personal design practice based on all their reflections. Students typically identified things like access to creative materials, open-ended challenges, teammates who build on each other's ideas, etc. The goal of this last task was to prepare students to assess an environment they may find themselves in to determine how well it could support design thinking.

3 Limitations

RDP can be an effective means for assessing students and providing feedback. However, there are some key limitations. One is the wide range of data from the weekly reflections. Students essentially reflect on an artifact of their choosing. Although the choice of artifact can prove insightful, each entry may or may not relate directly to the previous or subsequent week. This means that the overall narrative formed by all the reflections might have some disconnects. Another limitation is that the interviews are time intensive. Finding time to interview each student for 45 min might not be feasible.

There is an additional complication for researchers using RDP to collect data. Going through RDP impacts student learning. For example, if you use RDP to measure how much students learn through a design thinking course, the extra reflection done to complete RDP is going to affect student learning. As noted above, when we used RDP in a research study, we did not provide feedback to students until the end of the course—after the final post-assessment. This was done to minimize our impact as researchers. If RDP is purely a pedagogical tool, then instructors can give feedback at any time.

4 Modifications to RDP

RDP is quite flexible. Instructors and researchers seeking to modify RDP can start by articulating the core aspects of the learning experience they wish to evaluate. Then, using the design principles listed above, modify the materials outlined in this chapter. This will lead to a different set of pre/post questions, weekly reflection themes, and interview topics. Here are a few examples of changes one could make.

Increase the number of reflections. Having multiple reflections per week increases the amount of data. The cost is student time and the increased chances that students don't complete the reflections. The benefit is more rich data to draw on.

Build a more coherent overall narrative by scoping the topic of each weekly reflection. We allowed students to choose the artifact to photo. However, it would be possible to focus reflections on a specific activity like collaborative visualizations or a specific tool like journey mapping.

Ask about specific design-related topics during the interview. Navigating ambiguity or rapid experimentation are topics that students would easily be able to speak to by analyzing their weekly reflections.

Have students submit a video response to interviews questions. Instructors may not be able to interview every student, particularly in large courses. Students could submit video (or audio) responses to a list of interview questions. Alternatively, students could interview each other and submit the responses to instructors. In either case, instructors could choose particular segments of each student's response to review.

Create a group assessment by having individual teams use the RDP tools together. RDP provides space for individuals to notice their own design growth. However, teams could leverage this framework as well.

5 Conclusion

RDP collects both qualitative and quantitative data on students' progression through a design thinking course. It allows for pre and post measurement while connecting reflections to authentic work. The weekly reflections and interviews provide rich data instructors can use to assess student progress and give both formative and summative feedback. The format is extremely flexible and can be administered as part of a course or as a method for data collection. Our hope is that both instructors and researchers use RDP to advance the teaching and learning of design thinking.

References

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43*(5), 997.

Bott, N., Quintin, E. M., Saggar, M., Kienitz, E., Royalty, A., Hong, D. W. C., Liu, N., Chien, Y. H., Hawthorne, G., & Reiss, A. L. (2014). Creativity training enhances goal-directed attention and information processing. *Thinking Skills and Creativity, 13*, 120–128.

Bransford, J. D., & Schwartz, D. L. (1999). Chapter 3: Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education, 24*(1), 61–100.

Brown, A. L. (1978). Knowing when, where, and how to remember; a problem of metacognition. *Advances in Instructional Psychology, 1.*

Carroll, M., Goldman, S., Britos, L., Koh, J., Royalty, A., & Hornstein, M. (2010). Destination, imagination and the fires within: Design thinking in a middle school classroom. *International Journal of Art & Design Education, 29*(1), 37–53.

Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativity Research Journal, 17*(1), 37–50.

Chen, H. L., Powers, K., Prasad, K. V., Anderson, M., Royalty, A., Gilmartin, S. K., & Sheppard, S. D. (2018, October). A mixed methods approach to understanding how colleges, universities, and employers prepare and support undergraduates in engineering internships. In *2018 IEEE Frontiers in Education Conference (FIE)* (pp. 1–5). IEEE.

Chin, D. B., Blair, K. P., Wolf, R. C., Conlin, L. D., Cutumisu, M., Pfaffman, J., & Schwartz, D. L. (2019). Educating and measuring choice: A test of the transfer of design thinking in problem solving and learning. *Journal of the Learning Sciences, 28*(3), 337–380.

Cravens, A. E., Ulibarri, N., Cornelius, M., Royalty, A., & Nabergoj, A. S. (2014). Reflecting, iterating, and tolerating ambiguity: Highlighting the creative process of scientific and scholarly research for doctoral education. *International Journal of Doctoral Studies, 9*, 229–247.

Cross, N. (2006). Design as a discipline. *Designerly Ways of Knowing*, 95–103.

Domino, G. (1994). Assessment of creativity with the ACL: An empirical comparison of four scales. *Creativity Research Journal, 7*(1), 21–33.

Doorley, S., & Witthoft, S. (2012). *Make space: How to set the stage for creative collaboration.* Wiley

Dow, S. P., Glassco, A., Kass, J., Schwarz, M., Schwartz, D. L., & Klemmer, S. R. (2010). Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Transactions on Computer-Human Interaction (TOCHI), 17*(4), 1–24.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Psychology Press.

Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education, 94*(1), 103–120.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906.

Goldman, S., Carroll, M. P., Kabayadondo, Z., Cavagnaro, L. B., Royalty, A. W., Roth, B., Roth, B., Kwek, S.H., & Kim, J. (2012). Assessing d. learning: Capturing the journey of becoming a design thinker. In *Design thinking research* (pp. 13–33). Springer, Berlin, Heidelberg.

Goldman, S., & Kabayadondo, Z. (Eds.). (2016). *Taking design thinking to school: How the technology of design can transform teachers, learners, and classrooms*. Taylor & Francis.

Greeno, J. G., Moore, J. L., & Smith, D. R. (1993). Transfer of situated learning.

Guilford, J. P. (1967). The nature of human intelligence.

Hawthorne, G., Saggar, M., Quintin, E. M., Bott, N., Keinitz, E., Liu, N., Chien, Y.H., Hong, D., Royalty, A., & Reiss, A. L. (2016). Designing a creativity assessment tool for the twenty-first century: Preliminary results and insights from developing a design-thinking based assessment of creative capacity. In *Design thinking research* (pp. 111–123). Springer.

Kaufman, J. C., & Baer, J. (2004). Sure, I'm creative—but not in mathematics!: Self-reported creativity in diverse domains. *Empirical Studies of the Arts, 22*(2), 143–155.

Kelley, T., & Kelley, D. (2013). Creative confidence: Unleashing the creative potential within us all. Currency.

Kienitz, E., Quintin, E. M., Saggar, M., Bott, N. T., Royalty, A., Hong, D. W. C., Liu, N., Chien, Y. H., Hawthorne, G., & Reiss, A. L. (2014). Targeted intervention to increase creative capacity and performance: A randomized controlled pilot study. *Thinking Skills and Creativity, 13*, 57–66.

Marks, J., & Chase, C. C. (2019). Impact of a prototyping intervention on middle school students' iterative practices and reactions to failure. *Journal of Engineering Education, 108*(4), 547–573.

McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist, 52*(5), 509.

Royalty, A. (2018). Design-based pedagogy: Investigating an emerging approach to teaching design to non-designers. *Mechanism and Machine Theory, 125*, 137–145.

Royalty, A., & Roth, B. (2016). Developing design thinking metrics as a driver of creative innovation. In *Design thinking research* (pp. 171–183). Springer.

Royalty, A., Oishi, L., & Roth, B. (2012). "I use it every day": Pathways to adaptive innovation after graduate study in design thinking. In *Design thinking research* (pp. 95–105). Springer, Berlin, Heidelberg.

Royalty, A., Oishi, L. N., & Roth, B. (2014). Acting with creative confidence: developing a creative agency assessment tool. In *Design thinking research* (pp. 79–96). Springer.

Royalty, A., Ladenheim, K., & Roth, B. (2015). Assessing the development of design thinking: from training to organizational application. In *Design thinking research* (pp. 73–86). Springer.

Royalty, A., Chen, H.L., & Sheppard, S. (2018). Reflective Design Practice: A Novel Assessment of the Impact of Design-based Courses on Students. In *Proceedings of the ASEE/IEEE Frontiers in Education Conference*. San Jose, CA.

Royalty, A., Chen, H., Roth, B., & Sheppard, S. (2019). Measuring design thinking practice in context. In *Design thinking research* (pp. 61–73). Springer.

Royalty, A., Chen, H., Roth, B., & Sheppard, S. (2020). Reflective tools for capturing and improving design driven creative practice in educational environments. In *Design thinking research* (pp. 49–65). Springer.

Saggar, M., Quintin, E. M., Kienitz, E., Bott, N. T., Sun, Z., Hong, W. C., et al. (2015). Pictionary-based fMRI paradigm to study the neural correlates of spontaneous improvisation and figural creativity. *Scientific Reports, 5*, 10894.

Saggar, M., Quintin, E. M., Bott, N. T., Kienitz, E., Chien, Y. H., Hong, D. W., Liu, Y., Hawthorne, G., & Reiss, A. L. (2017). Changes in brain activation associated with spontaneous improvization and figural creativity after design-thinking-based training: A longitudinal fMRI study. *Cerebral Cortex, 27*(7), 3542–3552.

Schön, D. A. (1987). Educating the reflective practitioner.

Torrance, E. P. (1972). Predictive validity of the torrance tests of creative thinking. *The Journal of creative behavior*.

von Thienen, J., Royalty, A., & Meinel, C. (2017). Design thinking in higher education: How students become dedicated creative problem solvers. In *Handbook of research on creative problem-solving skill development in higher education* (pp. 306–328). IGI Global.

Measurement in Design Thinking: How to Improve Different Areas and Fields by Applying Design Thinking

Assessing the Impact of Design Thinking in Organizations: Foundations of a Framework

Selina Mayer, Martin Schwemmle, Claudia Nicolai, and Ulrich Weinberg

Abstract Understanding the impact of Design Thinking in organizations is fundamental for fostering its dissemination and development. This chapter thus deals with the assessment of Design Thinking impact in organizations. Providing an overview of the concept of "impact" and reviewing existing models of impact measurement, we derive important criteria for assessing the impact of DT. Moreover, the chapter presents an overview of existing research in the fields of DT and design while showing the lack of a proper framework currently in place. Synthesizing these streams, it proposes foundations for the impact assessment of DT in organizations.

1 Introduction

Organizations increasingly rely on Design Thinking (DT) to develop innovative offerings or drive transformation (Cooper et al., 2009; Micheli et al., 2019). Despite this growth of DT in both research and practice and while there are numerous success stories of Design Thinking around the globe and enthusiastic reports by workshop or sprint participants, assessing the impact of Design Thinking still is a challenge (Liedtka et al., 2017). Accordingly, practitioners report the usage of DT in various areas, such as organization design, employee engagement, learning, and other topics that are relevant to HR (Bersin et al., 2016). In the academic community, researchers have investigated the relationship of DT and organizational culture (Elsbach &

S. Mayer (✉) · M. Schwemmle · C. Nicolai · U. Weinberg
HPI School of Design Thinking, Hasso Plattner Institute for Digital Engineering, Campus Griebnitzsee, August-Bebel-Str. 88, 14482 Potsdam, Germany
e-mail: Selina.Mayer@hpi.de

M. Schwemmle
e-mail: martin.schwemmle@hpi.de

C. Nicolai
e-mail: claudia.nicolai@hpi.de

U. Weinberg
e-mail: uli.weinberg@hpi.de

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_14

Stigliani, 2018) or learning (Beckman & Barry, 2007), mirroring the raised interest in broadening the usage and impact of DT for an organization and its individuals. For instance, managers following the performance measurement paradigm “You can only manage what you can measure,” ask for DT KPIs (Key Performance Indicators). And, while DT teams need a certain freedom and autonomy, some mention a lack of clear goals for their projects as a reason holding them back, since they have no clear direction to follow.

Up-to-date scholars and practitioners discuss the added value of DT quite controversially. Johansson-Sköldberg, Woodilla, and Çetinkaya even sum it up as being “easy for the temporarily intensive discourse to be dismissed as hype or a fad.” (2013, p.121). With DT “coming of age” (Kolko, 2015), this argument might be refuted; however, the impact of DT on organizational performance still lacks empirical evidence (Micheli et al., 2019).

The reasons for this lack are manifold. They include arguments such as that DT is still rather young, why there might not exist enough data, that it still lacks a clear definition, making it hard to assess its influence, and that it gets its power through the difference from the usual organizational thinking, which is why, for instance, KPIs seem an inadequate way of approaching DT. At the same time, DT is understood as a problem-solving approach, which is why the requirement of a proof that it actually solves problems or solves them in an even better way than other approaches seems more than valid. In essence, the ongoing discussion about measuring DT impact, what impact means, and how it can be carried out, calls for a deeper understanding.

Taking a first step in addressing these questions, this chapter first provides an overview of the term and concept of “impact,” then demonstrates existing performance measurement approaches in the fields of management and creativity, and finally derives implications for measuring and evaluating the impact of Design Thinking. In a second step, this study gives an overview of the current literature dealing with the assessment of DT and Design in organizations. In concluding with a discussion and outlook, important foundations that need to underly a proper assessment of DT impact are presented. The provision of a general understanding of impact and a summary of existing and related research provide practitioners with a better understanding and overview and as well as offering them helpful arguments for discussing and developing the assessment of DT impact in their organizations. In a similar way, researchers are provided important foundational thoughts as the basis for further research.

2 Understanding the Impact of Design Thinking and the Implications for Measuring It

2.1 Impact and the Logic Model

In order to define impact, we reviewed literature in the field of performance measurement as well as in program evaluation. Performance measurement (and management) is described as “a collection of management processes supported and enacted through the use of tools and techniques such as scorecards, measures, targets, performance reviews and incentives that are developed centrally and cascaded throughout the organization” (Bourne et al., 2018, p. 2788). Two prominent approaches for performance measurement, the Balanced Scorecard and the EFQM model, are described in Sect. 2.3.1.

Program evaluation is an academic research field investigating how scholarly activities create impact for policies and practice (Pettigrew, 2011). Within this field, various approaches have been identified in order to evaluate certain activities or programs in a structured manner; the one that is used most often is the logic model (Kneale et al., 2015). “As understood in the program evaluation literature, logic models are one way of representing the underlying processes by which an intervention effects a change on individuals, communities or organisations.” (Kneale et al., 2015, p. 26).

Within the field of program theory and evaluation, the logic model describes how short- and long-term outcomes link with the adopted activities and initial goals (McLaughlin & Jordan, 1999). Logic models are adopted in various areas, such as allowing the evaluation projects sponsored by grants of foundations or the government (Foundation. Kellogg, 2004) or in the healthcare sector, for example, concerning programs aimed at preventing chronic diseases (Tucker et al., 2006). As the implementation of a new working mode or innovation process, such as Design Thinking, can be seen as a program, we describe in detail what the elements of program theory indicate for a program evaluation. Based on the different available variations (e.g., Julian et al., 1995; McLaughlin & Jordan, 1999), we identified six core components relevant to answering the question of effectiveness of programs and the assessment of impact, which are displayed in Fig. 1.

1. First, to explain the logic behind a program, the *Problem/Situation* needs to be defined, clarifying the current context and what problem is targeted by the initiative based on the goal or strategic intent of the organization.
2. The *Input* refers to the question of what resources are available to the program.
3. The logic model further states what *Activities* are planned and conducted in the program.
4. The *Output* describes specific results that the activities produce, often referring to short-term results.
5. Focusing more on long-term implications, the *Outcome* describes changes or benefits that result from the program.

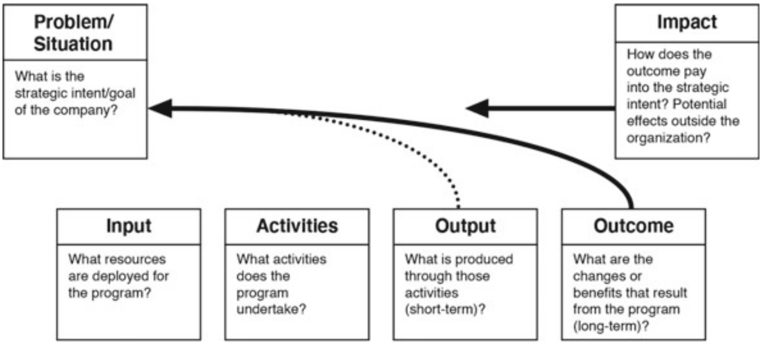


Fig. 1 Logic model and its core elements to understand impact

6. Last, the *Impact* refers to the question of how the outcomes relate back to the Problem/Situation, indicating also potential effects outside of the organization.

An example might help to make this a little bit clearer. Imagine a non-profit organization with the overall goal of saving animals and reducing species extinction (e.g., Problem/Situation). To address this problem, the organization launches a program, for which they provide money, staff, and equipment (e.g., Input). The program focusses on buying land, so the animals can keep their natural habitat (Activity). The amount of square meters of obtained land provides a measure for the output, whereas a rise in public awareness and therefore the acquired additional funding is a long-term outcome. So far the non-profit organization seems to be quite successful, looking at the achieved outputs and outcomes. But the reported measures do not indicate a reduction of the extinction rate of species, therefore indicating no impact. This example demonstrates that it is important to have a clear understanding of what we are talking about when we want to investigate the impact of Design Thinking.

The consideration of these different elements has broader implications when it comes to the differentiation of output, outcomes, and impact. In particular, it reveals an important differentiation between performance measures and performance evaluation (Harris Mulvaney et al. 2006; Brousselle and Champagne 2011). Measures can be applied before, during and after an activity, whereas the evaluation looks at the long-term outcomes compared to the beginning strategic intent, as can be seen in Fig. 1.

While one can apply measures on any of the six elements, impact is not measured, but assessed or evaluated requiring a relationship of the Problem/Situation and the outcome. We therefore understand impact as an evaluation of the question how much the output and outcomes of a certain activity help to fulfill a previously stated strategic intent or solve a problem. Hence, following this understanding and the logic model approach, to understand the impact of Design Thinking endeavors their strategic intent and goals need to be defined to relate the output and outcomes, respectively.

As an example, organizations often start the introduction of DT by freeing the capacity of 4–6 employees (Input), allowing them to be a team and participate in a DT

Workshop (Activity). In this workshop, the team creates a prototype (Output), which, in the ideal situation is further iterated and results in a marketable product, generating a certain amount of revenue (Outcome). But what does this tell us about impact? Not much, if it is not related to the original intent (Problem/Situation). For instance, if the intent was to build innovation capacity in the organization, this outcome of a marketable product might not allow drawing conclusions about capacities. The term impact therefore needs to be considered carefully here, because having created certain outputs and outcomes (e.g., marketable product) might not result in a desired impact (e.g., capacity building). On the other hand, during the innovation project work, the team could have built up various new capabilities, but as long as measures are not in place capturing outputs and outcomes for this strategic goal, the actual impact might get overlooked. Overall, evaluating the impact of DT always means identifying the direct and indirect outputs and outcomes and relating them to the strategic intent. Reporting output and outcome measures alone can be informative, yet they alone do not tell us about the impact.

2.2 Criteria of Frameworks—Model or Theory?

In addition to understand the term impact itself in more detail, we set out to investigate what frameworks already exist to model the impact of DT. In order to do this, we first need a deeper understanding of what we are actually looking for in distinguishing broadly used terms such as model or theory. We identified five essential criteria to constitute theory or models from literature (Whetten, 1989; Wiesche & Yetton, 2017) as displayed in Fig. 2. A model consists of the elements: (1) descriptions of empirical observations, (2) definitions of abstract variables, and (3) definitions of the relationships between these variables. A theory offers in addition to these three elements: (4) explanations for the existence of the relationships between variables as well as (5) boundary conditions for those. Often, exploratory studies focus on rich descriptions

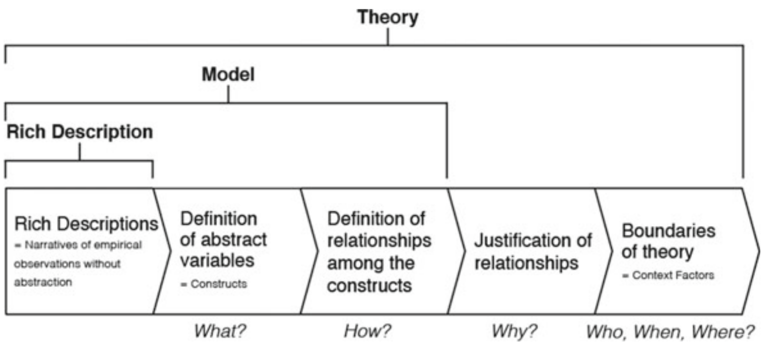


Fig. 2 Elements of a theory based on Wiesche and Yetton (2017)

of empirical observations, which are a valuable contribution. Nevertheless, these rich descriptions alone are not enough to qualify as a model or theory.

2.3 Existing Approaches of Measurement

Since measurement is a key necessity for organizations and their managers, academic and practitioner literature provide a variety of approaches for impact and success measurement for an organization. In the following, we introduce three major measurement approaches that all take a broader perspective, go beyond single KPIs, such as sales or profit, and seem particularly relevant for deriving implications to assess the impact of Design Thinking. These approaches are the Balanced Scorecard, the EFQM model, and a Theory of Organizational Creativity. These approaches have been selected because they already advance the notion of performance measurement to performance management.

2.3.1 The Balanced Scorecard (BSC)

In 1996, Robert Kaplan and David Norton published their famous book *The Balanced Scorecard*. At this time, organizations were quite focused on financial (and related performance) measures; however, as the authors write in their introduction, similar to a pilot who not only uses air speed, but also needs to monitor fuel gauge and altitude, they should consider new capabilities to adapt to the new operating environment of the information age (Kaplan & Norton, 1996). As a result of these considerations, they propose the Balanced Scorecard (BSC), a new framework that translated strategy into operations. The BSC puts the vision and strategy of the organization in the focus and then considers four perspectives answering different questions (Kaplan & Norton, 1996):

1. Learning and Growth: "To achieve our vision, how will we sustain our ability to change and improve?"
2. Internal Business Processes: "To satisfy our shareholders and customers, what business processes must we excel at?"
3. Customer: "To achieve our vision, how should we appear to our customers?"
4. Financial: "To succeed financially, how should we appear to our shareholders?"

All four dimensions contain objectives, measures, targets, and initiatives to turn the vision and strategy into practice. Hence, each organization needs to derive objectives in order to reach their goal and assign them to the respective dimension. In light of the aim of this chapter, three characteristics of the BSC are striking. First, the BSC combines an internal with an external perspective: While the Learning and Growth and Internal Business Processes perspective ensure that the organization has the right people, trains them correctly and has processes in place to create and deliver value, the Customer and Financial Perspective reflect the connection of the organization to the market, i.e., to build relationships with customers, sell products, and thereby earn money. Second, these elements have an inherent logical form—bottom to top, as shown in Fig. 3, which is elaborated on as the concept of "Strategy Maps" (Kaplan & Norton, 2004).

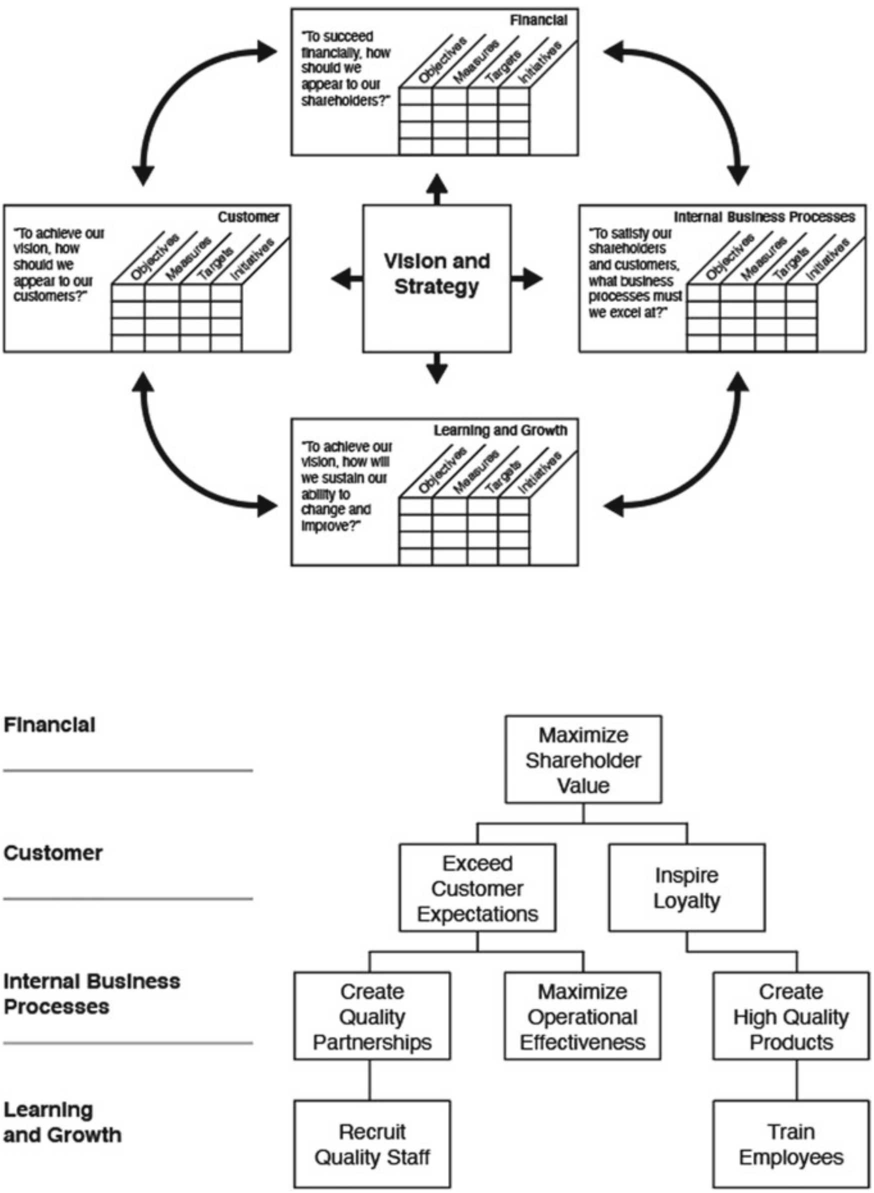


Fig. 3 Logic of the Balanced Scorecard and Strategy Map (Kaplan & Norton, 1996, pp. 9 and 31)

Within such a Strategy Map, not only are the four dimensions linked, but also the objectives that are part of the dimensions. Hence, the BSC not only considers a balanced view on the organization, but also considers causal relationships between these perspectives and their objectives. Third, the BSC links the spheres of Vision and Strategy and specific measures, i.e., a strategy with an operational level. Or, in terms of the logic model, outputs and outcomes with impact. As shown in the questions quoted above, each perspective is linked with the overall vision and strategy, underlining the aim to generate impact. To measure these dimensions, specific measures and initiatives are developed. These three characteristics of the BSC make it very suitable for assessing the impact of Design Thinking, since also Design Thinking influences the organization and its culture internally, while at the same time providing better products for customers and building relationships with users. Moreover, following the logic model, the BSC differentiates the levels of operation and strategy, which could mean that an organization measures the number of employees trained in DT or the number of workshops conducted, links that to customer satisfaction and, also checks if all these operational measures help to achieve the overall vision of becoming more sustainable. While research in the field of design has applied the BSC (Borja de Mozota, 2010), to the best of our knowledge, there have been no attempts to use the BSC for the assessment of DT impact.

2.3.2 The EFQM Excellence Model

The EFQM Excellence Model has been developed and is continually updated by the European Foundation for Quality Management. “Excellence” refers to performance, customers, people and society, and strategy. It is achieved through a number of enablers, including leadership, people, policy and strategy, partnerships and resources, and processes, as shown in Fig. 4 (EFQM 2003). Hence, like the BSC, the EFQM Excellence Model differentiates two spheres and takes a balanced and multidimensional approach comprising different success criteria. While all components of

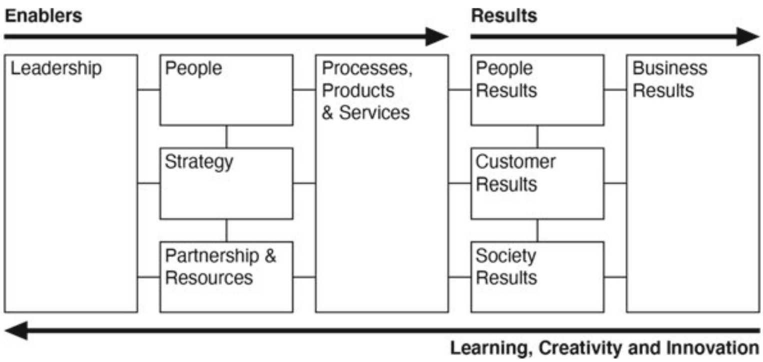


Fig. 4 Structure of the EFQM excellence model (EFQM 2003, p. 5)

the BSC are also reflected by this model, it in particular highlights the responsibility of leadership to provide the internal prerequisites of the organization and, moreover, not only considers customers, but also society as relevant for an organization's results. However, the clear orientation toward a vision and the continuous connection to strategy are much clearer in the BSC. Regarding measurement, the EFQM model recommends a self-assessment for the different areas, while the BSC defines clear criteria on the level of each perspective. Lastly, the BSC only provides a scaffolding for the organization's objectives and addresses why each BSC is unique, while the EFQM model has a more fixed structure with its specific elements.

In essence, the logic of the BSC seems more adequate as a general structure for the assessment of DT impact since it provides more flexibility to reflect the individual strategy of each organization. The inherent logic between the four perspectives is also a good indicator of the various levels within an organization, as introduced in the next section. However, this general structure of the BSC might be extended by, for instance, elements of leadership or societal impact, as suggested by EFQM.

2.3.3 Theory of Organizational Creativity

To acknowledge the innovative and creative power of Design Thinking, we want to complement the previously introduced management models with the introduction of a model dealing with creativity in organizations. This model was introduced by Woodman et al. (1993) to describe how creativity and innovation unfold in an organization from a systemic viewpoint and is shown in Fig. 5. It in particular highlights the three levels of individual, team, and organization, an aspect which was not in the focus of the two mentioned models, and does not contain any forms of measurement approaches. At the core of the model lies a so-called interactive understanding of creativity, understanding creativity as "the complex product of a person's behavior in a given situation" (Woodman et al., 1993, p. 294). This individual creativity is influenced by antecedents, such as past experiences, and through characteristics of the current situation. Moreover, the individual's cognitive abilities, personality, knowledge, and intrinsic motivation have an effect on creativity. This complex interaction not only happens on the individual level, but also on larger levels within the organization: Several individuals together form a group. However, group creativity is not simply the aggregate of all members' creative output, but rather takes into account the group composition, characteristics, and processes as further influencing factors. And several such groups then determine the creativity on the organizational level. Hence, "The gestalt of creative output (new products, services, ideas, procedures, and processes) for the entire system stems from the complex mosaic of individual, group, and organizational characteristics and behaviors occurring within the salient situational influences (both creativity constraining and enhancing) existing at each level of social organization." (Woodman et al., 1993, p. 296).

In addition to the previously mentioned models, this interactive and systemic perspective on creativity highlights the relevance of linking different levels within the organization and further considering the complex interactions inherent in these

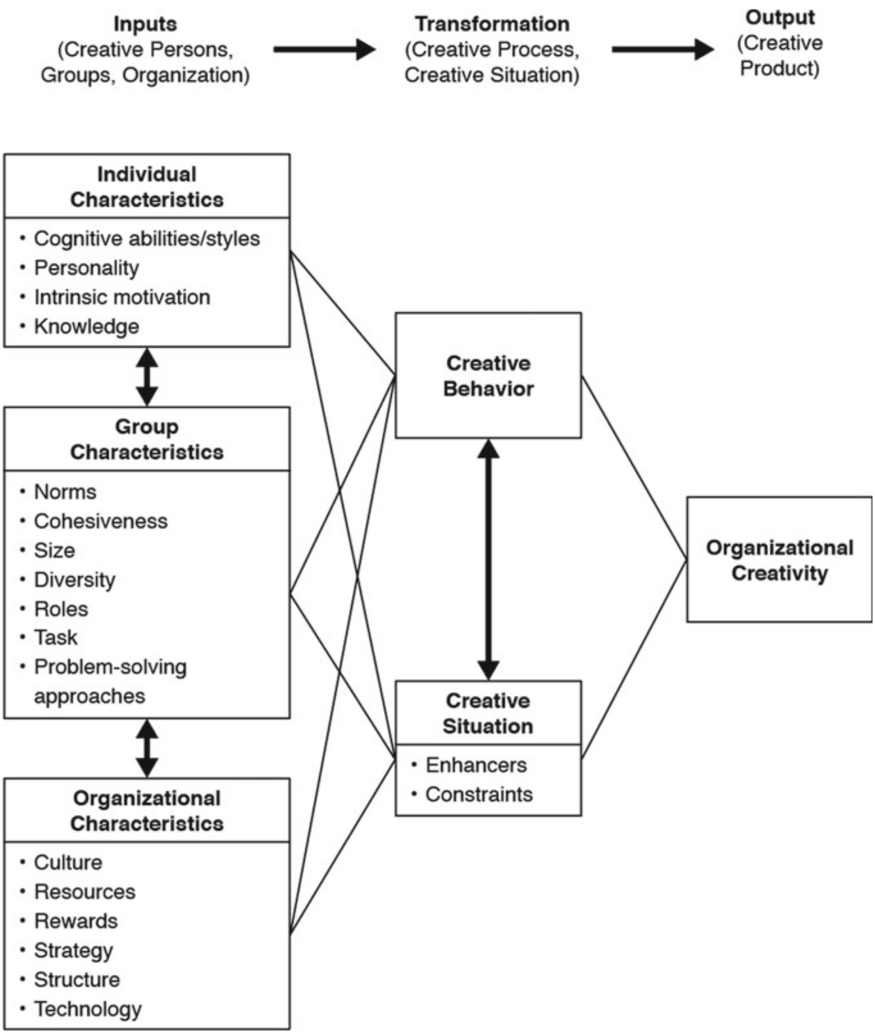


Fig. 5 A model of creativity in organizations (Woodman et al., 1993, p. 309)

interactions. As a consequence for the assessment of DT impact, different measures on different levels seem to be necessary since the successful application of DT on the individual level does not necessarily lead to unfolding effects on a team or organization level. Moreover, depending on the focus of the implementation of DT, it might make sense to adjust the focus on the impact measurement to fit the scope. Hence, if an organization only applies DT in one department, one cannot expect a broad influence on the organization level but must break down the overall strategies for a proper impact assessment.

3 Overview of the Current Assessment of DT and Design in Organizations

3.1 Overview of Existing Studies to Assess DT Impact

In order to get an overview of studies with frameworks describing DT impact, we screened the existing literature based on the previously described five criteria (Sect. 2.2) to evaluate the studies for offering models or theories. The five criteria are (1) descriptions of empirical observations, (2) definitions of abstract variables, (3) definitions of the relationships between these variables, (4) explanations for the existence of the relationships between variables, and (5) boundary conditions. We identified 15 relevant studies, for an overview see Table 1. As a first step, we explored the phenomenon of interest in the studies. We looked at what exactly the studies focused on in order to understand how well the described results might fit our phenomenon of interest, the impact of DT. This revealed three broad categories. First, DT and its elements as the phenomenon of interest itself. Better understanding DT itself remains an issue of imminent importance to the academic and practitioner community. Researchers investigated common elements, such as underlying principles, or most commonly used methods (Micheli et al., 2019), the roots of DT and potential future developments (Johansson-Sköldberg et al., 2013), or how practitioners enact DT (Carlgren et al., Carlgren, Rauth, et al., 2016) in order to better understand DT as a concept. The second topic of interest is around the implementation of DT. Guiding questions of such research are: What are factors influencing the implementation of DT (or design) in organizations and what hinders or enables a successful implementation (Carlgren et al., Carlgren, Elmquist, et al., 2016; Micheli et al., 2018)? And last, we identified studies investigating the impact of DT or of design. Starting with the impact of design, researchers show how design can add value for management (e.g., Borja de Mozota, 2010) or the role of design specifically for new product development (Perks & Cooper, 2005). Focusing more on DT but being broader in terms of impact areas, studies suggest an impact of DT by way of being a mechanism for brand ambidexterity (Beverland et al., 2015) and in enabling a learning process for innovation activities (Beckman & Barry, 2007). Further studies show DT's impact through cognitive bias reduction (Liedtka, 2015) or how DT is interrelated to organizational culture (Elsbach & Stigliani, 2018). We identified four studies with a focus on the impact of DT as related to performance. One study found an impact on the performance (e.g., idea generation and selection) of novice multidisciplinary teams (Seidel & Fixson, 2013). Another one showed the positive effects of DT in the US Federal Government (Liedtka et al., 2019). The study closest to investigating the impact of DT in organizations (in understanding how it relates to performance) looks at what DT activities lead to what kind of impact in organizations overall (Liedtka, 2018). Another study focused mainly on effects on the financial performance of IBM through the usage of DT (IBM, 2018; please refer to the next section for a more detailed description).

After gaining a profound understanding of the various phenomena of interest, we investigated what type of framework the existing research offers. Investigating DT as a phenomenon was the aim of three studies, all offering rich descriptions of the phenomenon, but lacking the description of relationships between constructs. Therefore, these study do not provide a model for DT overall and thus not for DT impact specifically (Johansson-Sköldberg et al., 2013; Carlgren et al. Carlgren, Rauth, et al., 2016; Micheli et al., 2019). The also applies to the studies investigating the implementation of DT or design (Carlgren et al. Carlgren, Elmquist, et al., 2016; Micheli et al., 2018).

The biggest cluster focusing on the impact of DT consists of two subareas. Four studies focused on the impact of DT in specific areas, such as culture, cognitive bias reduction, innovation as a learning process, or brand ambidexterity (Beckman & Barry, 2007; Beverland et al., 2015; Elsbach & Stigliani, 2018; Liedtka, 2015). Beckman and Barry provide a theory for innovation as a learning cycle, explaining how the different process steps of Design Thinking are related and why (Beckman & Barry, 2007). Yet they focus on the practice of DT and offer no indication or explanation for the various effects and values DT can bring to an individual, team, or organization. Two other studies provide models. Yet both focus on one specific area of impact, such as organizational culture (Elsbach & Stigliani, 2018) or brand ambidexterity (Beverland et al., 2015), not providing a model that helps to understand the bigger picture of DT impact.

The other four studies focus on the impact of DT in general or focus on the performance of teams or the organization. None of these offers a theory. Closest to providing a model comes Liedtka's working paper (Liedtka, 2018), giving rich narratives of observations as well as constructs and indications for relationships between the constructs. Yet, explanations of why and how exactly these constructs are related are missing. Only one of these 15 studies offers a theory (Beckman & Barry, 2007). This theory aims to explain how innovation results through learning. Therefore, the phenomenon of interest is not directly related to the impact of DT and it cannot be used as an explanation for the value of DT in organizations. Overall, a theoretical framework to explain the impact of DT in organizations is still needed (Table 1).

3.2 Forrester IBM Study in Focus

A rather prominent study, which for the first time delivered figures for DT impact, was published by IBM and carried out by the consultancy Forrester (IBM, 2018). According to its key result, the Return on Investment (ROI) for Design Thinking amounts to 301%, which means that every dollar invested in DT leads to a return of 3.01 dollars. Since this study is well-acknowledged by practitioners, we seek to introduce it in more detail and set it in the context of the aim of this chapter of assessing DT impact.

Table 1 Overview of DT models and theories

	Phenomenon of interest	Rich description	Definition of constructs	Relationships btw constructs	Justification of relationships	Boundaries of theory
Carlgren, Rauth, Elmquist (2016)	DT is phenomenon	√	√	x	x	x
Johansson-Sköldberg et al. (2013)	DT is phenomenon	x/√	x	x	x	x
Micheli et al. (2019)	DT is phenomenon	√	√	x	x	x
Micheli et al. (2018)	Implementation of design	√	√	x/√	x	x
Carlgren, Elmquist, Rauth (2016)	Implementation of DT	√	√	x/	x	x
Borja de Mozota (2010)	Impact of design	√	√	x/√	x	x
Perks, Cooper, Jones (2005)	Impact of design	√	√	x	x	x
Beckman and Barry (2007)	Impact of DT	√	√	√	√	√
Beverland et al. (2015)	Impact of DT	√	√	√	x	x
Elsbach and Stigliani (2018)	Impact of DT	√	√	√	x	x
Liedtka (2015)	Impact of DT	√	√	x	x	x
Seidel and Fixon (2013)	Impact of DT (performance)	√	√	x	x	x
Liedtka (2018)	Impact of DT (performance)	√	√	x/√	x	x
Liedtka et al. (2019)	Impact of DT (performance)	√	√	x	x	x
IBM by Forrester, 2017	Impact of DT (performance)	√	x/√	x	x	x

Other than the models introduced earlier, the understanding of the outcome of DT in this study is a purely financial one. While the authors acknowledge that there are effects of DT on culture, employee engagement, and the brand and also found evidence in their qualitative interviews for these factors, they are considered as "unquantified benefits." This perspective can be better understood if the study is considered in its larger context. As its title, "The Total Economic Impact™ Of IBM's Design Thinking Practice," indicates, this study follows the tradition of so-called economic impact analyses, whose aim is to show the economic impact of a project or organization. Regularly used in the field of culture, economic impact analysis for a music festival, for instance, shows how each dollar given as a subsidy will flow back into the region through income generated at the event (tickets, merchandise, food), but also through surrounding activities (accommodation, shopping, restaurant visits; Crompton et al. 2001). The approach used by Forrester in the IBM study follows a similar procedure, translating benefits of DT into dollar amounts, subtracting costs associated with DT, and correcting these through a risk premium. More specifically, Forrester conducted interviews with four of IBM's clients and 60 executives to construct a composite organization. Based on such an organization, the study calculates benefits, such the reduction of personnel costs due to faster project cycles and increased sales due to a faster time to market. Costs include personnel costs for the project teams or the fees paid to IBM as a consultancy. Again, these costs were not real numbers, but rely on the composite organization. The approach is well-documented in the appendix of the study, and the authors make very clear that the study measures marginal profit increases, better investments, and reduced risks, regarding labor costs and costs directly attributable to DT.

While this study is the first to provide a specific number for the impact of DT, its results refer to the impact of DT as implemented by IBM for their clients, it also relies on several assumptions and defines very specific areas for impact. In general, such a positive and high ROI of DT is promising. Yet, in addition, it would be of interest to apply the principles inherent to DT, such as gestalt view or human-centeredness (Micheli et al. 2019), also to the assessment of the impact of DT. This implies, for instance, considering a scorecard approach for impact that not only considers financial numbers, but also cultural impact. However, based on early evidence, such a larger scope might actually uncover more benefits than costs and might thus in fact demonstrate a larger impact of DT going beyond the ROI of 301%.

3.3 Design-Driven Companies

Another angle that provides first insights into the impact of DT in organization is to look at the broader concept of design and its impact. Here, the Design Management Institute (DMI), the largest global (non-profit) community of designers and innovators, has created the Design Value Index (Rae, 2016). This Index allows us to compare design-led companies with the S&P 500, showing that from 2005–20,015 design-led companies, have maintained significant stock market advantage, outperforming the S&P by 211% (Rae, 2016). Besides this financial measure, based on stock comparison, the DMI reports eight ways, in which design adds value to organizations (Rae, 2013). The first mechanism described is called *The WOW Factor*, which refers to the fact that "Great design helps make products and services more aesthetically pleasing, more compelling to use, and more relevant in a world that seems to change at an ever-increasing pace." (Rae, 2013, p.32). The second way design is adding value to organizations is through *Brand Expression*, allowing customers to connect to the brand in more depth. *Solving Unmet User Needs* is the third added value, emphasizing the power of empathy. With *Developing Better Customer Experiences*, design creates value by supporting seamless and differentiating experiences. For the fifth way, *Rethinking Strategy*, the DMI refers directly to DT as a way to understand complex problems, considering multiple solutions and therefore offering new strategic directions. Great design also influences how well interactions play into each other, referred to as *Hardware/Software/Service Integration*. With *Market Expansion Through Personal Development and User Understanding*, designers foster the acquisition of new types of customers all over the world. And last, *Cost Reduction* emerges when, for example, production process is reinvented.

These eight ways give great first insights in ways that design can impact organizations. Yet for our interest, it falls short in two ways. First, it focusses on design, which can be seen as closely related to DT, but yet is fundamentally different when it comes, for example, to the question of who is a designer and whether it is a distinct function or an overall embedded approach. And second, the described ways seem to mix various levels, when we consider the logic model approach. We see no clear distinctions between activities, outputs, outcomes or impact, making it difficult to identify elements of a model which might be transferable to the context of DT impact in organizations.

4 Discussion

This chapter has introduced an overview of the term impact, using the logic model approach. Furthermore, we described and compared existing performance measurement approaches (Balanced Scorecard and EFQM) and discussed the potential implications for measuring and evaluating the impact of Design Thinking. In addition, we gave an overview of the current literature dealing with the assessment of DT and design in organizations, showing the current lack of a theoretical framework explaining the impact of DT.

From the above-described status of the literature, we derived five key implications for the assessment of DT impact. First, there is no impact without knowing the (defined) strategic intent. Impact cannot be measured in terms of single indicators, but must be evaluated in terms of multiple indicators that are modeled with their direct and indirect relationships. These include the acknowledgment of higher-order concepts and mediators, meaning there needs to be an anchor point in the form of a specific problem or goal that should be tackled. Only the existence of such a goal allows us to evaluate whether the conducted activities can be seen as successful. Measuring DT activities and presenting measures for output and outcomes can be valuable in order to steer these activities, nevertheless they need to be clearly differentiated from the term impact, which includes and evaluation (Harris Mulvaney et al., 2006; Brousselle & Champagne 2011).

Second, when it comes to impact, it might be helpful to consider the inside and outside perspective. Indicated by Balanced Scorecard logic (Kaplan & Norton, 1996), it is helpful to consider the external connection to the market through Customer and Financial Criteria, whereas Learning and Growth and Internal Business Processes might help understand, how the organization is set up internally in order to perform in the best possible way. So looking into value added by DT initiatives, it might make sense to measure output and outcomes for these inside and outside perspectives as well as looking into the strategic intent and, respectively, the DT impact for these two sides.

Connected to these two sides is our third implication. Learning from models for creativity in organizations, it seems relevant to keep multiple levels in mind (Woodman et al., 1993). Both, for the inside and the outside views it is important to differentiate between individual and organizational factors when evaluating the impact of DT.

A fourth implication based on other performance-related evaluation frameworks such as the EFQM model and the Balanced Scorecard is the consideration of a wider range of valuable outputs and outcomes beyond financial factors. Investigating the role of leadership, for example, might be a fruitful step in order to gain a more in-depth understanding of the impact of DT.

The last and overarching implication is that there is a need for a model describing the impact of DT for organizations. Future research aiming for the creation of such a model would deliver contributions for practitioners and academics alike. Developing a theoretical model for capturing Design Thinking impact will offer three central contributions for practitioners. First, it provides an overview of the various dimensions and in particular highlights that the impact of Design Thinking goes beyond the established KPIs and also has to consider more soft factors. Thus, second, it provides practitioners who want to introduce Design Thinking in their organizations with helpful and valuable arguments and gives a first indication for their own measurement approaches. Third, summarizing existing research and presenting it in a comprehensive model might make research more accessible to the practitioner community. In a similar vein, the model delivers three important contributions for researchers. First, it tackles the unanswered question of DT impact by a theoretically sound model and definition. Second, and as a consequence, such a theoretical understanding provides the foundation for subsequent measurements and evaluations. Third, it provides an overview of existing work and allows other researchers to better locate and integrate their work into the field of DT impact.

Acknowledgements The authors want to thank Sascha Hoever for his support with the creation of the displayed illustrations. Additional thanks to Dr. Sharon Nemeth for her copyediting support.

References

Beckman, S. L., & Barry, M. (2007). Innovation as a learning process: Embedding design thinking. *California Management Review, 50*, 25–56. https://doi.org/10.2307/41166415

Bersin J, Solow M, Wakefield N (2016) *Design thinking—Crafting the employee experience* (pp. 67–77). Deloitte University Press

Beverland, M. B., Wilner, S. J. S., & Micheli, P. (2015). Reconciling the tension between consistency and relevance: Design thinking as a mechanism for brand ambidexterity. *Journal of the Academy of Marketing Science, 43*, 589–609. https://doi.org/10.1007/s11747-015-0443-8

Borja de Mozota, B. (2010). The four powers of design: A value model in design management. *Des Manag Rev, 17*, 44–53. https://doi.org/10.1111/j.1948-7169.2006.tb00038.x

Bourne, M., Franco-Santos, M., Micheli, P., & Pavlov, A. (2018). Performance measurement and management: A system of systems perspective. *International Journal of Production Research, 56*, 2788–2799. https://doi.org/10.1080/00207543.2017.1404159

Brousselle, A., & Champagne, F. (2011). Program theory evaluation: Logic analysis. *Evaluation and Program Planning, 34*(1), 69–78.

Carlgren, L., Elmquist, M., & Rauth, I. (2016). The challenges of using design thinking in industry—Experiences from five large firms. *Creativity and Innovative Management, 25*, 344–362. https://doi.org/10.1111/caim.12176

Carlgren, L., Rauth, I., & Elmquist, M. (2016). Framing design thinking: The concept in idea and enactment. *Creativity and Innovation Management, 25*, 38–57. https://doi.org/10.1111/caim.12153

Cooper, R., Junginger, S., & Lockwood, T. (2009). Design thinking and design management: A research and practice perspective. *Design Management Review, 20*, 46–55. https://doi.org/10.1111/j.1948-7169.2009.00007.x

Crompton, J. L., Lee, S., & Shuster, T. J. (2001). A guide for undertaking economic impact studies: The Springfest example. *Journal of Travel Research, 40*(1), 79–87.

EFQM. (2003). Introducing excellence. European Foundation of Quality Management.

Elsbach, K. D., & Stigliani, I. (2018). Design thinking and organizational culture: A review and framework for future research. *Journal of Management, 44*, 2274–2306. https://doi.org/10.1177/0149206317744252

Foundation. W.K. Kellogg (2004). Logic model development guide.

Harris Mulvaney, R. R., Zwahr, M., & Baranowski, L. (2006). The trend toward accountability: What does it mean for HR managers? *Human Resource Management Review, 16*(3), 431–442.

IBM. (2018). The total economic impact™ of IBM's design thinking practice.

Johansson-Sköldberg, U., Woodilla, J., & Çetinkaya, M. (2013). Design thinking: Past, present and possible futures. *Creativity and Innovation Management, 22*, 121–146. https://doi.org/10.1111/caim.12023

Julian, D. A., Jones, A., & Deyo, D. (1995). Open systems evaluation and the logic model: Program planning and evaluation tools. *Evaluation and Program Planning, 18*, 333–341. https://doi.org/10.1016/0149-7189(95)00034-8

Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard: Translating strategy into action.* Harvard Business Review Press.

Kaplan, R. S., & Norton, D. P. (2004). *Strategy maps: Converting intangible assets into tangible outcomes.* Harvard Business Review Press.

Kneale, D., Thomas, J., & Harris, K. (2015). Developing and optimising the use of logic models in systematic reviews: Exploring practice and good practice in the use of programme theory in reviews. *PLoS ONE, 10*, 1–26. https://doi.org/10.1371/journal.pone.0142187

Kolko, J. (2015). Design thinking comes of ages. *Harvard Business Review, 93*, 66–71.

Liedtka, J. (2015). Perspective: linking design thinking with innovation outcomes through cognitive bias reduction. *Journal of Product Innovation Management, 32*, 925–938. https://doi.org/10.1111/jpim.12163

Liedtka, J. (2018). *Exploring the impact of design thinking in action. Darden Working Paper Series.*

Liedtka, J., Salzman, R., & Azer, D. (2017). *Design thinking for the greater good: Innovation in the social sector*. Columbia University Press.
Liedtka J, Sheikh A, Gilmer C, Kupetz M, Wilcox L (2019) The use of design thinking in the U.S. Federal Government. Public Performance & Management Review 0:1–23. https://doi.org/10.1080/15309576.2019.1657916.
McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your programs performance story. *Evaluation and Program Planning, 22*, 65–72. https://doi.org/10.1016/S0149-7189(98)00042-1
Micheli, P., Perks, H., & Beverland, M. B. (2018). Elevating design in the organization. *Journal of Product Innovation Management, 35*, 629–651. https://doi.org/10.1111/jpim.12434
Micheli, P., Wilner, S. J. S., Bhatti, S. H., Mura, M., & Beverland, M. B. (2019). Doing design thinking: Conceptual review, synthesis, and research agenda. *Journal of Product Innovation Management, 36*, 124–148. https://doi.org/10.1111/jpim.12466
Perks, H., & Cooper, R. (2005). Characterizing the role of design in new product development: An empirically derived taxonomy*. *Journal of Product Innovation Management, 22*, 111–127.
Pettigrew, A. M. (2011). Scholarship with impact. *Brazil Journal of Management, 22*, 347–354. https://doi.org/10.1111/j.1467-8551.2011.00769.x
Rae, J. (2013). What is the real value of design?
Rae, J. (2016). The power & value of design continues to grow across the S&P 500. *DMI Review, 27*, 4–11.
Seidel, V. P., & Fixson, S. K. (2013). Adopting design thinking in novice multidisciplinary teams: The application and limits of design methods and reflexive practices. *Journal of Product Innovation Management, 30*, 19–33. https://doi.org/10.1111/jpim.12061
Tucker, P., Liao, Y., Giles, W. H., & Liburd, L. (2006). The REACH 2010 logic model: An illustration of expected performance. *Preventing Chronic Disease, 3*, A21.
Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review, 14*, 490–495.
Wiesche, M., & Yetton, P. W. (2017). Grounded theory methodology in information systems research. *MIS Quarterly, 41*, 685–701.
Woodman, R. W., Sawyer, J. E., & Griffin, R. W. (1993). Toward a theory of organizational creativity. *Academy of Management Review, 18*, 293–321.

Performance Measurement of Design Thinking: Conceptualisations, Challenges and Measurement Approaches

Thomas Haskamp

Abstract Design Thinking receives growing attention from both practitioners and scholars alike. While companies have adopted the methodology in various contexts, an increasing number of practitioners ask for specific advice on measuring Design Thinking in its different manifestations. Reviewing existing literature on Design Thinking and Innovation measurement, this chapter derives different streams of understanding on performance measurement of Design Thinking and Innovation. Following one stream, we investigate challenges that practitioners experience when Performance Measurement instruments are adapted for Design Thinking. Addressing these challenges, this chapter proposes a framework differentiating between measuring Design Thinking as a set of methods, as a process for innovation, and as a mindset driving transformation. This contributes as it sketches to different paths of measuring design thinking, which helps theory and practice to investigate how to quantify Design Thinking.

1 Introduction

Design Thinking (DT)—a human-centred approach to innovation making use of design methods—has gained increasing attention over the last years (Liedtka, 2015; Micheli et al., 2019). While the role of DT has evolved into a more strategic one (Kolko, 2015), scholars and practitioners alike have raised questions about the measurement of DT with different intentions (Brown, 2018; Liedtka, 2017; Rauth et al., 2014). Some have raised the question of the impact of DT (Rauth et al., 2014) seeing the proof of impact as the *"Holy Grail"* of DT. Others have investigated measurement looking on the team level with the attempt to understand and leverage team performance in design thinking projects (Royalty and Roth 2016b; Roth et al., 2020; Royalty et al., 2019). Another group of scholars in the community have looked on the organisational level trying to help companies understand the value of design (Sheppard et al., 2018; Westcott et al., 2013; Brown, 2018; Liedtka, 2017).

T. Haskamp (✉)
Hasso Plattner Institute for Digital Engineering, University of Potsdam, 14482 Potsdam, Germany
e-mail: Thomas.Haskamp@hpi.de

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_15

While these streams of research helped to get a better understanding of DT as a concept used in practice and academia, both, theory and practice claim the necessity to develop performance measurement instruments to more systematically steer and measure DT in organisations (Björklund et al., 2020; Wrigley et al., 2020; Mayer et al., 2021).

At the same time, while stakeholders in organisations want to measure, the performance measurement of DT seems challenging as the concept is difficult to grasp with classical performance measurement indicators (Schmiedgen et al., 2016). Surprisingly, research on innovation—a key promise of DT—shares a variety of metrics to measure (Dziallas & Blind, 2019).

As shown, while the need to measure DT is increasing due to its increasing diffusion in organisations, research has not addressed this topic accordingly. Furthermore, existing research cannot explain the existence of metrics for innovation against the non-existence of DT on the other side.

To address these topics and to better understand how performance measurement in the context of DT can be conceptualised, we investigated three research questions. To enhance our understanding and shed light on the opaque relationship between DT and Innovation, we first aim to conduct a literature review (Vom Brocke et al., 2009) to identify existing streams on measuring DT. This translates into the first research question (RQ):

RQ1: What are the existing conceptualizations of measuring Design Thinking?

Focusing on one particular stream, we will explore the challenges of using performance measurement instruments by relying on an exploratory qualitative design (Edmondson & Mcmanus, 2007) with practitioners. This translates into the following second research question:

RQ2: What are the challenges of measuring the performance of Design Thinking in practice?

Answering this question will not only enhance our understanding of measuring DT in practice. Combined with literature, it also serves as the foundation to propose a framework that intends to sketch different measurement approaches for different understandings of DT. This can be translated to the third research question:

RQ3: What are different measurement approaches for different understandings of Design Thinking?

Addressing these research questions contributes as it helps to improve our understanding of the performance measurement of DT. This sophisticated understanding can then be used to develop suitable performance measurement instruments.

This book chapter starts by providing an understanding of DT. Afterwards, borrowing from management literature, we introduce the concept of performance measurement. We then present our methodology, which is followed by our results structured along with the research questions. Afterwards, we discuss the findings we have arrived at, as we then conclude with an outlook on potential future work.

2 Design Thinking—Conceptual Understanding

For defining DT, we follow Brenner et al. (2016) who differentiate between DT as a set of methods, as a process, and as a mindset. While we are aware that different understandings and schools of thought have coined the term itself (see for example Johansson-Sköldberg et al., (2013) for a review), we find this definition helpful. The reason for this is that, in our study, it serves us as a typology for the different manifestations of DT in practice. Furthermore, recent understandings of literature on DT (Björklund et al., 2020; Micheli et al., 2019) can be mapped quite well towards the different levels of DT.

2.1 Design Thinking as a Set of Methods

This most simplified understanding of DT conceptualises the approach as a collection of methods from different domains that are used in the context of innovation (Brown, 2008; Seidel & Fixson, 2013). While certainly not exclusive, existing literature provides an extensive overview of existing methods and tools that are used under the umbrella of DT (Uebernickel and Brenner 2020; Lewrick et al., 2018). A recent literature review identified the following key methods most mentioned in academia in the context of DT: ethnographic methods, personas, journey maps, brainstorming techniques, mind maps, visualisation methods, prototyping techniques and design of experiments (Micheli et al., 2019). While one could see these methods as firmly embedded in the process steps (see 2.2. Design Thinking as a Process), many tools are also adopted themselves serving a specific problem-solving intention in a given context unrelated to DT.

2.2 Design Thinking as a Process

In terms of understanding DT as a process, different phase models are available (Brenner et al., 2016; Brown, 2008; Kelley & Kelley, 2013). Giving an overview of all models would take us beyond our intention here, therefore we will only introduce the process model of Brenner et al. (2016) which has strong roots at the ME310 course of Stanford University.

Although different models differentiate in different phases and process steps, the models share the intention to deliver an innovative solution (Kelley & Kelley, 2013) or a final prototype (Brenner et al., 2016).

To point out one particular understanding in-depth, we now turn to introduce the process model of Brenner et al. (2016), which builds on the micro and macro cycle of DT. The micro cycle (depicted in Fig. 1), consists of the recurring steps of (re)defining the problem, needfinding and synthesis, ideation, prototype and testing.

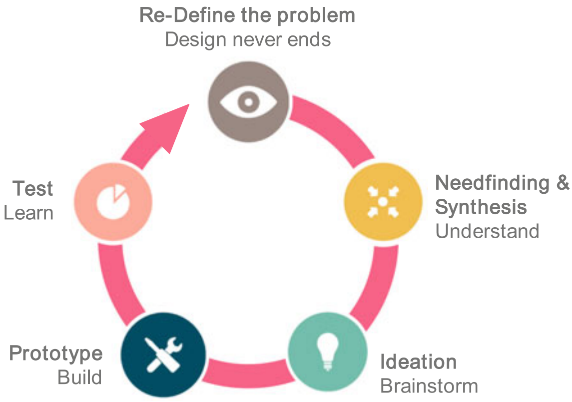


Fig. 1 Micro cycle of design thinking adapted from Brenner et al. (2016)

This micro cycle is embedded in the macro cycle, which follows two main parts. In a first diverging part, teams explore the design space, prototype potential critical functions and stretch the space of potential solutions by building dark horse prototypes. In a second converging phase, teams combine their learnings and insights and build different prototypes (funky, functional, x is finished) towards the final solution (Fig. 2).

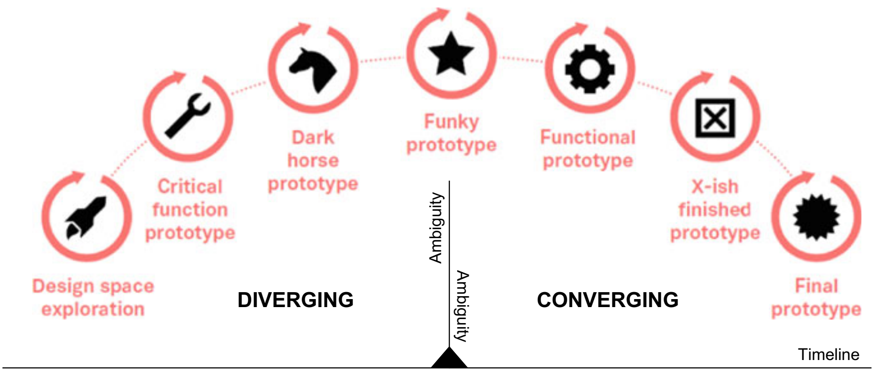


Fig. 2 Macro cycle of design thinking adapted from Brenner et al. (2016)

2.3 *Design Thinking as a Mindset*

Understanding DT as a mindset refers to its underlying principles that aim to drive the mental model of design thinkers (Carlgren et al., 2016; Micheli et al., 2019; Brenner et al., 2016). Principles attached to the concept are, for example, human-centeredness, customer-centricity, experimentation, diversity and problem framing

(Carlgren et al., 2016; Brenner et al., 2016). Understanding DT on this principle level and fostering these principles has made scholars call delete -inging it a “Social Technology” (Liedtka, 2020). Often with the understanding of DT as a mindset comes the intention and understanding of design as a catalysator of transformation. For example, in his foundational article (1969) Simon describes early on Design as the *“transformation of existing conditions into preferred ones (p. 4).”*

Nowadays, when this intention that evolved from Simon’s work is applied to managerial problems, *“design can be seen as a cultural transformation process within business”* (Björklund et al., 2020, p.2).

While for now DT, has been conceptualised accordingly, the following chapter will introduce some key concepts of performance measurement as the second key term.

3 Performance Measurement

Performance Measurement in management relates to the use of procedures and tools that intend to improve the efficiency and effectiveness of organisations. Thereby, organisations benefit on three levels (Micheli & Mari, 2014):

1. Formulation, implementation and review of organisational strategy
2. Communication of the results achieved to stakeholders, and strengthening of brand and reputation
3. Motivation of employees at all levels, creation of a performance improvement culture and fostering of organisational learning.

Putting these concepts into practice, management literature has proposed different approaches and dimensions of performance measurement (Neely et al., 2000). These include well-known measurement approaches such as the Balanced Scorecard (Kaplan & Norton, 1992) or Du Ponts Pyramid of Financial Ratios (Chandler, 1977). Following well-known management phrases such as: “What gets measured, gets done” or “If you cannot measure, you cannot manage” Performance Measurement has been implemented in different ways to many organisations (Micheli & Mari, 2014). This has not been without criticism as measurement relies on the key assumptions that behaviours and action always follow measurement and that measurement properties such as objectivity, accuracy, precision are taken for granted (Micheli & Mari, 2014).

Especially in the context of innovation, the use of performance measurement instruments has been the subject of disagreement (Chenhall & Moers, 2015). While some are sceptical, as measurement instruments may diminish creativity (Hennessey & Amabile, 2010) and lead to misleading incentives for employees (Ylinen & Gullkvist, 2014), others argue through adjustments in the design of such instruments, performance measurement can enhance information exchange in organisations (Davila et al., 2009) and ensure alignment in teams (Bedford et al., 2019).

3.1 The Triple P-Model of Performance

The use of the concept Performance Measurement requires a clear conceptualization of the term performance. While the term is used in a variety of contexts, the meaning of the concept can defer greatly from the context in which it is used (Tangen, 2005). For this study, we rely on the "Triple P-Model of Performance (Tangen, 2005). We found it helpful to increase concept clarity and to clearly distinguish between the different terms often used, such as performance, productivity, efficiency and effectiveness (Fig. 3).

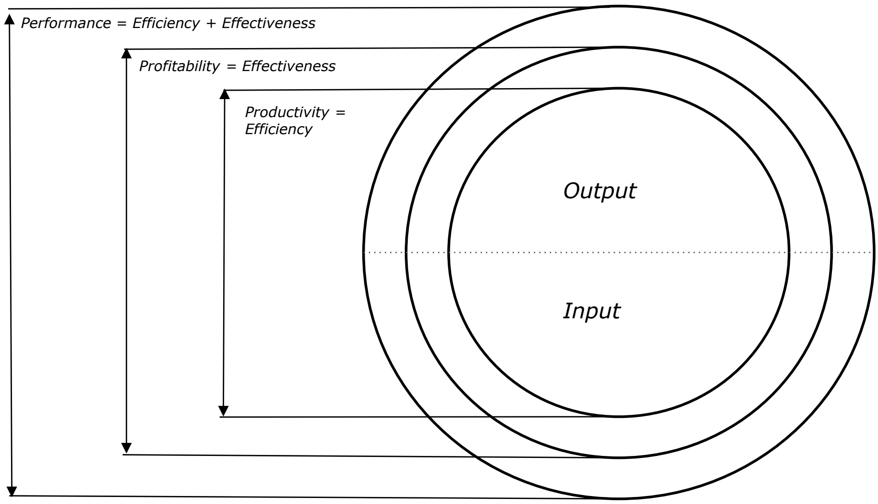


Fig. 3 Triple P-model of performance adapted from Tangen (2005)

The model explains performance considering two dimensions: efficiency and effectiveness. While efficiency (doing things right) relates to the ratio of input and outputs and thus looks at productivity, effectiveness (doing the right things) considers whether the organisation can translate their output into something that has value and generates an outcome.

4 Methodology

To address our three research questions, our research design consisted of three phases. In the first phase answering RQ1, we conducted a literature review following Vom Brocke et al. (2009). By doing this, we intended to identify existing streams on measuring DT and to shed light on the measurement of DT and the measurement of innovation.

	PHASE I **Literature Review** *Research Question 1*	PHASE II **Exploratory Interviews** *Research Question 2*	PHASE III **Framework Development** *Research Question 3*
Goal	• Gain an overview of different understandings of design thinking measurement and innovation measurement	• Identify challenges of measuring design thinking	• Development of measurement approaches for design thinking
Activities	• Literature Review on search terms design thinking, innovation, measurement	• Interviews with design thinking practitioners in digital innovation units	• Analysis of data from exploratory interviews and existing literature
Outcomes	• Three different understandings of measuring design thinking	• Challenges of measuring design thinking	• Framework of measuring design thinking in organisations

Fig. 4 Research process

In the second phase, targeting RQ2, we adopted a qualitative exploratory approach to identify challenges of using performance measurement instruments by running interviews with practitioners (Edmondson & Mcmanus, 2007).

In the third phase addressing RQ3, we combined our findings with the literature and derived a framework suggesting measurement approaches depending on the context of the use of DT (Fig. 4).

4.1 Literature Review

To gain an overview of the different understandings of the term measurement in relationship to DT and innovation, we conducted a literature review (Brocke et al., 2009). Thus, as the first step, we defined the scope of the review and conceptualised our topic around the two key terms of innovation and DT and added measurement as well as potential synonyms. Afterwards, we conducted a literature search. Therefore, we adopted two search strategies. Firstly, we searched for peer-reviewed publications with the keywords "design thinking" AND "measurement" OR "measure" OR "indicator" OR "metrics" in the title using a meta-database that included databases such as Web of Science, Scopus, Ebsco, etc..

We received 239 available publications of which four discuss measuring DT. Although these were published in journals or conferences below a C-Ranking following VHBs journal list, we decided to use these papers as they were the only ones our research identified on the topic. In a second step, we replaced the term design thinking with innovation which revealed 1.421 publications. Here again, we focused on publications that were ranked at least on a C-level which resulted in 18 relevant publications. Thus, in total 22 papers were subject to our analysis and synthesis.

4.2 Exploratory Interviews

In the second phase of our study, we conducted an exploratory study to identify challenges of using performance measurement instruments by running 20 exploratory interviews with practitioners (Edmondson & Mcmanus, 2007). The choice for this approach was made as not much knowledge on our study object—performance measurement of DT—existed. Thus exploratory interviews are a robust choice to answer RQ2 (Eriksson & Kovalainen, 2015). The first intention of these interviews was to investigate what kind of performance measurement instruments practitioners adopted. However, we quickly learned during the first interviews that performance measurement methods specifically for DT were non-existent. Besides exploring the applied measurement instruments, we explored the reasons for the lack of measurement methods that resulted in challenges of using performance measurement instruments.

Sample Selection: In terms of sample selection, we decided to look for interview partners in Digital Innovation Units (DIUs). These are "*organisational units with the overall goal to foster organisational digital transformation by performing digital innovation activities for existing and novel business areas*" (Barthel et al., 2020). The reason for this was that existing research highlights the application of DT for their exploratory endeavours (Barthel et al., 2020; Holotiuk and Beimborn, 2019). Many companies have set up such DIUs as part of their digital transformation initiatives with different intentions and organisational setups (Svahn et al., 2017; Barthel et al., 2020). We, therefore, approached several DIUs across industries and conducted 20 interviews between March and September 2020 with DIU members in Germany and Switzerland. Interviewees were positioned across hierarchies having roles such as Innovation Coach, CIO, CTO, Innovation Manager, DIU Lead or Business Development Manager. We aimed to identify a diverse set of participants in different industries to reveal different perspectives on measuring DT. This may lead to contrasting results, which enhances the validity of the study (Edmondson & Mcmanus, 2007). The average interview duration was 48 min and the interviews held 6 years of experiences within their jobs. To ensure that participants were able to express themselves as well as possible, interviews were conducted in their native language.

The interview guideline was semi-structured, constantly iterated and followed different areas. First, we asked the subjects about their understanding of DT and how they applied the methodology. Afterwards we asked questions about the general setting of the DIU trying to understand the context in which the method was applied. We then posed questions on how they reported progress on their DT activities and if they had any measurement approaches in place. This question part was adapted to the hierarchy level of the interviews. While some interviewees were in the role to measure and report progress itself, others higher up the hierarchy were in a position to do both, gather status updates from teams but also, depending on their situation, report to other stakeholders. As the interviews quickly revealed, only rarely did interviewees use specific measurement approaches for DT, we decided to explore the reasons for this more deeply. Thus, our interview guide was constantly iterated, as the case for such a study design is to refine and iterate findings (Gioia et al., 2013).

Data Analysis: For analysing our qualitative data, we adopted a grounded theory approach aiming to provide thick descriptions of new phenomena which have been recognised as a valuable contribution (Wiesche et al., 2017).

Table 1 Exemplifying data structure for challenge I

Quote	First order concepts	Second order themes	Aggregated Dimension
"Why we started this, well my boss did not really define that in a clear way." (Innovation Coach)	Unclear DT intention	Intention of DT not clear	Intention and expectations concerning the use of design thinking
"Nevertheless, I think that we [as human-centered innovation unit] are usually approached to work out solutions and not to change working methods."(Innovation Manager)	DT is used to develop solutions	Intention of DT is innovation	
"After all, Design Thinking is all about cost-effectiveness, knowing the customer's needs and delivering the technical solution. Profitability is one element." (DIU Lead)	DT is used to deliver cost-effective and profitable solutions		
"Design Thinking, there has been considerable pressure in recent months to show how anything - in other words, how success can be measured. Whether it's savings or something else, because there were no clear KPIs or because the KPIs that had been developed before were no longer considered important". (Innovation Manager)	DT is under pressure to show benefit and potential savings	DT intention should be measurable and success needs to be defined	
"So, it's quite possible that even with the application of Design Thinking, you know, you may end up with a completely different set of requirements than you started with." (CIO)	Results of using DT are difficult to predict	DT results are difficult to plan for in advance	

First, all interviews were transcribed and coded following the gioia methodology (Gioia et al., 2013). Thus, quotes were summarised to first order concepts which were then grouped into second order themes. These second-order themes were grouped into aggregate dimensions which were interpreted as challenges. Emerging codes and dimensions were discussed in the team and iterated several times. An example of our data structure for the first challenge can be seen in Table 1.

4.3 Framework Development

In the third phase of our study, relying on both the results of the literature review as well as our findings from the exploratory interviews, we developed our DT measurement framework. From theory, the different conceptualizations of DT following Brenner et al. (2016) were used. Adding our findings from the exploratory study and addressing the identified challenges from practitioners, the framework aims to point towards different paths in which context DT can be measured. While these paths are not exhaustive, grounded in our data and in theory, they point towards suitable measurement approaches that consider a robust understanding of DT and the intention of its users.

5 Performance Measurement of Design Thinking

This results chapter follows our three introduced research questions. We first show the findings of our literature review and explain existing approaches on the measurement of DT and accordingly, their streams of literature (RQ1). Addressing our second research question, we introduce three challenges that practitioners in our sample experienced. Last but not least, we show our framework answering RQ3.

5.1 Literature Review: Existing Approaches and Streams

The literature review analysing existing research related to the terms of DT and innovation and their performance measurement resulted in three different understandings of measuring DT which answers RQ1.

We present the three streams that we found and a selection of representative literature in Table 2. While the overview of different streams within the literature is not exclusive, it allows us to closer conceptualise the term of measurement in the context of DT and it also provides some interesting learnings. Thereby, we explain each stream shortly.

The first stream of literature, *Performance Measurement of Innovation* is quite mature and literature in this stream presents several approaches, including specific

Table 2 Streams on measurement of design thinking and innovation

Stream	Description	Representative literature (selection)
Performance measurement of innovation	This stream of literature discusses the application of performance measurement instruments such as indicators and metrics to innovation and research development projects in companies	(Adams et al., 2006; Chiesa et al., 2009; Dewangan & Godse, 2014; Dziallas & Blind, 2019; Godener & Soderquist, 2004; Hagedoorn & Cloodt, 2003)
Performance effects and impact of design thinking	This stream of literature investigates whether the use of DT translates into performance outcomes such as company performance etc	(Liedtka, 2015; Nagaraj et al., 2020; Nakata & Hwang, 2020; Roth et al., 2020)
Measuring design thinking on the organisational level	This stream of literature discussed or mentions different measurement approaches how organisations can use metrics to steer DT	(Royalty and Roth 2016b; Royalty et al., 2019; Royalty and Roth 2016a; Schmiedgen et al., 2016; T. Björklund et al., 2020; Westcott et al., 2013; Wrigley et al., 2020; Carlgren et al., 2016)

measurement ideas. For example many (Adams et al., 2006; Dewangan & Godse, 2014; Dziallas & Blind, 2019; Godener & Soderquist, 2004) presents metrics (e.g. Innovation Spendings, Amount of Patents, Success Rate, Total cost R&D) and dimensions (e.g. Knowledge measurement, Process measurement, Customer measurement) to measure innovation efforts in companies. However, many of these studies conceptualise the term innovation rather to the concept of New Product Development (e.g. Godener & Soderquist, 2004). In comparing metrics within these papers with principles of DT, one only could see limited overlap. While these existing papers provide a few metrics, for example, the principle customer-centricity (e.g. NPS), we could not identify metrics that would relate to the principle of diversity or experimentation in these papers.

The second stream of literature *"Performance Effects and Impact of Design Thinking"* analyses the effects that DT has in organisations. For example, some studies relate DT to other constructs as psychological safety (Roth et al., 2020) to explain statistical performance effects. Similarly, others in this direction operationalise DT mindsets and actions and investigate effects on new product and service performance (Nakata & Hwang, 2020). While this work is important and relevant to analyse the impact of DT on the construct level, this understanding of measurement was not the subject of this study as our intent lies in developing performance measurement methods that help organisations to track and measure their activities.

We found this intended understanding closely related to the third stream of research, which is particularly relevant for the understanding of measurement in this study. The Design Value Scorecard (Westcott et al., 2013), for example, provides a framework to quantify the impact of design on business and also to develop different maturity levels of DT. While the work of the Design Management Institute has provided tools and methods to assess the maturity of DT in organisations, tools and metrics to assess, steer and manage DT initiatives are lacking.

The work conducted by Royalty and Roth (2016b), Royalty et al. 2019, Schmiedgen et al. (2016) as part of the Hasso Plattner Design Thinking Research program has made valuable first steps. For example Schmiedgen et al. (2016) proposed different learnings on measurement DT. For example, they mention the challenges of defining DT and having appropriate measures for DT. Royalty and Roth (2016a, 2016b) have developed ecologies and metrics based on different elements of DT such as empathy measures (number of days gone without interacting with a customer, number of users spoken with, number of categories of people spoken with) a novelty/value grid for reframing and iteration measures (number of prototype iterations). Björklund et al. (2018) propose to use metrics based on the maturity of design in an organisation and distinguish between external and internal performance metrics.

All of these efforts, especially the third stream, which closely followed our intentions, provided valuable steps towards measuring DT. However, a research gap seems to be evident when it comes to using performance measurement instruments to steer and measure DT in companies systematically (Björklund et al., 2020; Wrigley et al., 2020; Mayer et al., 2021).

5.2 *Exploratory Study: Challenges of Measuring Design Thinking Performance*

From the analysis of our interviews, we identified three major challenges of measuring DT activities. The first one deals with the intention and expectations that are raised when DT is used in environments. Often the intention does not seem to be clarified from the beginning. Second, once the intention of DT is defined, we saw issues emerging when it came to finding a suitable measurement approach.

Third, in connection with the two introduced challenges, tracing the effects of the application of DT proved to be a significant hurdle. This is because of the difficulties that arise in assigning the proper credit due for the results that emerge.

Challenge I: Intention and Expectations concerning the use of Design Thinking
The intentions and expectations when DT is applied in business contexts often seem unclear to all involved stakeholders. Sometimes, there was no clear intent at all for using DT. In other areas, DT was used with both exploratory and exploitative intentions and also to achieve innovation and transformation at the same time. While all these goals and intentions seem quite similar, their outcomes and expectations vary greatly.

In our sample, some situations were mentioned where the reason behind using DT was unclear. One interviewee stated, for example, when asked about his intention: *"Why we started this? Well, my boss didn't really define that in a clear way" (Innovation Coach).* This quote shows that in some cases people are curious about the methodology and initiate projects where outcomes are not always clear and well-defined.

While sometimes the intention is not set very clearly, at some point a hidden expectation emerges. This is the case when someone from the management level asks about outcomes, having in many cases already something in mind that he or she wanted to see. Often team members using DT explained that this was quite surprising for them as no clear intention was defined in the beginning.

Another interesting point that came up is the flexibility of DT being applied in many different contexts and situations. For example, one may use DT methods in projects that are intended to explore new markets through ethnographic methods, or in another project that aims to initiate a cultural transformation in the company, or even as the first phase in developing a specific product. While this is not a problem itself, it translates into a challenge when it comes to measurement. For example, one interviewee explained that the exploratory nature of DT could lead to different outcomes than what was expected or even planned for in the beginning of the projects. In particular, the person said: *"We did not specifically use [Measurements], it is difficult, I think, to do that. […]. You know, the DT process setup causes you to rethink and rethink and rethink as you're going through every stage. So, it's quite possible that even with the application of design thinking, you know, you may end up with a completely different set of requirements than you started with" (CIO).* Thus, in this exploratory mode, practitioners seem to avoid working with metrics that predefine the solution too precisely.

In a more exploitive setting, in which DT methods are used as part of the specific development of a product, it seems much easier to use and define metrics such as the Net Promote Score. This is again completely different when it comes to initiating transformation activities in companies. One interviewee for example explained: *"Nevertheless, I think that we [as human-centred innovation unit] are usually approached to work out solutions and not to change working methods" (Innovation Manager).* This quote shows that in an exemplary use context of a DIU, while the methodology seems to have been applied in both very different contexts, it is difficult to focus only on one part. This can be seen as well in the following quote where a team used DT to push the organisational transformation while it wasn't their task to do so: *"Much changed in the organisation. But after two years they say: But that [transformation activity] wasn't your assignment" (DIU Lead).* Interestingly there are two aspects to this quote. The team conducting these activities claims that they had an impact and contributed to the overall organisational goal of transformation. Not being recognised for this may lead to disappointment from their side. From the management perspective the team did not deliver on their promise to deliver innovations. This miscommunication could only have been avoided if both sides would have agreed on specific targets and goals.

Challenge II: Finding the right measurement approach for Design Thinking As indicated, besides setting a clear objective for using DT, finding a suitable measurement approach is challenging. Our data showed that organisational members need to present short-term outcomes to secure organisational survival today. Still, often DT-related projects (innovation, transformation, exploration) have long-term impacts and cannot deliver up on short-term metrics. Another issue relates to the use of phase adequate measurement metrics. While comparable and standardised quantitative measurement approaches are requested from management, especially early activities can often only be assessed upon looking at flexible and case-dependent artifacts (such as a Business Model Canvas) or through a qualitative approach (showing user quotes, testing results, etc.).

DT is often applied in contexts such as innovation and transformation. To show the measurable impact of these intentions usually requires hard work and a substantial period of time. Interviews revealed that often it is quite challenging to argue for long-term activities while management wants to see short-term effects. For example, one interviewee explained: *"Especially in front of the CFO, who also wants to see results in the short-term, it is sometimes difficult to argue for the long-term"(DIU Lead).* While this exemplifying quote illustrates the challenges quite well, it also points to the issue of selecting the right assessor and evaluator for such activities on the organisational level. As the nature of such exploratory activities is quite different, it might require a special skill set to determine the quality of such work that is much different than how to assess a sales project in the company. Another interviewee confirmed this impression: *"We create a new idea or a new project proposal. Even if it is good, we get asked, what is the present value?" (DIU Lead).*

Along with the different phases of applying DT following one of the process models, teams can deliver other forms of results that are subject to measurement. Within the data, it was possible to recognise that while DT project teams tend to provide qualitative findings, the management teams prefer to rely on quantitative and preferable tangible outcomes. One interviewee for example explained: *"I can just say we did 20 interviews and got 10 times smarter. But yes, nice. Believe in it or don't believe in it" (DIU Lead).* As the quotes reveal, these kinds of metrics have challenges when it comes to credibility. This is stressed when it comes to higher management: *"The higher you go in the hierarchy level in a corporation, the more hard KPIs are needed" (CTO).* The quote shows that hard (meaning financial metrics) metrics seem to have high credibility while softer metrics (such as number of interviews) are seen as less credible. This points towards the perception of different measurement approaches in organisations. It implies considering not only the measurement approach, but also the mindset of the person assessing these kinds of activities.

In general, teams conducting such activities are open to show accountability but wish reasonable expectations as the following quote shows: *"The expectations were that the management wanted numbers. I have worked towards a situation where this is no longer the case. So, we have created free space. It needs pressure. Pressure's okay. But it doesn't need the pressure of sales at a very early stage. It's counter-productive"(DIU Lead).* Setting up these "free spaces" and equipping them with the right measurement approach that satisfies both stakeholders along different phases requires a compassionate approach that does not limit such exploratory endeavours while still providing accountability and a sense of control.

Challenge III: Gaining credits for Design Thinking activities
A third major challenge to measure DT activities is related to the limited ability to trace the efforts and resulting benefits of applying DT. As the method is used often early in projects, tracing of impact on projects is difficult.

Success has many fathers, and failure only one: being applied in early project stages, often projects are handed over to the responsible department for the market of the product. This handover is critical in the evaluation of success and failure. Thus, teams apply DT early and develop as an output a prototype or an MVP. When this output is handed over to the department to be further implemented and realised, accountability for outcomes is difficult to assign. If the project turns out to be a failure, was that because the methodology failed and the MVP already was problematic or was it because of a problem in the team responsible for implementing the prototype? Answering and tracing projects requires substantial efforts and is sometimes very challenging.

This can be also translated to assumed time savings that you can realise by applying withdelete with DT. As one interviewee explained: *"Well, actually the effort came at the end of projects [because we had to adjust to users' needs, we did not plan for in the beginning], and with using Design Thinking now, the effort came at the beginning of the project. So, it seems to be more effort if you just look at the short-term" (Innovation Coach).* The quote shows that with applying DT you often try to anticipate topics. Thus, interviewing stakeholders early might make it easier to secure their buy-in for the implementation later. However, if the time investment pays off is difficult to say as it is hypothetical to argue with savings that are hard to prove.

5.3 Framework Development: Conceptualising Design Thinking Performance Measurement

Relying on the literature review and the qualitative data gathered to answer RQ1 and RQ2, we now propose a framework that conceptualises measuring DT along three different paths answering RQ3.

The framework intends to use our sophisticated understanding of DT derived from literature and reflects the conceptualizations' specific intentions. Once the particular purpose of DT is defined, an adequate performance measurement approach can be developed.

The need for such a framework seemed to be apparent, bearing in mind that one main reason for the lack of measurement approaches was the lack of a clear definition of intent (Challenge I). Thus, addressing this challenge by relying on DT's different understandings and connecting these understandings with the intentions builds a robust base for developing specific measurement approaches.

However, the framework is not intended to cover the use of DT exhaustively. We instead understand it as the first conceptual framework that helps to structure the field. Thus, we synthesise existing research on performance measurement and DT, align them with our findings from the qualitative interviews. Accordingly, the specific development of the measurement approach needs to be conducted as part of further endeavours. The model is displayed in Fig. 5.

	Phase 1: Understanding of Design Thinking	Phase 2: Intent of Design Thinking	Phase 3: Measurement of Design Thinking
Path 1 Method	**Set of methods** - Approach: DT is understood as set of methods from different domains e.g. design, business (Brown, 2008). - Example methods: interviews, personas, journey maps, prototypes, field experiments (Micheli et al. 2019)	**Method-depending intent:** - Approach: Each method adopted with DT has its own intent which can differ on the specific use context. - Example: Interviews can be used to explore user needs (Uebernickel and Brenner 2020)	**Method-depending measurement approach:** - Approach: Measurement based on rigour and intent of method performance. - Example metrics for method exploratory interviews: amount of interviews, interviews per identified need, interviews per stakeholder group
Path 2 Innovation	**A process** - Approach: DT is understood as process that contains certain phases to pursue a goal (Brenner et al. 2016; Brown, 2008). - Example process phases: Design Space Exploration, Critical Function, Dark Horse, Funky prototype (Brenner et al. 2016)	**Process intents to deliver innovation:** - Approach: Each process and the phase result in intermediate outcomes and final deliverables (Brenner et al. 2016; Kelley and Kelley 2013). - Example: Design Space Exploration aims to explore user needs and results in opportunity areas (Brenner et al. 2016)	**Phase depending measurement approach:** - Approach: Measurement based on activities and outcomes of process/phase. - Example metrics for Design Space Exploration phase: number of stakeholder interactions, number of opportunity areas generated, time investment per identified need
Path 3 Transformation	**A mindset** - Approach: DT is understood as a mindset fostering certain values in organisations (Simon 1969; Björklund et al. 2020). - Example Values: Human-centeredness, customer-centricity, experimentation, diversity (Carlgren et al. 2016; Brenner et al. 2016; Micheli et al. 2019)	**Transformation of organisation:** - Approach: Deploying DT drives the transformation process in a firm fostering the values of DT (Björklund et al. 2020). - Example: Intensive interaction with customers as part of DT drives value of customer-centricity	**Transformation goal depending measurement approach:** - Approach: Measurement based on transformation goal and related outcomes and activities. - Example metrics for value customer-centricity: Share of employees qualified in Design Thinking, Amount of Customer Interactions per employee

Fig. 5 Framework for measuring design thinking

Path I: Method

The first path understands DT as a collection of methods, e.g. interviews, personas, journey maps, prototypes. These methods can be applied as part of a process but also independently. If the latter happens, it can be to solve a specific problem. For example, looking at the interview technique used in the context of DT. Interviews are used for instance, in the early phases as an instrument of understanding user needs (Uebernickel and Brenner 2020). This particular objective related to such a method can be measured by for example counting its number and calculating ratios such as the number of interviews per need. Thus, when using such a method on its own, it is possible to come up with different metrics that can be seen as a proxy for a team performing this specific method. Beyond the number of interviews, one could measure the effectiveness of a team running interviews by looking at how many interviews they could gather per identified need. A high number of interviews that explicitly mentioned the need could also signal its importance. By looking at specific methods isolated from their organisational context, it appears feasible to derive specific metrics to evaluate such a method's quality by looking at the intention of the subject. Thus, analysing each method's specific objective and generating metrics around this method might help to develop a measurement approach for practitioners that use DT methods disconnected from any given process model.

Path II: Process

The second path conceptualises DT as a process that leads to innovative solutions and prototypes. In our interviews, we found that many DIUs used methods that follow different phases of innovation. For each phase, one could derive objectives that can be measured on an outcome basis (What is the result of the stage?) and the process itself (How are the activities of the stage executed?). While the latter would relate to efficiency measurements, the outcome basis measurement approach would reflect effectiveness. Relying on such phases does not only address challenge III, which was about defining clear responsibilities that allow teams to gain responsibility, it also makes measurement easier to execute. Looking, for example, at the introduced process model and the first Design Space Exploration phase, one could develop the following metrics. As the Design Space Exploration phase is targeted to explore the problem space, one could measure efficiency through the number of stakeholder interactions (e.g. users, experts), assuming that more stakeholder interactions lead to a more sophisticated understanding of the problem space. Furthermore, to address the phase's effectiveness, one could assess critical artifacts that are part of the agreed deliverables. While such a "measure" is qualitative in its characteristics, in such early phases, deploying quantitative measures might result in the negative effects of performance measurement mentioned in the previous chapter.

Beyond the challenge of developing suitable metrics for such early phases, a second topic related to the organisational setup assessing "performance" in such contexts seems necessary. Evaluating such exploratory endeavours requires different skills that are necessary for exploitive projects. For example, while managers can apply financial metrics such as cost savings in exploitive projects with low uncertainty, using such metrics in uncertain environments might lead to unreliable estimations. Thus, setting up the right decision-making structures and selecting the right individuals assessing these tasks is another crucial element that emerged from our data.

Path III: Transformation

The third part contains DT as a mindset. In our sample cases, companies implementing DT often aimed to embed the principles in the organisation to drive (digital) transformation endeavours. Doing so, and adopting DT for this purpose brings with it two important implications.

First, DT can only be assessed as part of the bigger picture, that requires a holistic management including steering such organisational transformation efforts. Furthermore, one should be aware that such transformation efforts require long-term dedication while a short-term financial impact is rare.

Second, DT in such contexts seems to be often used for building capabilities (Carlgren et al., 2014). To give an example, one could consider a case where a company aims to drive customer-centricity. In terms of measurement, one could measure and identify practices that contribute to this goal. If running DT introduction workshops is one of the practices that strengthen the employee's skills in sensing user and market needs, one could measure the share of employees that have been qualified or that took part in such workshops. These metrics are the first idea for measuring such transformative endeavours. Deconstructing the transformation goals into specific practices and developing metrics to steer and measure these practices would also likely be beneficial. The role of DT in this process is a rich area for future research.

6 Discussion

Our findings provide insights on all three research questions raised. We started our investigation of existing research streams by conducting a literature review on the measurement of DT and innovation (RQ1). This resulted in three streams of research conceptualising DT measurement in different contexts. Following one stream and, specifically, with the gap of missing performance measurement instruments to steer DT in organisations, we conducted 20 exploratory interviews to identify challenges in measuring DT in organisations (RQ2). Building on these challenges and our insights from the literature, we propose a framework that points towards three paths of developing measurement approaches depending on the conceptualization of DT (RQ3).

While answering our research questions, we came across multiple topics to discuss from which we will focus on the two of the most prominent ones based on our data.

Measuring Design Thinking through Innovation Metrics

An interesting topic to discuss is measuring DT through metrics from the innovation literature. As introduced through our literature review results, existing innovation management literature has done much work on performance measurement (Adams et al., 2006; Dziallas & Blind, 2019). While substantial parts focus on New Product Development in a non-digital world, literature seems to have two blank spots. First, one in the area of performance measurement in the context of digital innovation (Hund et al., 2019) and second, in the very early innovation activities—the so-called fuzzy front end (Dziallas & Blind, 2019). In both areas, DT is often combined with other agile methods (Dobrigkeit, de Paula, & Carroll, 2020; Berghaus and Back, 2017). The shift of the activities of companies initiated through digital transformation expressed in methods such as DT seems to acquire a new set of measurement tools that can respond to more flexible and faster processes (Hund et al., 2019). This might explain why several metrics on innovation exist, but metrics on DT are still scarce.

Measuring Design Thinking as part of companies transformation efforts

The latter point connects well with another topic that emerged mainly in our interviews with practitioners. Often, DT is used as part of a company's transformation endeavours. This seems reasonable as many of the attributes assigned to DT, e.g. customer-centricity, speed (Brenner et al., 2016; Micheli et al., 2019) are relevant for many companies in their intention to pursue the transformation of the own organisation (Vial, 2019). While much research is available on maturity models measuring companies' digital transformation activities (Kontić & Vidicki, 2018), how to measure and steer such transformation efforts seems to be less explored.

7 Conclusion

This book chapter aimed to investigate the topic of performance measurement in the context of DT. By relying on a literature review, we described different understandings of these two terms. Following one particular stream and subsequently, a specific research gap of using performance measurement instruments to measure and steer DT in organisations, we conducted exploratory interviews to understand what challenges practitioners experience in measuring DT in a given context. The challenges revealed that DT often lacks a clear intention, making it challenging to develop suitable measurement approaches. Addressing this topic, we propose a framework for measuring DT in organisations and describe the first elements of how performance can be measured following these different conceptualisations.

Of course, our research has limitations. Developing performance measurement instruments in the introduced contexts where DT is applied (Digital Transformation, Innovation) first requires the insight that both contexts are complex and driven by uncertainty. This makes measurement particularly difficult. Thus, rather than proposing conclusions on this topic our work is exploratory by nature and aims to provide some coherent terms and concepts when discussing it. This leads to the limitation that our findings should be understood as preliminary. Secondly, our sample stems from Germany and Switzerland and has investigated DT in the particular context of DIUs. Thus, the generalizability of our findings may offer rich potential for further studies.

Future research on this topic could understand these limitations as starting points. Furthermore, adopting design science research approaches that aim to develop specific artefacts along our paths could be an interesting avenue in exploring the role and measurement in the particular contexts more deeply. Doing so can help practitioners and scholars to measure, understand and steer DT for innovation and transformation more efficiently and effectively. Additionally, exploring how organisations measure and steer their digital transformation efforts including the use of DT in this context might be helpful to further develop our understanding of Performance Measurement of DT.

Acknowledgements The results shown in this chapter are the outcome of a team effort. In particular we need to mention Falk Uebernickel and Danielly de Paula as part of the core team for intense discussions on how to measure DT in organisations. Furthermore, we would like to thank Tiziano Sabbioni and Céline Stalder for writing their master thesis as part of this project. Additionally, we appreciate the collaboration with the HPDTPRP project Design Thinking Scorecard, namely Selina Mayer with whom we were able to publish a conference paper for the HICSS 21 which is an important element of this chapter. In writing this chapter, we would like to thank Sharon Nemeth for her language support as well. Last but not least, we thank Prof. Hasso Plattner and the HPI-Stanford Design Thinking Research Program Committee for their support.

References

Adams, R., Bessant, J., & Phelps, R. (2006). Innovation management measurement: A review. *International Journal of Management Reviews, 8*(1), 21–47.

Barthel, P., Fuchs, C., Birner, B., & Hess, T. (2020). Embedding Digital Innovations in Organizations: A Typology for Digital Innovation Units. *Proceedings of the 15th International Conference on Wirtschaftsinformatik*, 780–795. Presented at the 15th International Conference on Wirtschaftsinformatik , Potsdam.

Bedford, D. S., Bisbe, J., & Sweeney, B. (2019). Performance measurement systems as generators of cognitive conflict in ambidextrous firms. *Accounting, Organizations and Society, 72*, 21–37.

Berghaus, S., & Back, A. (2017, December 10). Disentangling the Fuzzy Front End of Digital Transformation: Activities and Approaches. *Proceedings of the 38th International Conference on Information Systems*, 1–17. Presented at the International Conference on Information Systems (ICIS) 2017, Seoul, South Korea. Association for Information Systems.

Björklund, T. A., Hannukainen, P., & Manninen, T. (2018). Measuring the impact of design, service design and design thinking in organizations on different maturity levels. In *ServDes2018. Service design proof of concept, proceedings of the ServDes. 2018 conference, 18–20 June, Milano, Italy*, pp. 500–511. Linköping University Electronic Press.

Björklund, T., Maula, H., Soule, S. A., & Maula, J. (2020). Integrating design into organizations: The coevolution of design capabilities. *California Management Review, 62*(2), 100–124.

Brenner, W., Uebernickel, F., & Abrell, T. (2016). Design thinking as mindset, process, and toolbox. In *Design Thinking for Innovation* (pp. 3–21).

Brown, B. (2018). The Total Economic Impact™ of IBM's Design Thinking Practice. *Forrester Consulting: Massachusetts, MA, USA,* 1–45.

Brown, T. (2008). Design thinking. *Harvard Business Review, 86*(6), 84–92, 141.

Carlgren, L., Elmquist, M., & Rauth, I. (2014). Design thinking: Exploring values and effects from an innovation capability perspective. *The Design Journal, 17*(3), 403–423.

Carlgren, L., Rauth, I., & Elmquist, M. (2016). Framing design thinking: The concept in idea and enactment: Creativity and innovation management. *Creativity and Innovation Management, 25*(1), 38–57.

Chandler, A. D., Jr. (1977). *The visible hand: The managerial revolution in American business.* Harvard University Press.

Chenhall, R. H., & Moers, F. (2015). The role of innovation in the evolution of management accounting and its integration into management control. *Accounting, Organizations and Society, 47*, 1–13.

Chiesa, V., Frattini, F., Lamberti, L., & Noci, G. (2009). Exploring management control in radical innovation projects. *European Journal of Innovation Management, 12*(4), 416–443.

Davila, A., Foster, G., & Oyon, D. (2009). Accounting and control, entrepreneurship and innovation: Venturing into new research opportunities. *European Accounting Review, 18*(2), 281–311.

Dewangan, V., & Godse, M. (2014). Towards a holistic enterprise innovation performance measurement system. *Technovation, 34*(9), 536–545.
Dobrigkeit, F., de Paula, D., & Carroll, N. (2020, November). InnoDev Workshop: A One Day Introduction to Combining Design Thinking, Lean Startup and Agile Software Development. *Conference Proceedings*, 1–10. Presented at the IEEE 32nd Conference on Software Engineering Education and Training.
Dziallas, M., & Blind, K. (2019). Innovation indicators throughout the innovation process: An extensive literature analysis. *Technovation, 80–81*, 3–29.
Edmondson, A. C., & Mcmanus, S. E. (2007). Methodological fit in management field research. *AMRO, 32*(4), 1246–1264.
Eriksson, P., & Kovalainen, A. (2015). *Qualitative methods in business research: A practical guide to social research.* SAGE.
Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods, 16*, 15–31. https://doi.org/10.1177/1094428112452151.
Godener, A., & Soderquist, K. E. (2004). Use and impact of performance measurement results in R&D and NPD: An exploratory study. *R and D Management, 34*(2), 191–219.
Hagedoorn, J., & Cloodt, M. (2003). Measuring innovative performance: Is there an advantage in using multiple indicators? *Research Policy, 32*(8), 1365–1379.
Hennessey, B. A., & Amabile, T. M. (2010). Creativity. *Annual Review of Psychology, 61*, 569–598.
Holotiuk, F., & Beimborn, D. (2019). Temporal Ambidexterity: How Digital Innovation Labs Connect Exploration and Exploitation for Digital Innovation. *Proceedings of the 40th International Conference on Information Systems*. Presented at the International Conference on Information Systems, Munich, Germany.
Hund, A., Drechsler, K., & Reibenspiess, V. A. (2019). The current state and future opportunities of digital innovation: A literature review. *Proceedings of the 27th European Conference on Information Systems*. Presented at the 27th European Conference on Information Systems, Stockholm & Uppsala, Sweden, June 8–14.
Johansson-Sköldberg, U., Woodilla, J., & Çetinkaya, M. (2013). Design thinking: Past, present and possible futures. *Creativity and Innovation Management, 22*(2), 121–146.
Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard-measures that drive performance. *Harvard Business Review, 70*(1), 71–79.
Kelley, T., & Kelley, D. (2013). *Creative confidence: Unleashing the creative potential within us all*. Crown.
Kolko, J. (2015). Design thinking comes of age. *Harvard Business Review*. Retrieved from https://enterprisersproject.com/sites/default/files/design_thinking_comes_of_age.pdf.
Kontić, L., & Vidicki, Đ. (2018). Strategy for digital organization: Testing a measurement tool for digital transformation. *Strategic Management, 23*(2), 29–35.
Lewrick, M., Link, P., & Leifer, L. (Eds.). (2018). *Das design thinking playbook*. (2nd ed.). Vahlen, Franz.
Liedtka, J. (2015). Perspective: Linking design thinking with innovation outcomes through cognitive bias reduction. *Journal of Product Innovation Management, 32*(6), 925–938.
Liedtka, J. (2017). Evaluating the Impact of Design Thinking in Action. *Proceedings: AMIA Annual Fall Symposium,* 10264. Presented at the AMIA Fall Symposium 2017 (1): 10264. Academy of Management.
Liedtka, J. (2020). Putting technology in its place: Design thinking's social technology at work. *California Management Review, 62*(2), 53–83.

Mayer, S., Haskamp, T., & de Paula, D. (2020, December 31). Measuring what Counts: An Exploratory Study about the Key Challenges of Measuring Design Thinking Activities in Digital Innovation Units.*Proceedings of the 54th Hawaii International Conference on System Sciences*, 4951–4960. Presented at the Hawaii International Conference on System Sciences (HICSS) 2021, Virtual.

Micheli, P., & Mari, L. (2014). The theory and practice of performance measurement. *Management Accounting Research, 25*(2), 147–156.

Micheli, P., Wilner, S. J. S., Bhatti, S. H., Mura, M., & Beverland, M. B. (2019). Doing design thinking: Conceptual review, synthesis, and research agenda. *Journal of Product Innovation Management, 36*(2), 124–148.

Nagaraj, V., Berente, N., Lyytinen, K., & Gaskin, J. (2020). Team design thinking, product innovativeness, and the moderating role of problem unfamiliarity. *Journal of Product Innovation Management, 10*, 123.

Nakata, C., & Hwang, J. (2020). Design thinking for innovation: Composition, consequence, and contingency. *Journal of Business Research, 118*, 117–128.

Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M., & Kennerley, M. (2000). Performance measurement system design: Developing and testing a process-based approach. *International Journal of Operations & Production Management, 20*(10), 1119–1145.

Rauth, I., Carlgren, L., & Elmquist, M. (2014). Making it happen: Legitimizing design thinking in large organizations. *Design Management Journal, 9*(1), 47–60.

Roth, K., Globocnik, D., Rau, C., & Neyer, A.-K. (2020). Living up to the expectations: The effect of design thinking on project success. *Creativity and Innovation Management, 62*, 100.

Royalty, A., Chen, H., Roth, B., & Sheppard, S. (2019). Measuring design thinking practice in context. In C. Meinel & L. Leifer (Eds.), *Design thinking research: Looking further: Design thinking beyond solution-fixation.* (pp. 61–73). Springer International Publishing.

Royalty, A., & Roth, B. (2016a). Developing design thinking metrics as a driver of creative innovation. In H. Plattner, C. Meinel, & L. Leifer (Eds.), *Design thinking research: Making design thinking foundational.* (pp. 171–183). Springer International Publishing.

Royalty, A., & Roth, B. (2016b). Mapping and measuring applications of design thinking in organizations. In H. Plattner, C. Meinel, & L. Leifer (Eds.), *Design thinking research: Taking breakthrough innovation home.* (pp. 35–47). Springer International Publishing.

Schmiedgen, J., Spille, L., Köppen, E., Rhinow, H., & Meinel, C. (2016). Measuring the impact of design thinking. In H. Plattner, C. Meinel, & L. Leifer (Eds.), *Design thinking research: Making design thinking foundational.* (pp. 157–170). Springer International Publishing.

Seidel, V. P., & Fixson, S. K. (2013). Adopting design thinking in novice multidisciplinary teams: The application and limits of design methods and reflexive practices. *Journal of Product Innovation.*

Sheppard, B., Sarrazin, H., Kouyoumjian, G., & Dore, F. (2018). *The business value of design.* Retrieved from McKinsey Quarterly website: https://www.mckinsey.com/business-functions/mckinsey-design/our-insights/the-business-value-of-design.

Svahn, F., Mathiassen, L., & Lindgren, R. (2017). Embracing digital innovation in incumbent firms: How Volvo cars managed competing concerns. *The Mississippi Quarterly, 41*(1).

Tangen, S. (2005). Demystifying productivity and performance. *International Journal of Productivity and Performance Management, 54*(1), 34–46.

Uebernickel, F., & Brenner, W. (2020). *Design thinking: The handbook*. World Scientific Publishing Company Pte Limited.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems, 28*(2), 118–144.

Vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A., & Others. (2009). Reconstructing the giant: on the importance of rigour in documenting the literature search process. ECIS 2009, 9, 2206–2217.

Westcott, M., Sato, S., Mrazek, D., Wallace, R., Vanka, S., Bilson, C., & Hardin, D. (2013). The DMI design value scorecard: A new design measurement and management model. *Design Management Review, 24*(4), 10–16.

Wiesche, M., Jurisch, M. C., Yetton, P. W., & Krcmar, H. (2017). Grounded theory methodology in information systems research. *The Mississippi Quarterly, 41*(3), 685–701.

Wrigley, C., Nusem, E., & Straker, K. (2020). Implementing design thinking: Understanding organizational conditions. *California Management Review, 62*(2), 125–143.

Ylinen, M., & Gullkvist, B. (2014). The effects of organic and mechanistic control in exploratory and exploitative innovations. *Management Accounting Research, 25*(1), 93–112.

Measuring the Impact of Project-Based Design Engineering Courses on Entrepreneurial Interests and Intentions of Alumni

Sheri D. Sheppard, Helen L. Chen, George Toye, Felix Kempf, and Nada Elfiki

Abstract As the field of engineering design has grown, educators have offered students experiential opportunities to engage in industry-sponsored projects that emphasize innovation and creativity and highlight entrepreneurial pathways. At Stanford University, alumni of these courses have gone on to engage in a range of professional endeavors, representing a variety of engineering functions and organizational roles. Much of the evidence of how these curricular efforts contribute to entrepreneurial interests lie primarily in anecdotal examples and stories about alumni career trajectories. Given this gap, this chapter describes the process of developing and implementing a survey instrument focused on gathering feedback and insights from course alumni. The findings inform a more nuanced understanding of the relationship between intensive project-based design experiences and alumni intentions and interests in entrepreneurial outcomes.

1 Introduction

For over half a century, Stanford University's *ME310: Project-Based Engineering Design Innovation & Developmen*t course sequence has engaged Stanford graduate students in industry-sponsored projects where they learn to navigate various phases of integrated design thinking through engineering fabrication. The core tenet of the educational pedagogy of ME310 is the emphasis on the design thinking process in application, implementation, as well as documentation. This is operationalized in the curriculum which emphasizes connections with industry through sponsored projects as well as global collaboration with an international academic partner.

S. D. Sheppard · H. L. Chen (✉) · G. Toye · F. Kempf · N. Elfiki
Department of Mechanical Engineering, 475 Via Ortega, Stanford, CA 94305, USA
e-mail: hlchen@stanford.edu

S. D. Sheppard
e-mail: sheppard@stanford.edu

G. Toye
e-mail: toye@stanford.edu

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_16

Carleton and Leifer (2009) and Carleton (2019) recorded the history of ME310 at Stanford and show how its development as a course has paralleled the evolution of the field of engineering design. In recent decades, in synchrony with the growing demands of today's design engineering professionals, ME310 projects have advanced an innovation focus both in process and outcome. Correspondingly, entrepreneurship education in engineering has been increasingly viewed as an important approach to strengthening the field's ties to innovation (Byers et al., 2013; Gilmartin et al., 2019; Jin et al., 2015). ME310 alumni are helping students navigate the chasm between the academic course setting and the environment of practicing professionals in industry. Many alumni maintain their affiliation with the course by serving as dedicated project advisors from beginning to end, giving guest lectures, providing feedback to students at course milestone events, and becoming lifelong resources to the growing ME310 social network.

The impact of this emphasis on innovation and creativity, as demonstrated in the prototypes created by the ME310 students, suggests potential opportunities and pathways for entrepreneurship. A survey administered by Eesley and Miller (2012) to Stanford University alumni, faculty, and selected staff in 2011 estimated that almost 40,000 active companies are associated with Stanford, suggesting an approximate annual revenue of $2.7 trillion and 5.4 million employees. Stanford's role in creating an ecosystem of entrepreneurship and creativity is well-documented and given that education is the primary mission of the university, this culture of generating unique solutions to meet market needs and enterprise creation can be viewed as important outcomes of the environment of design thinking oriented courses such as ME310.

The expansion of the ME310 format and the pedagogical approach to courses and projects in other countries and contexts, starting in 2004, has also resulted in many examples and anecdotal stories about successful student teams and alumni who have gone on to leverage their ME310 prototypes into commercial products (Kubota, 2018). However, the impact of these intensive design experiences has not been studied from the perspective of the intentions and interests in entrepreneurial outcomes. Entrepreneurial intent is defined as a "state of mind that directs attention, experience, and action toward a business concept, [setting] the form and direction of organizations at their inception" (Ajzen, 2002). We see "entrepreneurial intent" as a measure of an individual's entrepreneurial interests, given that entrepreneurial intent is predictive of future entrepreneurial behaviors (Ajzen, 2002; Krueger & Carsrud, 1993). Our survey-based study of ME310 alumni allows us to investigate how the ME310 course may have contributed to realization of entrepreneurial intent, through graduates' involvement in and creation of new endeavors.

1.1 Research Questions

By examining the relationship between the ME310 curriculum and alumni engagement in entrepreneurship and innovation, we aim to develop a fuller picture of how course-based training in design thinking can lead to a variety of professional endeavors (including starting new companies). The research objective of this study is to assess the impact of intensive course-based design experiences such as ME310 on the intentions and interests in entrepreneurial outcomes of alumni.

We accomplish this goal through a survey instrument designed to address three main research questions:

1. *What career paths have ME310 students pursued since graduating from Stanford?*
2. *What are the alumni's attitudes and perspectives on the various components of the ME310 curriculum?*
3. *What are the current levels of innovation and entrepreneurial self-efficacy of ME310 alumni?*

This chapter describes the process of designing this survey instrument and identifying insights that can be generalized to include other engineering design experiences. We seek to better understand how design-based education relates to and enables a variety of futures for its graduates.

2 Setting the Context: Stanford University's ME310

Exploration of the long-term impact on alumni of any academic course requires a solid understanding of both the course's background and situational context. This is especially true for ME310 with its extensive history of over fifty years at Stanford University.

Programmatically, the ME310 A/B/C 3-quarter course sequence is a graduate level, academic year-long, multi-disciplinary, project-based learning, design engineering, student experience in mechanical engineering that is grounded in real-world corporate/industry interests and opportunities. The breadth of content in ME310's project portfolio extends across a multiplicity of intersecting technical domains (aeronautical, astronomical, chemical, civil, electrical, mechanical, biological, material science, etc.) that are also inclusive of important human psychology, economics, and business model considerations. ME310 is one of a few focused depth sequence options that is required for the Master's degree in mechanical engineering design. Students are organized into teams, and each team then works together on its own unique project. This year-long project engagement offers opportunities for design explorations and iterations of greater depth and breadth than with shorter quarter/semester length courses.

2.1 History of ME310

In the first half of ME310's history, the projects mainly focused on the design of mechanical systems that were internally meaningful to corporate sponsors and were relatively well defined (e.g., test and manufacturing; methods, physical mechanisms, tools, or systems). In 1992, the new teaching team of Professors Larry Leifer and Mark Cutkosky introduced a greater emphasis on design methodologies and supporting technologies to ME310.

This shift resulted in projects that were more open-ended and incorporated more human-centered design thinking perspectives. Coincidentally and influentially, a new Defense Advanced Research Project Agency (DARPA) research project was being undertaken by these course instructors to explore how design and manufacturing processes might advance with the development of the commercial Internet (aka the "information superhighway"). This seeded interest in how computer electronic communications technologies could facilitate and enhance team collaboration, both locally (in-person) and remotely (globally distant). ME310 became the opportunistic research testbed for this research project.

Under this DARPA research program, new computer software tools for distributed collaboration were developed. To experiment and test these tools, ME310 installed the first academic course worldwide Web site, connected to Stanford's campus Internet network and accessible worldwide 24 × 7. In the intervening quarter century, ME310 has continued to advance (and experiment with) the use of computer- and Internet-based technologies to support students' design work—from establishing a shared computer cluster in the dedicated course loft workspace in the mid-1990s, where students could do much of their joint teamwork, to distributing network capable portable laptop computers to teams, and further on to today's reality where every student has in their smartphone a portable hand-sized computer with screen, touch keyboard, extensible digital storage, high resolution photo/video capture, text, email and real-time conferencing capabilities via ubiquitous worldwide Internet connectivity.

In 2003, DARPA research and international collaborations showed that the engineering design profession was becoming increasingly global. Engineering design teams from around the world could actively collaborate across time zones; design work could continue around the clock with handoffs to work shifts located elsewhere in the world. To prepare students to work globally, ME310 extended its initial local student team model to include close collaborations with students from global academic partners (i.e., international universities). This global team arrangement was phased in over four years, with increases in the number of academic partners until every project team would be paired with a global team. To build effective collaborative relationships and heighten cultural awareness, students would travel internationally to meet their global teammates at the beginning, middle and end of the project.

Given the interactive nature of team-based design processes, the course structure around global partnerships evolved to intentionally acknowledge and foster the social dynamics of design teams. Student team members often worked on their own dedicated projects with few interactions across project teams. The good or poor quality of social relationships among teammates were seen to be predictive of whether the team's design project performance would be boosted or hindered. New activities outside of normal class times were designed to enhance the social connectivity between local teammates and classmates. For example, a social gathering to partake in food and drink together as a community after class every Thursday evening was introduced and became a tradition called "Slightly Unorganized Design Sessions" (SUDS). This regular weekly ritual is still practiced at Stanford, and at the global academic partner universities, though not in real-time due to time zone differences.

In summary, the design of our study of ME310 alumni needed to account for curricular and societal changes over time as well as student body and generational changes, especially given the changing profile of graduate students admitted to Stanford's Mechanical Engineering graduate program in recent decades. International student enrollment in ME310 was rare in the course's earlier years but has risen and become more prominent in recent years. For example, in 2017 almost half of the students completing ME310C were international students creating a challenge as well as an opportunity to incorporate more diverse backgrounds and perspectives. Over the decades, the student cohorts in ME310 have transitioned generationally as represented by the research on *Baby Boomers, Gen X, Gen Y/Millennials*, and now *Gen Z* (Pew Research Center, 2015). Members of these social generations respond to the course goals differently, and often have contrasting personal and professional aims (Bialik & Fry, 2019; LeDuc, 2019). Finally, the survey also had to address economic conditions such as the bursting of the 2001 dot-com bubble and the 2008–2009 recession, which are likely to have influenced students' mindsets and priorities when they took the course. We expected students to have different takeaways and perspectives from their course experiences that would likely influence their subsequent academic or professional job pursuits.

3 Survey Design and Deployment

While surveys of graduates of American colleges and universities date back to the 1930s (Ewell, 2005), the growth of the field of institutional research in the 1960s (Olsen, 2000) and subsequently, the easy availability of online survey software in the 2010s, all have contributed to the use of this method for data collection (Chen et al., 2012). One unique aspect of the current study is the focus on engineering graduates' relationship with a specific course rather than the broader program, department, or institution.

The ME310 course framework makes this study of alumni particularly interesting, although not without challenges. These challenges range from the pragmatic (i.e., securing approval from the Stanford Institutional Review Board and getting access to alumni contact information), to the realistic (i.e., creating a survey of reasonable length does not permit a high level of granularity in capturing every job change), to the perplexing (i.e., ME310 has evolved over time—how do we capture this in a single survey? How do we trigger a respondent's memories without unduly influencing these memories?), to the creative (i.e., how do we incentivize graduates to complete the survey?).

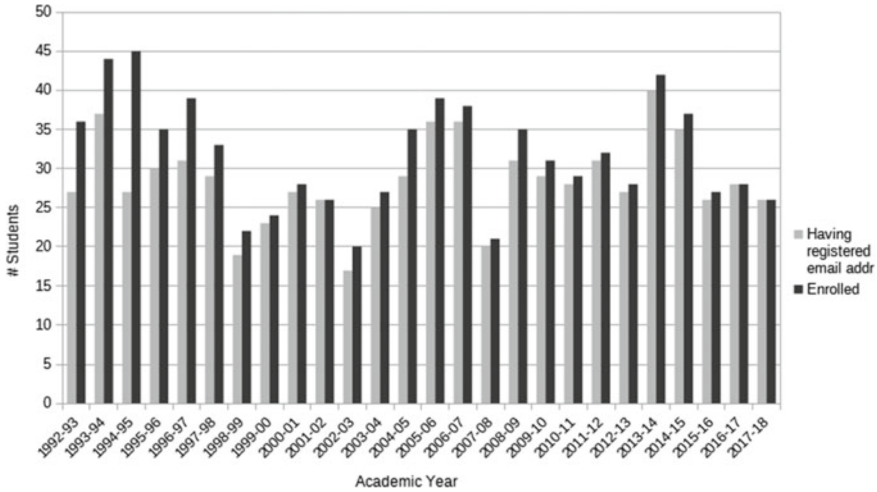


Fig. 1 Mapping of ME310 enrollment from 1992–93 to 2017–18

3.1 Selection of ME310 Alumni Respondents

These challenges led us to consider whether we would aim for our survey going out to five decades of alumni, or to a subset of that population. In the end, we decided to focus on course graduates from the class of 1992–93 through the class of 2016–17. The main reason for this is that circa 1992 was a notable pivotal inflection point in ME310's course history in terms of the type of projects being undertaken, the course putting greater emphasis on design methodologies (and underlying support tools), and a growing course focus on the global and social nature of modern design and design teams.

In order to determine the names and email addresses of the ME310 alumni members over the past 25 years, we worked internally with the Alumni Relations and Student engagement arm of the Stanford School of Engineering as well as the university's institutional research office. A list of around 800 ME310 alumni who completed the year-long ME310 A/B/C sequence in its entirety was generated. Figure 1 provides a more detailed description of how the enrollment of ME310 students has evolved during our study period.

3.2 Survey Outreach and Recruitment

In order to combat survey fatigue while maximizing the survey response rate and response quality, careful attention was given to the email invitations to the alumni. Three mails were sent during the survey administration period during the last two weeks of July 2020: the initial invitation with survey link to all alumni, a first reminder to those who had not completed the survey about one week later, and a final reminder about the close of the survey again to the subset of non-completers two days before the announced survey deadline. (The actual survey was formally closed five days after the deadline.)

Each outreach email was carefully crafted to appeal to the graduates' altruism and willingness to give back to the Stanford community. Given the disruptions due to COVID-19, these emails emphasized how the input of the graduates would help the teaching team better understand the impact of the course. The perspectives of the alumni were critical since the course is at an important crossroads where new approaches are imperative in order to meet the needs of the next generation of design engineers.

A sense of nostalgia about the ME310 experience was fostered by including pictures of Professors Larry Leifer and Mark Cutkosky along with a personal message from the current teaching team in these communications. In addition, two incentives were provided in order to encourage participation. Every survey respondent was offered a commemorative ME310 pin and a chance to win a copy of the book *ME310 at Stanford University: 50 Years of Redesign (1967–2017)* (Carleton, 2019) as well as to receive notifications about future research invitations to participate in an interview and to learn about the ME310 research findings. A critical component of the ME310 experience is the community and, in addition to the research efforts, the survey respondents were invited to join a new LinkedIn group set up by the ME310 teaching team, designed to strengthen the connections among the alumni through professional networking and invitations to ME310 events and activities.

3.3 Designing the Survey Instrument

As a reminder, our study objective was to assess the impact of intensive course-based design experiences such as ME310 on the intentions and interests in entrepreneurial outcomes and behaviors of alumni. Furthermore, the goal of this research was to produce findings that were actionable and/or affirming to the ME310 teaching team while also creating a positive "survey user experience" that might contribute to continued expansion of the ME310 network.

To this end, we worked closely with the ME310 teaching team throughout the survey development process to understand the curricular goals, course strategies, milestones and activities. These components were integrated with outcomes and measures related to innovation and entrepreneurship in engineering education. The final survey instrument was organized around the themes associated with our three research questions: 1) education and career; 2) the ME310 experience; and 3) attitudes and perceptions of self-efficacy related to innovation and entrepreneurship.

RQ1: What career paths and plans have ME310 students pursued since graduating from Stanford? (31 survey items).

Our research team debated the level of granularity in asking our survey participants about their career paths. Those who had finished ME310 in 2017 might be in the same "post-graduation" job at the time of completing our survey, whereas a graduate from 1993 might have had several career changes, many different jobs, and might even be retired when they took the survey.

We asked participants about their first job after completing ME310, and their current or most recent employment status. The first job questions gave us a look at how job opportunities and prospects of ME310 graduates may (or may not) have changed over the 25 years due to industry changes (who could have imagined Tesla in 1993!) and curricular changes. The "current job" questions allowed us to begin to assess how varied the career paths may be for mechanical engineering graduates, particularly for those with 10 or more years of experience since ME310. In order to consistently define and characterize the fields, industry sectors, and roles, we used survey items that were developed for an earlier survey—the *Engineering Majors Survey*—which was based on a taxonomy of descriptors from the US Department of Labor (Gilmartin et al., 2017).

Following the completion of the survey, subsequent interviews with 39 survey participants led by co-author N. Elfiki provided an opportunity to probe more deeply into the career choices and influences along a career path, particularly as related to individuals with innovation and entrepreneurial interests and intentions.

RQ2: What are the alumni's attitudes and perspectives on the various components of the ME310 curriculum? (44 survey items).

The quality of the survey responses was dependent on alumni's recollections of their student course experience. Given the large span of years and the age of the oldest alumni in our study group, we assumed we would need to help respondents reconnect with those memories. ME310's pedagogical focus on engineering design processes emphasizes the sequence of team-based design strategies, and the practice of necessary skills for professional leadership. Students spend one to three weeks learning each new design strategy and building an accompanying design idea prototype for each milestone. We believed that incorporating course elements such as names and descriptions of key milestones (i.e., CFP, DarkHorse) and pictures of popular activities (i.e., Paper Bike) throughout the survey would not only help jog memories but also help alleviate survey fatigue and streamline the flow of the survey. Figure 2 illustrates the "cascade" of milestones and design strategies from the most recent offerings of ME310 that we were aiming to capture in a survey format.

Some of the ME310 survey questions were focused on recalling and assessing the impact of core design strategies and skills whereas other questions were focused on recalling specifics of their project. Because of the course's emphasis on the social dynamics of design, additional questions were related to the team- and classmate (locally and globally) interactions and performance, and the durability of teammate relationships.

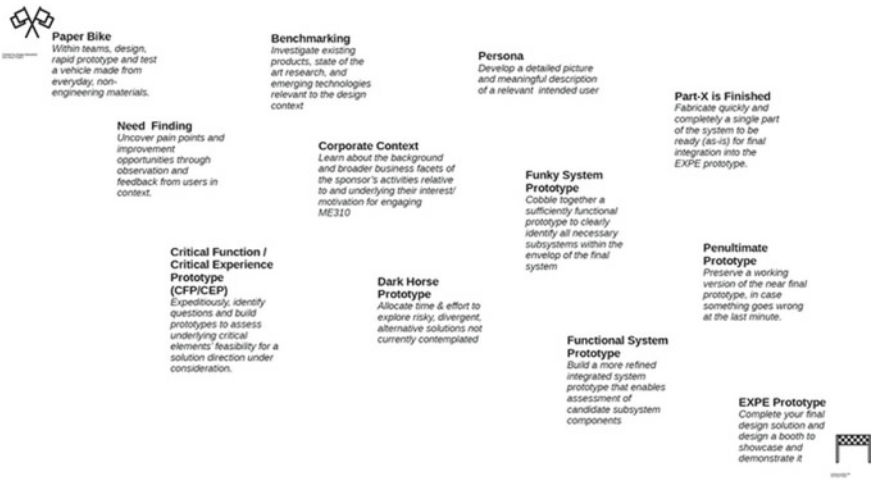


Fig. 2 ME310 milestones and design strategies (2016–2020)

Example question topics:

- What ME310 course elements do you remember from ME310? Have you found valuable in your life/career? (e.g., design process, public communications, collaboration, project and team management)
- What ME310 course elements do you remember from ME310? Have you found valuable in your life/career? (e.g., challenging assumptions, building quick prototypes, taking risks with radical design ideas)
- Have you worked on a project with a global partner team?
- Since ME310, how many connections have you maintained with (e.g., your project's Stanford teammates, your project's global teammates, etc.)?
- On a scale from 0–10, how likely are you to recommend ME310 to current Stanford students? [Net Promoter Score]

RQ3: What are the current attitudes and perspectives around self-efficacy related to innovation and entrepreneurship*?* (47 survey items).

Social Cognitive Career Theory (SCCT) was developed by Lent. Brown, and Hackett (1994) and later, Lent and Brown (2006) served as the theoretical framework for our engagement with engineering students' attitudes, interests, goals, and performance as related to innovation-related work and their "self-efficacy," defined as one's own beliefs in their ability to perform a specific task or action (Bandura, 1986). In prior work (Gilmartin et al., 2017), the SCCT model and the factors that influence students' decisions about academic and career intentions have highlighted self-efficacy measures focusing on innovation (Schar et al., 2017a, b), entrepreneurship (DeNoble, Jung, and Ehrlich, 1999), and design thinking (Schar, 2020), as detailed in Table 1.

Table 1 Descriptions of self-efficacy measures

Measure	Description
Innovation Self-Efficacy (ISE.5) [5 items]	Confidence in one's ability to innovate, i.e., to engage in specific behaviors that characterize innovative people. Measured on a five-point Likert scale ranging from "Not confident" (0) to "Extremely confident" (4)
Engineering Task Self-Efficacy [4 items]	Confidence in one's ability to perform integral technical engineering "tasks" such as "analyzing the operation or functional performance of a complete system." Measured on a five-point Likert scale ranging from "Not confident" (0) to "Extremely confident" (4)
Entrepreneurial Self-Efficacy (ETSE) [10 items]	Confidence in one's ability to pursue a new venture opportunity, representing two dimensions related to developing "new product and market opportunities" and "coping with unexpected challenges." Measured on a five-point Likert scale ranging from "Not confident" (0) to "Extremely confident" (4)
Design Thinking Self-Efficacy [6 items]	Confidence in one's abilities in the areas of empathy, reframing, ideation, prototyping, and testing. Measured on a continuous scale from "cannot do at all" (0), "moderately can do" (50), "highly certain can do" (100)

The survey also included demographic items asking about the year that the ME310 course was taken, gender, and undergraduate institution. Several open-ended questions ensured that respondents had a space to include any observations about their ME310 experience that were not explicitly asked about or general opinions that they felt would be useful for the researchers to know.

Since the survey was designed for such a diverse cohort of alumni, the survey was piloted extensively with members of the Designing Education Lab and other collaborators in order to ensure that the questions and the flow of the survey were comprehensible and meaningful, and the completion of the survey was within the desired time frame of 15–20 min to complete. The survey was iteratively improved and refined based on the feedback received.

4 Findings

4.1 Demographics of Survey Sample and Response Rate

In total, 301 student alumni participated in the online survey, yielding an overall response rate of 41%. As pointed out by Carleton and Leifer (2009), there are various reasons for why the total population of student alumni is not available to survey. The most important reasons include deceased alumni or lost contact with Stanford. In particular, since the survey was conducted online and was primarily distributed via email, we heavily relied on valid and up-to-date contact information as well as self-initiative from the recipients to fill out the survey.

Out of the 301 survey participants, 12% of survey participants did not complete the survey; a survey was considered complete if it was successfully submitted. Some 89% of the participants submitted their survey within the first hour after starting the survey, indicating that the majority of people were likely to have filled out the survey in a single pass. The median progress of unfinished surveys was 16% and only 5 participants who did not complete the survey answered more than 50% of the questions. Therefore, for the purposes of consistency and comparability, we only consider complete survey submissions in our analysis. Subsequently, the effective sample size under consideration is 266.

Nevertheless, a complete survey does not necessarily mean that the responses are free from missing data which occurs, for example, when survey participants progress through the survey while leaving some questions unanswered[1]; these are so-called missing data. Missing values in survey data always impose challenges and must be dealt with. As pointed out by (Brick & Kalton, 1996), there exists a considerable amount of research focusing on the question of how to deal with missing data in survey research, including methods of weighting and imputation. Due to the overall limited scope of missing data, we proceeded by replacing missing values with the median answer if the question was of numerical nature.

Our dataset consists of a unique and heterogeneous network of alumni who graduated from Stanford's ME310 module between 1993 and 2017. Despite the extraordinary course history and multi-generational survey audience, the 266 survey respondents were evenly distributed across year groups and clusters as visualized in Fig. 3. An even distribution of survey participants across year groups allowed us to gain robust and valuable insights into how ME310 was perceived over the years. The gender distribution showed that 77% of the survey participants are male and 22% are female. Less than 1% did not disclose their gender.

[1] In fact, 25% of all questions asked in the survey contain missing values. This may seem significant at first, however, the highest percentage of missing data per question across all questions is 3% and only occurs once. The mean of missing values across all other questions is 2%. We conclude that the magnitude of missing data in the dataset is within the expected boundaries of human error. Moreover, we find that the majority of missing data stems from only four survey participants. While the reason for why participants left certain questions unanswered (either intentionally or unintentionally) remains unknown, the overall scope of missing data in the dataset is in fact very limited, boosting the robustness of our statistical analysis.

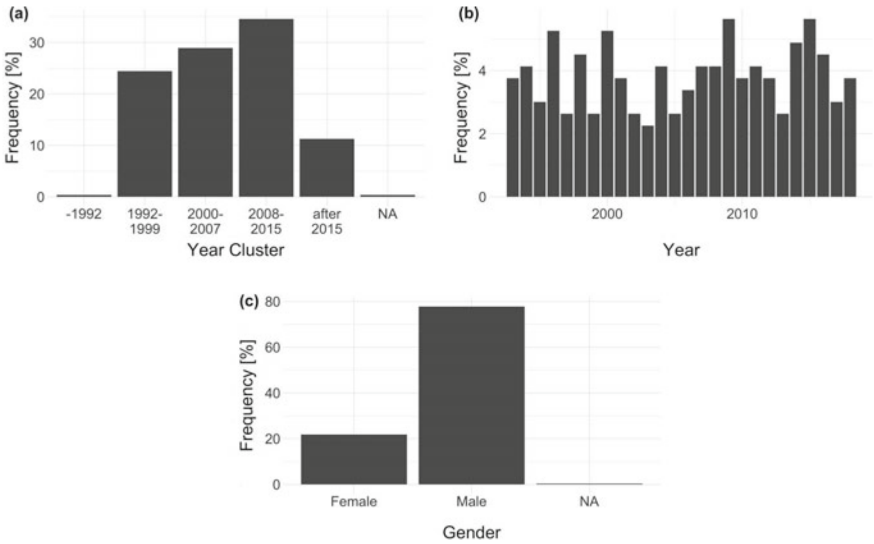


Fig. 3 Distribution of survey participants per year cluster **a** and year **b**, distribution of gender **c**

4.2 Preliminary Course Response and Self-Efficacy Results

The overall course assessment was overwhelmingly positive. Over 90% of the survey participants stated that they either loved or at least rated their overall ME310 experience positive in connection with their Stanford teammates and/or global partners. Subsequently, the data highlighted an extraordinary course appraisal. While 86% described the course as challenging, over 73% of the surveyed alumni also described their experience as *fun, engaging* or *useful*. Only 2% associated negative attributes with their ME310 experience such as *boring* or *forgettable*. Moreover, despite the technical content of the course, a crucial aspect of ME310 was the social interactions with project team members or classmates in general. We find that an astonishing 92% of the survey participants are still in contact with at least one classmate, and over 70% still have contact with at least one project team member, underlying the extraordinary and lifelong bonding element of ME310.

One of the central goals of the survey was to gain a better understanding of ME310 alumni's current self-assessments on innovation self-efficacy (ISE), engineering task self-efficacy (ETSE) and entrepreneurial self-efficacy (ESE) and design thinking self-efficacy (DTSE). Our approach of investigating all four self-efficacy measures establishes a multifaceted profile of the ME310 graduates. For each self-efficacy measure, a Cronbach's alpha was calculated to evaluate the internal reliability of the responses.

In Table 2, the overall mean of the Innovation Self-Efficacy scale is 3.00 (0.40) with a Cronbach's alpha of 0.78. Interestingly, the item concerning building a large network of contacts had the greatest variation in responses as well as the lowest

Table 2 Innovation self-efficacy (ISE)

Survey questions	Cronbach's alpha/factor loadings (Total sample n = 266)	Mean (SD)
	α = 0.72	3.00 (0.40)
1—Ask a lot of questions	0.67	3.34 (0.73)
2—Generate new ideas by observing the world	0.61	3.11 (0.86)
3—Experiment as a way to understand how things work	0.68	3.18 (0.75)
4—Build a large network of contacts with whom you can interact to get ideas for new products or services	0.71	2.30 (1.08)
5—Connect concepts and ideas that appear, at first glance, to be unconnected	0.67	2.97 (0.82)

average score of confidence. In contrast, participants' confidence regarding asking a lot of questions displayed a particularly strong confidence score with the lowest variance.

Table 3 describes the engineering task self-efficacy measure with a Cronbach's alpha of 0.85 and an overall mean and standard deviation of 2.90 (0.12). Respondents reported a high level of self-confidence with regard to designing a new product or project that meets specific requirements.

Cronbach's alpha for the Entrepreneurial Self-Efficacy items is 0.87 as noted in Table 4. Interestingly, questions that concern new market opportunities for a product, bringing product concepts to market, or what the business will look like displayed the highest variation in survey responses.

Table 3 Engineering task self-efficacy (ETSE)

Survey questions	Cronbach's alpha/factor loadings (Total sample n = 266)	Mean (SD)
	α = 0.85	2.90 (0.12)
1—Design a new product or project to meet specified requirements	0.82	3.11 (0.84)
2—Conduct experiments, build prototypes, or construct mathematical models to develop or evaluate a design	0.84	2.99 (0.91)
3—Develop and integrate component sub-systems to build a complete system or product	0.81	2.79 (1.00)
4—Analyze the operation or functional performance of a complete system	0.82	2.87 (0.85)
5—Troubleshoot a failure of a technical component or system	0.83	2.90 (0.92)

Table 4 Entrepreneurial self-efficacy (ESE)

Survey questions	Cronbach's alpha/factor loadings (Total sample n = 266)	Mean (SD)
	α = 0.87	2.70 (0.32)
1—See new market opportunities for new products and services	0.85	2.32 (1.04)
2—Design products that solve current problems	0.86	2.81 (0.86)
3—Discover new ways to improve existing products	0.86	2.98 (0.80)
4—Create products that fulfill customers' unmet needs	0.86	2.71 (0.91)
5—Identify new areas for potential growth	0.85	2.55 (0.96)
6—Tolerate unexpected changes in business conditions	0.86	2.69 (0.98)
7—Bring product concepts to market in a timely manner	0.86	2.36 (1.07)
8—Work productively under continuous stress, pressure and conflict	0.86	3.03 (0.85)
9—Determine what the business will look like	0.86	2.19 (1.11)
10—Persist in the face of adversity	0.87	3.13 (0.83)

The results in Table 5 describe the Design Thinking Self-Efficacy (DTSE) measure. The scale for DTSE differs from the ISE, ETSE, and ESE Likert scales in order to preserve its comparability with prior work on this construct originated by

Table 5 Design thinking self-efficacy (DTSE)

Survey questions	Cronbach's alpha/factor loadings (Total sample n = 266)	Mean (SD)
	α = 0.75	78.0 (4.4)
1—Sense how another person feels and what they might be thinking	0.74	72.0 (21.0)
2—Look at problems in the world from different angles	0.68	80.0 (17.0)
3—Generate a wide variety of ideas	0.69	81.0 (18.0)
4—Build a prototype solution that satisfies user needs	0.74	75.0 (20.0)
5—Accept feedback on your work and make changes	0.73	83.0 (15.0)
6—Enhance the lives of people by finding a better way to do things	0.68	76.0 (18.0)

Schar (2020). Future analyses may explore how these various self-efficacy scales could be standardized to allow for greater comparability and exploration across constructs and datasets.

Overall, we find that the surveyed student alumni tended to display very high confidence across all of the self-efficacy measures. While the above analysis was conducted on the entire sample, future research will investigate potential differences and similarities in the self-efficacy measures across ME310 cohort years, as well as compare these scores with other engineering populations.

5 Future Work

The next phase of qualitative research has already begun with the use of these preliminary results related to the self-efficacy measures and career paths. The Entrepreneurial Self-Efficacy scores, academic and career pathways, and demographic profiles have been used to identify and select a cross-sectional sample of alumni from the last 5 to 20 years to participate in a brief 30-min interview. The criteria for selection were based on respondents' scores on Innovation Self-Efficacy and two dimensions of the Entrepreneurial Self-Efficacy measure established by DeNoble et al. (1999), developing new product and market opportunities and coping with unexpected challenges. Both of these measures align with key course elements and prototyping strategies in ME310.

These semi-structured interviews focus on "career maturing," with an interest in identifying alumni who continue to be involved with innovation and design in their post-ME310 careers. The purpose of these interviews was to learn how specific professional and academic experiences influence and contribute to alumni's confidence, interests, and career goals, thereby strengthening their interests, intentions and actions toward outcomes related to entrepreneurship and innovation.

While we recognize the relationships that exist among engineering, innovation and entrepreneurship in technological, social, and economic realms, a deeper understanding of the critical role that education plays is still needed. In the current study, we focused on the impact of ME310, a globally recognized course within the innovation and entrepreneurial educational ecosystem of Stanford, and future analyses will continue to draw out the details of these connections. We note that the instruments, methods, and processes for analysis and dissemination established in this research have the potential to become a model that could be adapted and replicated by other comparable courses, such as Stanford's Smart Product Design Fundamentals course (ME218) and potential collaborators in the Hasso Plattner Institute of Design–this extended work has already begun.

Future outcomes and deliverables from this study will inform several areas: (1) strengthening the validity and reliability of key variables in the ME310 Alumni Survey such that the instrument could be adapted and administered to other courses, alumni populations, and institutions outside of Stanford; (2) contributing to the engineering and entrepreneurship education literature on the innovation-related self-efficacy measures and the applicability of Social Cognitive Career Theory as a framework for understanding career goals in innovative work; and (3) more pragmatically, articulating the nuances related to the design and implementation of a course such as ME310 and how it can be improved in order to better support and benefit past, current, and future students.

Acknowledgements We have an extended team to say "thanks" to: Professors Larry Leifer and Mark Cutkosky (who helped us understand more about ME310), Crystal Pennywell and Elizabeth Mattson (our administrative colleagues who help us with budgeting and staff appointments), Mark Schar (who generously shared his Design Thinking Self-Efficacy measure), Hung Pham (a visiting researcher who participated in Stanford-Ignite Research Experiences for Teachers program in the summer of 2020), Timo Bunk (a visiting student researcher from the Technical University of Munich, who is extending this work to additional Stanford courses), Stanford Designing Education Lab colleagues who gave feedback and encouragement, and Elizabeth and Lucy Higgins who assisted with the mailing of the incentives. We are also grateful for those who participated in the pilot phase of survey design, to the many ME310 graduates who jumped in to fill out the final survey, and to those who agreed to be interviewed -- you are the essential ingredient of this research. Finally, we acknowledge the generous support of this work from the Hasso Plattner Design Thinking Research Program.

References

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology, 32*(4), 665–683

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.

Bialik, K., Fry, R. (14 Feb 2019). Millennial life: How young adulthood today compares with prior generations. Pew Research Center. https://www.pewsocialtrends.org/essay/millennial-life-how-young-adulthood-today-compares-with-prior-generations/

Brick, J.M., & Kalton, G. (1996). Handling missing data in survey research. *Statistical Methods in Medical Research, 5*(3), 215–238

Byers, T., Seelig, T., Sheppard, S., & Weilerstein, P. (2013). Entrepreneurship: Its role in engineering education. *Summer Issue of the Bridge on Undergraduate Engineering Education, 43*(2), 35–40

Carleton, T. (2019). *ME310 at Stanford university: 50 years of redesign (1967–2017).* Innovation Leadership Publishing.

Carleton, T., Leifer, L. (2009). Stanford's ME310 course as an evolution of engineering design. In *Proceedings of the 19th CIRP design conference—competitive design*, Cranfield University, 30–31 March 2009, pp. 547. Available at: https://dspace.lib.cranfield.ac.uk/handle/1826/3648

Chen, H.L., Grau, M., Brunhaver, S., Gilmartin, S., Sheppard, S., Warner, M. (2012). Designing the deployment of the pathways of engineering alumni research survey (PEARS). In *Proceedings of the annual meeting of the american society of engineering education*, TX.

De Noble. A., Jung. D., Ehrlich. S. (1999). Entrepreneurial self efficacy: The development of a measure and its relationship to entrepreneurial action. *Frontiers of entrepreneurship research.* Wellesley. Babson College. pp. 73–87.

Eesley, C., & Miller, W. (2012). *Impact: Stanford university's economic impact via innovation and entrepreneurship.* Stanford University. Available at SSRN: https://ssrn.com/abstract=2227460 or http://dx.doi.org/https://doi.org/10.2139/ssrn.2227460

Ewell, P.T. (2005). Alumni studies as instruments of public policy: The U.S. experience. *New Directions for Institutional Research, 126*, 19–29

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

Gilmartin, S., Thompson, M., Morton, E., Jin, Q., Chen, H.L., Colby, A., & Sheppard, S.D. (2019). Entrepreneurial intent of engineering and business undergraduate students. *Journal of Engineering Education, 108*, 316–336

Jin, Q., Gilmartin, S.K., Sheppard, S., & Chen, H.L. (2015). Comparing engineering and business undergraduate students' entrepreneurial interests and characteristics. *The Journal of Engineering Entrepreneurship, 6*(2), 1–24

Kubota, T. (June 14 201). Stanford engineering design course reaches its 50th year. Stanford News. https://news.stanford.edu/2018/06/14/engineering-design-course-reaches-50th-year/

Krueger, N.F., & Carsrud, A.L. (1993). Entrepreneurial intentions: Applying the theory of planned behaviour. *Entrepreneurship and Regional Development, 5*(4), 315–330

LeDuc, D. (May 20 2019). Who is generation Z: Meet the post-millennial and most diverse generation. Pew Trust Magazine. https://www.pewtrusts.org/en/trust/archive/spring-2019/who-is-generation-z

Lent, R. W., & Brown, S.D. (2006). On conceptualizing and assessing social cognitive constructs in career research: A measurement guide. *Journal of Career Assessment, 14*(1), 12–35

Lent, R.W., Brown, S.D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior, 45*, 79–122

Olsen, D. (2000). Institutional research. *New Directions for Institutional Research, 111*, 103–111

Pew Research Center. (Sep 3 2015). The whys and hows of generations research. https://www.pewresearch.org/politics/2015/09/03/the-whys-and-hows-of-generations-research/

Schar, M. (2020). Design thinking self-efficacy. [Unpublished presentation].

Schar, M., Gilmartin, S.K., Harris, A., Rieken, B., Sheppard, S. (2017a). Innovation self-efficacy: A very brief measure for engineering students. *Proceedings of the 2017 ASEE annual conference and exposition.*

Schar, M., Gilmartin, S.K., Rieken, B., Brunhaver, S.R. Chen, H.L. Sheppard, S. (2017b). The making of an innovative engineer: Academic and life experiences that shape Engineering Task and Innovation Self-Efficacy. *Proceedings of the 2017 ASEE annual conference and exposition.*

Joining Forces: Applying Design Thinking Techniques in Scrum Meetings

Franziska Dobrigkeit, Christoph Matthies, Ralf Teusner, and Michael Perscheid

Abstract The most prominent Agile framework Scrum is often criticized for its amount of meetings. These regular events are essential to the empirical inspect and adapt cycle proposed by Agile methods. Scrum meetings face several challenges, such as being perceived as boring, repetitive, or irrelevant, leading to decreased cooperation in teams and less successful projects. In an attempt to address these challenges, Agile practitioners have adopted teamwork, innovation, and design techniques geared toward improving collaboration. Additionally, they have developed their own activities to be used in Scrum meetings, most notably for conducting retrospective and planning events. Design thinking incorporates non-designers and designers in design and conceptualization activities, including user research, ideation, or testing. Accordingly, the design thinking approach provides a process with different phases and accompanying techniques for each step. While these techniques are often not new, they are revised and customized for teams with little design experience. These design thinking techniques can support shared understanding in teams and can improve collaboration, creativity, and product understanding. For these reasons, design thinking techniques represent a worthwhile addition to the Scrum meeting toolkit and can support Agile meetings in preventing or countering common meeting challenges and achieving meeting goals. This chapter explores how techniques from the design thinking toolkit can support Scrum meetings from a theoretical and practical viewpoint. We analyze Scrum meetings' requirements, goals, and challenges and link them to groups of techniques from the design thinking toolkit. In addition, we review interview and observational data from two previous studies with software

F. Dobrigkeit (✉) · C. Matthies · R. Teusner · M. Perscheid
Hasso Plattner Institute for Digital Engineering, Campus Griebnitzsee, August-Bebel-Str. 88, 14482 Potsdam, Germany
e-mail: franziska.dobrigkeit@hpi.de

C. Matthies
e-mail: christoph.matthies@hpi.de

R. Teusner
e-mail: ralf.teusner@hpi.de

M. Perscheid
e-mail: michael.perscheid@hpi.de

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_17

development practitioners and derive concrete examples. As a result, we present initial guidelines on integrating design thinking techniques into Scrum meetings to make them more engaging, collaborative, and interactive.

1 Introduction

Agile software development processes, especially Scrum and Scrum-hybrids (i.e., Scrumban or Scrum/XP), are standard practices in modern software companies. Team meetings employed in these methods facilitate empirical process control, create regularity, and decrease the demand for additional unplanned or unstructured sessions (Rubin 2012; Schwaber & Sutherland, 2017). The literature regarding Scrum provides detailed descriptions of each meeting and *what* should be accomplished in them (Schwaber & Sutherland, 2017). However, few prescriptions or best practices on *how* these goals should be achieved exist. Consequently, teams face several challenges concerning the facilitation and perceived effectiveness of meetings including:

- Not all meetings are equally relevant for all participating team members, leading to a decrease in participant engagement (Akif & Majeed, 2012).
- Planning and review meetings may be perceived as a waste of development time and as too "simple" (Cho, 2008).
- Regular meetings featuring identical structures create monotony which can negatively affect outcome quality (Kua, 2013).
- A lack of follow-through regarding meeting outcomes leads to a feeling that "the same things are discussed over and over" (Przybyłek & Kotecka, 2017).
- The transparency afforded by Agile approaches can be abused leading to feelings of being exposed or inadequate in team members (Conboy et al., 2011).

Knowledge of these common meeting issues allows addressing, resolving, and preventing problems before they become hindrances to team collaboration (Cho, 2008). Researchers and practitioners have proposed and collected various activities and best practices for different phases and contexts of the Scrum process. Much of these focus on Retrospective (Baldauf, 2018; Jovanović et al., 2016; Ng et al., 2020) or Sprint Planning meetings, e.g., the common *Planning Poker* in Scrum (Haugen, 2006; Rubin 2012), but other activities exist (Gray et al., 2010; Hohmann, 2006).

Design thinking (DT) has emerged as an approach that can support software development by providing interactive and engaging techniques for collaborative problem analysis and solution development. It can improve creativity, product understanding, collaboration, and empathy toward customers and the team (Dobrigkeit et al., 2020). Employing DT as a preceding phase to Agile software development is a popular scenario (dos Santo Braz et al., 2019).

However, in line with previous studies (Pereira et al., 2018), our research has also shown that the application of DT in later stages of the development cycle is valuable for development teams (Dobrigkeit & de Paula, 2019). Developers especially value the change in working style and the required customer-centered mindset. As such, DT techniques provide an opportunity to add further useful activities to the Scrum meeting toolbox. They can help address common challenges, achieve meeting requirements for attendees, or aid in reaching meeting outcomes that have the support of all team members. We, therefore, explore how techniques from DT can be applied to support Scrum meetings.

2 Background and Related Work

In this section, we briefly describe the key concepts of Scrum and DT as well as existing research into their activities and techniques. Additionally, we describe existing approaches to combine Agile Software Development and DT.

2.1 Scrum

Scrum is an Agile process framework for managing work on complex products in short iterations (*Sprints*), with an emphasis on software development (Schwaber & Sutherland, 2017). It defines the different roles of a Scrum team, the rules to follow, the artifacts to create and the meetings that steer and organize a Sprint (Rubin 2012; Schwaber & Sutherland, 2017). A Scrum team consists of the Product Owner (PO), the Scrum Master (SM), and the Development Team. The PO is responsible for managing the product vision and the *Product Backlog*, which contains the required work items needed to achieve the vision. The SM facilitates the Scrum meetings and supports the team in overcoming issues. The development team develops the (software) product. Work happens in an iterative, incremental manner, carried out in cycles called Sprints (Deemer et al., 2012). There are five main Scrum meetings that facilitate team cooperation and collaboration (Rubin 2012):

- *Backlog Refinement*, where work items are prepared
- *Planning*, where the work items for the next Sprint are decided
- *Daily Scrum*, provides a regular team status update
- *Review*, concerned with inspecting the product and collecting feedback
- *Retrospective*, focuses on process improvement.

We provide a detailed description of each meeting with its requirements and outcomes in Sect. 4.

2.1.1 Meeting Challenges

Studies on the subject of the challenges involved in running Scrum meetings have identified multiple issues of relevance. The literature notes that Scrum events are not equally relevant for all participating team members, leading to a decrease in engagement (Akif & Majeed, 2012). Similarly, meetings such as Sprint Planning and Review, which were perceived by participants as “simple,” have been described as a waste of valuable development time (Cho, 2008). If Scrum meetings are repeatedly held in the same fashion and with similar structures, apathy and monotony felt by attendees can negatively affect the quality of meeting outcomes (Kua, 2013). In case studies of multiple teams, retrospective meetings were judged to be ineffective by participants, as the same issues were repeatedly discussed in subsequent meetings (Przybyłek & Kotecka, 2017). Furthermore, the downsides and drawbacks of the full transparency and visibility provided by Scrum events have been highlighted, such as exposing developer shortcomings publicly. This can prove to be counterproductive, as team members can be made to feel inadequate or can lead to unhealthy environments where this transparency is abused (Conboy et al., 2011).

The prescriptive or “dogmatic” nature of Scrum events has been criticized, e.g., regarding the Daily Scrum meeting, which is strictly time-boxed and should ideally always be performed every day at the same time with the entire team. These requirements can break down in real-world circumstances featuring distributed teams, meeting setup times, flexible working schedules, and varying team sizes (Meyer, 2014).

2.1.2 Existing Activities for Scrum Meetings

Various activities and best practices have been proposed in the related literature for different phases and contexts of the Scrum process to address common Scrum meeting challenges. In particular, Scrum process facilitators have introduced tools and meeting agendas that help in achieving the main goals of Scrum meetings, e.g., for the Retrospective (focusing on creative prompts for feedback collection) and Sprint Planning (focusing on collecting effort estimations).

Much previous research is available on Retrospectives, and the activities that teams can employ during these meetings and studies continue to be published (Ng et al., 2020). Retrospective meetings have been in use by teams since before Agile methods became widespread. Team activities or “games” that participants play to keep sessions from becoming stale and repetitive have been used in Retrospectives since their inception (Kerth, 2000). These interactive games are designed to encourage the required reflection and team collaboration needed for the meeting. Derby and Larsen describe the purpose of Retrospective activities as helping a team “think together” (Esther & Larsen, 2006). Retrospective activities have been shown to positively impact the creativity, involvement, and communication of team members (Przybyłek & Kotecka, 2017).

Sprint Planning and Backlog Refinement meetings are also closely associated with best practices in agile processes. These activities are concerned with the requirement of producing effort estimates for work items. Unstructured group estimation processes pose the risk of being highly influenced by company politics, group pressure, or dominant personalities, which may reduce estimation performance (Haugen, 2006). To allow teams to come up with shared estimates while reducing bias-related antipatterns, tools such as *Planning Poker* were introduced (Ralph, 2013). Through the use of structured meeting agendas and categories of answers, these activities can help ensure that all team members participate and that their opinions are heard, regardless of group influence (Haugen, 2006).

Scrum practitioners have recommended different activities for use in Scrum meetings, but few of these have garnered widespread attention and adoption.

2.2 Design Thinking

The term Design Thinking is used ambiguously and can describe a) a cognitive style of working dominant with designers, b) the idea of including design practices into all parts of an organization or c) an approach to creating new and innovative products and services. In this article, we understand DT as an approach to new product and service development, as it is taught at Stanford or the HPI School of Design Thinking Potsdam. In this approach, the design process and methods from design and other disciplines have been remolded in a way that non-designers can apply them to successfully create new products and services. Within this understanding of DT, the DT process is described as an iterative and flexible multi-step process with recommended techniques for each step. While authors do not agree on the number of steps, the steps of different processes can be mapped to (a) steps that focus on learning about the problem and the user, (b) ideating on a solution, and (c) prototyping and testing this solution with users. Detailed descriptions of these processes are available (Brenner et al., 2016; Brown, 2008; Thoring & Müller, 2011; Wölbling et al., 2012). There commended techniques accompanying each process step, provide practitioners with multiple ways to facilitate each step (cmp. Martin & Hanington, 2012; IDEO, 2003; Kumar, 2012). We give a detailed overview of such techniques in Sect. 5.

Research on DT found that its advantages include increased collaboration, empathy, product differentiation, and cost savings due to reduced redesign work and shorter lead time to development (Carlgren & Rauth, 2014).

The working style for most DT techniques is very interactive and research with development teams suggests that developers welcome this variety in work style when applying DT. Additionally, research suggests that developers value and implement DT techniques when they have the required knowledge (Dobrigkeit & de Paula, 2019) and that developers can increase empathy toward users and team members, as well as shared product understanding and collaboration (Dobrigkeit et al., 2020).

2.3 Design Thinking and Agile Software Development

Research on how to integrate DT and software development suggests different approaches. Lindberg et al. present four models of how DT can support software development (Lindberg et al., 2012). We previously suggested adding two more models and expanding the list to six (Dobrigkeit & de Paula, 2019). Table 1 provides an overview of these six integration possibilities.

In line with these proposals, researchers have suggested different integration approaches covering the range of different models, running DT as a preceding,

phase to agile development, as suggested in the first three models seems the most common form: the Integrated Design Thinking and Lean Development Approach (Hildenbrand & Meyer, 2012), the Nordstrom Innovation Lab Process (Grossman-Kahn & Rosensweig, 2012), InnoDev (Dobrigkeit et al., 2018) or Converge (Ximenes et al., 2015).

Continuous approaches of applying DT and software development are proposed by the Integrated Design Thinking and Agile Framework for Digital Transformation (Gurusamy et al., 2016), the Human-centered Agile Workflow (Glomann, 2018) and DT@XP (Sohaib et al., 2018), which suggest regular DT workshops or phases as part of each development cycle. Ad hoc DT implementations in later stages of the development cycle are proposed by InnoDev (Dobrigkeit et al., 2018) and Converge (Ximenes et al., 2015). These approaches propose spontaneous DT workshops with low planning overhead, when new features are added to the product, when tackling identified blockers, or supporting agile meetings. Toolbox approaches that implement these concepts were previously suggested (Dobrigkeit et al., 2020; Pedersen, 2016).

Adding DT techniques to the toolbox of Scrum meeting facilitators would be in line with approaches that follow the ad hoc DT model or the toolbox model.

Table 1 Overview of models integrating Design Thinking and software development processes adapted from Dobrigkeit and Paula (2019)

Name	Description
Split project model	DT as separate project before development
Overlapping teams model	DT as initial process phase with one or more development team members participating
Unified project model	DT as initial process phase and a large overlap of DT team and development team
Continuous DT model	Smaller DT phases are regularly integrated into agile development cycles
Ad-hoc DT model	DT as supporting tool that development teams can apply in form of DT workshops or DT phases if required
Toolbox model	Methods developers can use to overcome problems they can't solve with IT methods

3 Methodology

We performed a two-step study to explore how techniques from the DT toolkit can be applied to address the challenges of Scrum meetings. First, we compared the requirements, challenges, and existing method recommendations from the Scrum literature with various methods described in the DT literature. In a following step, we supported this theoretical comparison with empirical evidence and concrete examples collected from a meta-analysis. In this meta-analysis, we reviewed data collected in two former studies in which the topic of *DT techniques in Scrum meetings* had emerged.

The first original study *(OS-1)* was a ten-month longitudinal study in a global software company with a team of three Scrum teams who were experienced in both Scrum and DT. In this study, we explored when, where, and how DT was used during agile development. We observed several of the team meetings during the ten-month period and interviewed eleven team members regarding their experiences using DT in their team. Additionally, we interviewed five employees from two further teams in the same company. This study and its results are further described in Dobrigkeit and de Paula (2019).

In the second original study *(OS-2)*, we facilitated workshops introducing DT tools in agile development teams from six different companies. In this study, we explored the comprehensibility and applicability of a DT toolbox to everyday agile development with DT beginners. To this end, we prepared a collection of worksheets for twelve different DT tools, which we validated with six different companies. We observed one company in detail over a period of twelve weeks, in which they were introduced to one method each week and were asked to apply that method at least twice during the week. At the end of each week, we interviewed the team in a group interview and asked all team members to fill out a method survey.

The five other companies were introduced to our toolkit in a workshop format in which we introduced them to three of the methods from our toolkit and afterward conducted a group interview. The interview focused on finding out the experiences made within the workshop and if, where, and when team members would make use of these methods again. This study and its results are described in detail in a previous publication (Dobrigkeit et al., 2020).

3.1 Research Context

As the focus of OS-1 was on understanding where, when, and how DT supports agile development, we chose the global software company because it had embraced DT as well as agile development practices for several years. The three teams participating in OS-1 were chosen because they were reported to be using DT in their daily work and accordingly all interviewees in this study had prior experience with DT, albeit to a different extent.

Table 2 Teams interviewed and observed for this meta-analysis

Original Study	Team	# People	DT Exp.	Scrum Exp.	Short description
OS-1	Team A	25	medium to high	high	German-based team, incl designers developers and managers of a national software company developing a software for real estate management.
OS-1	Team B	3	high	high	US-based team, 1 designer, 1 developer and 1 manager of a national software company that develops proof-of-concept prototypes for new technologies or unsolved customer problems.
OS-1	Team C	2	high	high	US-based team, 1 designer and 1 manager of a national software company that acts as DT consultants to company customers, supporting them during their own DT efforts through training, coaching, and co-creation.
OS-2	Team D	10	low to medium	high	Team of programmers and designers from a medium sized, German software development company.
OS-2	Team E	6	low to medium	high	Team of programmers working in data analysis, artificial intelligence, machine learning and computer vision from a German research organization.
OS-2	Team F	8	low to medium	high	Team incl. a sales-manager, a designer and a mixed group of front-end and back-end developers from a German company developing team collaboration and security solutions.
OS-2	Team G	9	low to medium	high	Software developers of different teams from a German company developing project management software.
OS-2	Team H	8	low to medium	high	Diverse team incl. UX designers, Agile coaches and developers of a German company developing business applications.

The six companies chosen for OS-2 had little or no prior experience with DT, with the focus of this study being to explore the applicability of a toolbox for DT novices.

The participants of these two studies featured a broad range of DT experience. Table 2 provides an overview of the teams that we observed and interviewed for this study.

3.2 Data Collection and Analysis

For the theoretical comparison, we extracted the requirements and outcomes of each of the Scrum meetings as well as existing activities recommended for these meetings from the seminal literature. We concentrated on the descriptions by the Scrum originators (Schwaber & Sutherland, 2017; Sutherland & Schwaber, 2007) and highly referenced introductory literature that presents and describes the basic Scrum events (Deemer et al., 2012; Kniberg, 2015; Meyer, 2014; Rubin 2012). We compared these requirements with techniques frequently mentioned in the DT literature to create a mapping of Scrum meetings and DT techniques that can support them.

For the meta-analysis, we re-analyzed the data collected during OS-1 and OS-2.

In OS-1, we wrote observation notes during various Scrum meetings of the observed Team A over a period of ten months and recorded and transcribed eleven interviews with members of that team as well as five additional interviews with members of two further teams (Team B and C) from the same company. In OS-2, we collected observation notes over a period of twelve weeks within Team D and during five workshops with teams E–H. Additionally, we recorded and transcribed twelve group interviews in Team D (one after each week) and one group interview with each of the other four teams following their workshops. In this meta-analysis, we iteratively coded the available materials with a focus on the use of DT during Scrum meetings. We thus derived concrete examples to support our initial mapping. As a result, we created initial guidelines on how to integrate DT techniques in each of the Scrum meetings for Scrum Masters and teams that want to improve their meetings' level of interactivity and collaboration or simply desire more variety.

4 Goals and Requirements of Scrum Meetings

The Scrum meetings, sometimes referred to as *ceremonies* (Meyer, 2014), conducted by teams during a development iteration form the core of the Scrum software development method. These meetings mark points in the process where team members regularly discuss iteration progress and process aspects as well as work on development artifacts. The Scrum Guide describes the purpose of these meetings as creating "regularity and to minimize the need for meetings not defined in Scrum" (Schwaber & Sutherland, 2017). Schwaber and Sutherland point out that all meetings are essential and that not including a prescribed Scrum meeting results in "reduced transparency" and a "lost opportunity to inspect and adapt" (Schwaber & Sutherland, 2017). In the following, we present and characterize the five central Scrum meetings, with a particular focus on their requirements and outcomes.

4.1 Product Backlog Refinement

In the *Product Backlog Refinement* meeting, also referred to as *Backlog Grooming* (Kniberg, 2015), the team focuses on preparing product backlog work items for the following one or two Sprints (Deemer et al., 2012). Refinement includes activities such as analysis of requirements, adding required details to work items, estimating implementation effort, or splitting large work items (Schwaber & Sutherland, 2017). Furthermore, outdated or redundant items should be removed from the product backlog to keep its components as clear and actionable as possible. While the refinement of work items can form part of a Planning meeting, depending on teams' preferences and context, most of this work should be done before the Planning meeting (Kniberg, 2015). Separating these meetings allow the subsequent Planning to take advantage of well-analyzed and estimated items, making it more efficient (Deemer et al., 2012).

Many of the details of how Product Backlog Refinement is conducted are up to the team. Scrum does not feature a defined time box or even a frequency for the meeting. However, the Scrum Guide notes that the meeting should not consume "more than 10% of the capacity of the Development Team" (Schwaber & Sutherland, 2017).

Attendees: Development Team, Scrum Master, Product Owner.

Point in Process: Custom, depending on the Scrum implementation.

Requirements: A shared understanding of both the technical effort to implement a solution and the business value it delivers. This allows establishing cost estimates and enables discussions on priorities.

Outcome: A set of well-understood, clearly described, thoroughly analyzed, and estimated set of user stories in the Product Backlog (Deemer et al., 2012).

4.2 Sprint Planning

A *Sprint Planning* meeting marks the start of a new development iteration in Scrum. They are used to prepare for the upcoming Sprint and are typically divided into two parts: the first concerned with *what* to build, the second with the specifics of *how* to implement it (Deemer et al., 2012). The Product Owner communicates their vision of what is most critical regarding the project's next steps. Together with the Development Team, the PO discusses the highest priority items (from the Product Backlog), which represent those features that are most interesting for the near future. The contexts and details for these high-priority items are reviewed by team members, providing the rationale and a shared understanding of the value inherent in the planned features.

Based on these discussions and insights, the team may then define a shared, relatively stable *Sprint Goal* (Deemer et al., 2012) and the next items to be worked on (Rubin, 2012; Sutherland & Schwaber, 2007). The Sprint Goal summarizes the objective that the team aims at achieving within the Sprint, which ideally has a cohesive theme. It is based on the latest product increment, the projected capacity, and the past performance of the team (Deemer et al., 2012). To forecast the functionality that can be developed during the Sprint, the entire Scrum team collaborates to understand the work needed to achieve the Sprint Goal (Schwaber & Sutherland, 2017). The Planning meeting provides an opportunity to flesh out further details of the work items planned for the first days of the Sprint, e.g., by creating additional subtasks for work items or bugs. When effort estimates are not yet available (e.g., as work items were recently updated), the teamworks together to provide these assessments. Considering the priority and effort estimates, the Development Team decides on the amount as well as the specific tasks it will complete in the upcoming Sprint. The Development Team has the ability and responsibility to additionally include relevant work items with potentially lower priority, e.g., in the case of dependencies (Deemer et al., 2012).

By the end of the Sprint Planning meeting, the Development Team should have formed a consensus on the plan of action and should be able to explain how it intends to accomplish the Sprint Goal and build the anticipated product increment (Schwaber & Sutherland, 2017).

Attendees: Development Team, Scrum Master, Product Owner (optional for breaking down items into tasks).

Point in Process: At the beginning of a Sprint.

Requirements: The Product Owner representing the customer and the development team must share a mental model and a deep (technical) understanding of the work ahead and the business value it presents (Conboy et al., 2011).

Outcomes:

- A coherent Sprint Goal, summarizing the development work to be done.
- A list of items to be worked on during the next Sprint, ideally including the highest priority ones.
- A plan of how to deliver the planned functionality in a product increment.

4.3 Daily Scrum

The *Daily Scrum* meeting is a short team meeting to enable collaboration, in which the Development Team inspects and synchronizes their work as well as outline the next aims on a daily basis (Schwaber & Sutherland, 2017). The primary justification for a daily meeting during a Sprint is the general Agile principle that direct contact is critical to project success (Meyer, 2014). In the Daily Scrum, progress toward the specified Sprint Goal and the remaining work items of the Sprint Backlog are reviewed, with a particular focus on identifying impediments to the team's productivity (Deemer et al., 2012). Necessary In-depth discussions of identified issues are postponed (Rubin 2012) and the Scrum Master helps team members resolve obstacles in the following steps (Deemer et al., 2012; Sutherland & Schwaber, 2007). Daily Scrum meetings are designed to be time-boxed to 15 min, to improve communication, and to possibly eliminate further meetings by facilitating quick decision-making (Schwaber & Sutherland, 2017).

The structure of the Daily Scrum meeting is decided by the Development Team with a focus on producing an actionable plan for the day. It represents an opportunity for the members of the Development Team to make realistic commitments to each other (Meyer, 2014). The Scrum Guide proposes the following three questions for every participant to structure the meeting (Schwaber & Sutherland, 2017):

- What did I complete yesterday that helped meet the Sprint Goal?
- What will I work on today to help meet the Sprint Goal?
- Do I see any impediments or blockers toward the Sprint Goal?

The Daily Scrum meeting is not intended to be used as a detailed status reporting meeting; instead, it focuses on quick information sharing on what is happening across the team. Its tone should ideally be pleasant and enjoyable while staying informative. The meeting is also known as a *Stand-up* as one of the original ideas to ensure a short session was to require participants to stand (Meyer, 2014).

Attendees: Development Team (Product Owner and Scrum Master participate as developers if they are actively working on Sprint Backlog items).

Point in Process: Every day of a Sprint.

Outcome: An understanding between team members of the progress toward the Sprint Goal and the impediments that need to be tackled (Meyer, 2014).

Requirements: As the meeting is strictly time-boxed, concise recollections and summaries of recent work by attendees as well as individual reflections on encountered impediments are essential.

4.4 Sprint Review

Held at the end of the Sprint, the *Sprint Review* meeting is used to reflect on the work of the completed Sprint (Meyer, 2014). It focuses on inspecting the created product increment and eliciting feedback (Sutherland & Schwaber, 2007) and is the time to showcase the team's work. It is, therefore, also referred to as a *Sprint Demo* (Kniberg, 2015), where a hands-on inspection of the real software takes place, and the Development Team gets credit for their accomplishments (Kniberg, 2015). During the meeting, the team members present the Sprint results to key stakeholders and customers and discuss progress. The meeting provides an opportunity for the Product Owner to gain insight into the status of both the current product and the rest of the team. Other team members gain updates from the Product Owner and the market situation (Deemer et al., 2012).

To be demonstrable, Sprint's work must result in a working product increment and meet the team's agreed quality bar. In the case of software development, this might refer to a system that is integrated, tested, user documented, and "potentially shippable" (Deemer et al., 2012). The Sprint Review is the second to last meeting of a development iteration. The entire group collaborates on the next steps so that the Sprint Review provides valuable input to the Sprint Planning meetings of the next sprint (Schwaber & Sutherland, 2017). The Product Backlog is updated based on the completed work items or shifts in business value or the market (Schwaber & Sutherland, 2017).

Attendees: Development Team, Scrum Master, Product Owner, other project stakeholders as appropriate.

Point in Process: After work on the Sprint has concluded.

Outcome: A revised Product Backlog based on collected feedback (Schwaber & Sutherland, 2017).

Requirements: Knowledge of the status of work items. An understanding of how the developed functionality integrates with the rest of the product. Strategies to collect actionable feedback regarding the software increment.

4.5 Retrospective

The *Retrospective* meeting is the last Scrum meeting of a Sprint, taking place after the Sprint Review and before the next Sprint Planning. It is focused on inspecting and improving the executed development process of teams and is concerned with creating a plan for improvements to be enacted in the future (Schwaber & Sutherland, 2017). The meeting helps identify and rank aspects of the completed Sprint that have proven themselves and should be continued and those that should be discontinued (Deemer et al., 2012; Sutherland & Schwaber, 2007). Assumptions that led the team astray should be identified and their origins examined. Improvement opportunities can be identified for all process aspects, including people, relationships, work process, and tools, with the overall goal of making the development process and the employed practices more effective and enjoyable (Schwaber & Sutherland, 2017).

The Scrum method itself does not prescribe the detailed contents of Retrospectives and many activities as well as agendas for conducting the meeting have been proposed (Baldauf 2018). By the end of the Retrospective, the team should have found a consensus on the concrete improvement actions it will attempt to implement in the next Sprint. The most impactful improvements should be addressed as soon as possible and may be added to the next iteration's Sprint Backlog.

Attendees: Development Team, Scrum Master, Product Owner.

Point in Process: At the very end of a Sprint.

Outcome: Set of ordered action items regarding the team's work processes to be enacted in the next Sprint (Meyer, 2014).

Requirements: Detailed recollections of the enacted process of the concluded Sprint, including relevant interactions, dependencies, highlights, and encountered issues. Reflection on the severity of these items.

5 Techniques in the Design Thinking Toolkit

As described before, the DT approach uses a multitude of techniques that support each step of the DT process. These techniques originate from diverse areas, like quality management, research in creativity and design, research in communication, ethnography, and informatics (Brenner et al., 2016). There is no complete list of DT techniques, and practitioners add new techniques to their toolkits from other disciplines. So while no exact definition of a DT method exists, some methods are frequently mentioned and described by researchers (Carleton et al., 2013; Dobrigkeit et al., 2018; Wölbling et al., 2012) and practitioners (Martin & Hanington, 2012; IDEO, 2003; Kumar, 2012). Method collections, such as (Martin & Hanington, 2012; Carleton et al., 2013; IDEO. IDEO, 2003; Kumar, 2012) both from research and practice provide an overview of possible techniques and are often organized by steps in the design or DT process. For this section we propose a slightly different categorization, focusing on the tasks that can be achieved with a group of techniques. In the following, we describe the categories of techniques that are typically found in a design thinker's toolkit and give a few concrete examples for each category. Table 3 provides an overview of the presented categories with some examples.

5.1 Warm-Up Techniques

Warm-ups also called ice-breakers or energizers, are a known concept for facilitators of all kinds of team-based workshops. Warm-ups can serve a variety of functions. In the very literal sense of warming-up, these methods can help to energize participants after a longer presentation or early in the morning by providing several minutes of physical activity, e.g., Dancing in the Dark. For this activity, participants block their sight, e.g., with a post it or their hats. Then the moderator starts music that is suitable for dancing and asks the participants to perform different types of dancing and movements, for example, Tango, Belly Dance or a simple spin. In a new team or at the beginning of a workshop, warm-ups can help participants get acquainted with each other by facilitating the sharing of specific knowledge. For example, the exercise Two Truths and One Lie asks each person in the group to state three things about themselves, two of which are true and one is a lie. Afterward, the rest of the group has to guess which one is the lie. During a workshop warm-up techniques can help to get the group into the right mindset for the following exercises. If for example the next exercise for the group is to brainstorm ideas, a warm-up that practices brainstorming can help to get the creativity flowing. The paper clip exercises are such a warm-up. It asks each participant to write down as many possible usage scenarios for a big paper clip as they can over a certain time frame, e.g., 5 min.

Similarly, if building a prototype is the next step a warm-up that requires building something gets people into the activity. For example, in the marshmallow challenge teams are competing to build the highest possible structure out of tape, spaghetti and a marshmallow. By doing so they learn to build, fail and iterate together as a team. If at some point in time the team is in a negative mood, warm-ups can provide a fun activity that helps to create a positive atmosphere before carrying on. If a team is stuck, they provide distraction, making it possible to return to the task later.

Table 3 Categories of Design Thinking techniques and common examples

Method Category	**Examples**
Warm-up Techniques	Dancing in the Dark, Two Truth and One Lie, Paperclip Exercise, Marshmallow Challenge
Energizing group exercises designed to help participants ease into an atmosphere of teamwork.	
Unpacking Techniques	Charetting, Semantic Analysis, Mind Mapping
Group exercises designed to share and discuss the existing understanding and assumptions.	
Desk Research Techniques	News Search, Market Research, Technology Research
Research within books, published research literature or the internet,e.g. by searching for market trends, patents, news, etc.	
Field Research Techniques	Extreme User or Expert Interviews, Fly on the Wall, Shadowing, Contextual Inquiry, Immersion
Techniques geared towards learning directly from the user or stakeholder.	
Knowledge Sharing Techniques	Story Telling, Role Playing
Techniques geared towards sharing experiences, information and results from research or testing with the team.	
Knowledge Organization Techniques	Clustering, Venn-Diagram, 2x2Matrix, Timeline
Techniques geared towards organizing and visualizing information from research and testing.	
Knowledge Consolidation Techniques	Personas, PoV Statement, Journey Maps, Storyboards
Techniques geared towards condensing information from research or testing in a meaningful way usable for later process stages.	
Idea Generation Techniques	How to and How Might we Question, Brainstorming
Techniques that prompt the generation and sharing of ideas towards a specific prompt or question.	
Prototyping Techniques	Sketches, Role Plays, Storyboards, Wire Frames
Techniques that represent an aspect of an idea or a product in a meaningful way to gather feedback or allow for testing, e.g., by telling a story or presenting a schematic visual.	
Testing Techniques	Think Aloud, Concept Testing, A/B Testing, Usability Testing
Techniques, in which feedback or experiences of users are gathered through direct engagement with a prototype or the product.	
Feedback Techniques	I Like I wish, Feedback Capture Grid, Empathy Map
Techniques developed to elicit and collect structured feedback using different categories.	
Facilitation Techniques	Check-in & Check-Out, Team Rules, Planning sessions, Reflective sessions, Voting techniques
Techniques usually used by coaches and facilitators to ensure effective team work.	

5.2 Unpacking Techniques

Techniques used for unpacking are designed to share and discuss the existing understanding and assumptions on a topic. While doing so, teams can reach a shared understanding and uncover knowledge gaps or assumptions that have to be verified. Especially in the beginning of the process, they can help to understand and analyze the design challenge. Common unpacking activities include charetting, semantic analysis or simple brainstorming techniques such as mind mapping. Charetting is an activity that can be used for unpacking but also for ideation. During unpacking, the team first brainstorms relevant users or contexts for their design challenge. In a second step, the team picks one of these users and brainstorms potential issues relevant to that user. In the third step, the team picks the most interesting challenge and brainstorms potential solutions. Steps 2–3 are repeated for further users. Following these brainstorming tasks, the team discusses what they have discovered and rephrases the design challenge to reflect the new understanding. For a semantic analysis of the given design challenge, the team highlights important passages or words from the design challenge and discusses experiences, thoughts, open questions, associations and assumptions in detail. The information discussed is collected on sticky notes or the whiteboard. Thus, the team explores different aspects of the challenge and reaches a common understanding. At the end of the analysis, the team often rephrases the design challenge to reflect their new understanding. A less structured approach to unpacking is mind mapping, whereby the team brainstorms whatever comes to mind when thinking of the design challenge and then organizes the information in a mind map with the challenge in the center.

5.3 Desk Research Techniques

Desk research techniques, in contrast to field research, include analysis activities that can be done from one's desk. These include reviewing studies, reports or patents in published research literature as well as online sources and gray literature. Furthermore, market or technology trends can be explored usually through specific search sites on the Internet.

5.4 Field Research Techniques

Field research techniques are designed to learn directly from users and stakeholders. Several of the techniques originate from qualitative and ethnographic research in social sciences. These techniques can be separated into interview techniques, observation techniques, and self-immersion. Interview techniques elicit information from users or stakeholders by asking them questions. Common interview techniques are extreme user or expert interviews; or group interviews. Observation techniques allow the team to learn from the users or stakeholders in their actual environment by observing (and maybe asking them questions), e.g., in their workplace. Observation techniques common in DT are Fly on the Wall, Shadowing or Contextual Inquiry. Immersion techniques aim for the team to make relevant experiences themselves if possible. These techniques include trying out a service or product by themselves, or simulating specific conditions, e.g., bad eyesight, for an extended period of time.

5.5 Knowledge Sharing Techniques

Knowledge sharing techniques support a DT team in distributing the information collected through research and testing activities. A common technique for knowledge sharing is storytelling. In this activity, the involved researchers describe the details of their research to the rest of the team in the form of a story. The other participants are asked to listens attentively and to collect the important information on sticky notes. The form of a story should enable presenting the gathered information in a simple format and is supposed to provide reference points for anchoring key facts in the long-term memory of listeners.

5.6 Knowledge Organization Techniques

DT teams collect a vast amount of information that has to be organized in meaningful ways in order to learn from it. Common techniques to do so include clustering, Venn Diagrams, 2×2 Matrices, Timelines or process diagrams. With clustering, the team aims to group the information and looks for relationships or contradictions between clusters. Diagrams like matrices, Venn Diagrams or timelines help the team to organize their findings and highlight specific aspects, such as time or product categories and thereby uncover hidden needs and problems, or knowledge gaps.

5.7 Knowledge Condensation Techniques

In order to further work with the knowledge gained through research and testing, the team needs to condense it in such a way, that it can easily be accessed in later process stages. Personas, Point of View Statements, Storyboards, or Journey Maps are common techniques to condense knowledge. Personas represent a certain group of users with their needs, problems and wishes. The Point of View Statement condenses those problems from the problem domain the team currently wants to tackle. Storyboards or Journey Maps can be used to condense the experiences of many users into one typical experience.

5.8 Idea Generation Techniques

An integral part of every design process is to develop and discuss solution concepts. Many techniques from design or other creative disciplines can be used to develop ideas within a team. Common techniques include How to or How might we questions as brainstorming prompts, and a myriad of brainstorming techniques. The How to and How might we questions are often related to a Point of View Statement or a Persona derived from knowledge consolidation. They represent a specific problem to solve (How might we) or a specific aspect of the problem or the solution (How to). With regard to brainstorming techniques, different techniques work for different people and teams. Some brainstorming techniques let individuals brainstorm silently first and follow with a sharing and another group brainstorming session. Other brainstorming techniques use group brainstorming only. Some brainstorming techniques incorporate bodily activity. Some brainstorming techniques even use further prompts to generate ideas such as images or objects. Additionally DT prescribes a set of brainstorming rules.

5.9 Prototyping Techniques

In order to further develop an idea or test it with team members, users or stakeholders the idea needs to be prototyped. Prototypes can be a number of things, including simple sketches, product packages, role-plays, storyboards, wireframes, or functioning prototypes. Prototypes can have a different level of abstraction, ranging from low-fidelity prototypes designed to test basic interactions or ideas to high fidelity prototypes to test details of the interaction and functionality. Additionally, prototypes may cover the complete solution concept or single aspects of the concept.

5.10 Testing Techniques

Evaluating concepts with the actual user is an important aspect of DT. Several techniques to test concepts from areas such as UX Design, Human-Computer Interaction, Design, and Software Development are part of the DT toolkit. Common techniques include the Think Aloud Method, Concept Testing, A/B Testing, or Usability testing. With the Think Aloud Method, the user is asked to look at or try out the prototype and voice his thoughts out loud while doing so. The tester can ask clarifying questions. Concept Testing usually involves a low-fidelity prototype to explain the concept and an interview to gather feedback. A / B Testing is used to derive feedback for different alternatives be it features or concepts. Usability Testing aims to test how easy a product is to use.

5.11 Feedback Techniques

Feedback techniques can help to elicit and provide feedback, as well as document feedback. Most feedback techniques provide a specific set of categories in which to organize the feedback. For example, I like, I wish asks participants to provide feedback about what they liked and what they wish would have been different. Sometimes I like, I wish has an additional category called What if or How to in which participants can provide concrete examples to improve in the future. Other common feedback techniques include Five Finger Feedback, the Empathy Map, and the Feedback Capture Grid.

5.12 Facilitation Techniques

Facilitation techniques aim to ensure conditions in which the team can effectively work toward their objectives. These techniques include those that help a team, manage conflicts, reflect on their process or manage their resources, e.g., time or energy. Common techniques include: team rules, which a team agrees on, check-ins that allow for team members to transition into the group work and let the team know what might affect their working today, check-outs or reflective sessions similar to retrospectives that provide space for teams to reflect on their process and their teamwork. Voting techniques, such as dot voting or thermometer voting, can help to arrive at a decision if no consensus can be reached. Planning sessions help a team to decide on how to divide their time and workforce. And the famous Time Timer helps keeping things time-boxed.

6 Mapping of Scrum Meetings and Techniques from the Design Thinking Toolkit

By linking the collected requirements for each of the Scrum meetings and the goals andeffectsofthedifferenttypesoftechniquesintheDTtoolkitwederivedamapping between the two. Table 4 presents our theoretical mapping and the following sections provide explanations and examples for each entry.

6.1 Techniques Generally Applicable to Scrum Meetings

DT warm-ups, short energizing activities, can be employed before longer meetings to motivate participants. If several Scrum meetings are facilitated on the same day warm-ups can help the team transition from one meeting to the next and provide

Table 4 Mapping of Scrum meetings and Design Thinking techniques that could be applied to achieve the necessary meeting requirements and goals

Theoretical mapping of Scrum meeting and applicable Design Thinking techniques	
General	• Warm-Ups • Facilitation techniques
Product Backlog Refinement	• Artifacts from knowledge organization techniques • Artifacts from knowledge consolidation techniques • Artifacts from prototyping techniques • Unpacking techniques • Idea generation techniques • Low-fidelity prototyping techniques • Voting techniques
Sprint Planning	• Artifacts from knowledge organization techniques • Artifacts from knowledge consolidation techniques • Artifacts from prototyping techniques • Unpacking techniques • Idea generation techniques and their artifacts • Voting techniques
Daily Scrum	• Unpacking techniques • Knowledge sharing techniques • Establishing and visualizing meeting rules
Sprint Review	• Testing techniques • Feedback techniques
Retrospective	• Unpacking techniques • Knowledge sharing techniques • Feedback techniques • Knowledge organization techniques • Knowledge consolidation techniques • Idea generation techniques • Low fidelity prototyping techniques • Retrospective as DT workshops

a welcome break. On such meeting days, facilitation techniques such as check-ins and check-outs and establishing meeting rules can help the team transition into and out of the meeting work mode. Additionally the Time Timer can help staying within planned time boxes.

6.2 Product Backlog Refinement

During Refinement, developers need to understand the users' requirements, come up with solutions on how to implement these and document them. To support the understanding of the product vision as well as work items in general, artifacts from knowledge organization and consolidation techniques, which describe and document findings from user research, such as Personas, Journey Maps, Storyboards can be employed by the PO to help the team understand the underlying needs and business aspects. Similarly, artifacts from prototyping techniques such as wireframes and other low-fidelity prototypes can be employed by the PO to help the team understand aspects of the functionality. Additionally, when discussing how to implement a feature, unpacking methods such as charetting or creating a mind map can help to form a shared understanding of the feature. When discussing possible implementations, idea generation and prototyping techniques such as different types of brainstorming or sketching sessions can support the generation and discussion of possible solutions. The artifacts from these techniques, e.g., wireframes, diagrams, Journey Maps or Storyboards help to document the refined solution and can feedback to understanding the feature during planning or development. Finally, voting techniques can support estimation and prioritization of work items.

6.3 Sprint Planning

The main goal of this meeting is to make a plan and commit to it. As this is not a creative or problem-focused meeting, DT techniques are less helpful. However, some DT artifacts and techniques can support the meeting goals. As previously mentioned, several DT artifacts from knowledge organization and condensation as well as from prototyping techniques can support the PO in explaining the necessary work and can help the team understand the needs, business aspects, and functionality behind the feature. Similarly, to the Product Backlog refinement, unpacking techniques help to achieve a shared understanding if the feature is not clearly enough defined yet. Additionally, voting techniques, such as Dot Voting or Thermometer Voting, can support the team in prioritizing and deciding which work items to include in the following sprint. And finally, the team can make use of idea generation techniques when formulating and discussing the sprint goal.

6.4 Daily Scrum

The daily Scrum, as the short status update is another meeting that is not creative and problem-focused. However, preparing the Daily Scrum meeting by quickly unpacking recent achievements and identified impediments as a group or as individuals can help stay in time. Establishing rules similar to DT's *brainstorming rules* can ensure that discussions are deferred. A possible adaptation might be:

- Defer judgment and discussions.
- Stay focused on the topic.
- One conversation at a time.
- Observers defer talking.

Additionally, these rules can be visibly placed around the meeting area as commonly seen with brainstorming rules in DT workspaces.

6.5 Sprint Review

As this meeting focuses on inspecting results and collecting feedback, actual customer representatives or stakeholders are present at these meetings who might not be familiar with the details relevant to software development. DT testing techniques, such as the Think Aloud technique can support developers in eliciting targeted, actionable feedback. Additionally, feedback techniques such as the Feedback Capture Grid, the Empathy Map or the Five Finger Feedback can support customers in giving structured feedback, which can be used to adapt work items or gauge the business value of features.

6.6 Sprint Retrospective

Creative activities for Retrospectives are already commonplace (Jovanović et al., 2016) and the toolbox of DT techniques can add to these. The Retrospective focuses on improving the work process and, as such, a large part of the DT process and its techniques can be applied to this meeting.

In fact, entire Retrospectives can be facilitated in the form of a DT workshop, including the understand or empathize step to collect attendees thoughts, a clustering step to identify major feedback items, a voting to decide which issues to work on for the next Sprint, an ideation phase to come up with solutions as well as a prototyping phase to sketch out the solution. A testing phase implements the solution in the following Sprint. During such a DT workshop as well as during "normal" retrospectives unpacking, knowledge sharing, and feedback techniques can support the team in reflecting on the past sprint. Knowledge organization and condensation techniques help organize the feedback and identify improvements for the upcoming sprint. Idea generation and prototyping techniques can support the generation and documentation of action items.

7 Concrete Examples from Our Studies

After building a theoretical mapping between the Scrum meetings and their requirements and different categories of techniques from the DT toolkit we analyzed the data from our two original studies in order to derive concrete examples from practice and compare them to our mapping. For each of the Scrum meetings, Table 5 provides an overview of the DT techniques observed and mentioned in our study. As some of these techniques were not mentioned before, Table 6 briefly describes the specific DT techniques and artifacts mentioned in this section.

7.1 Techniques Generally Applicable to Scrum Meetings

The adherence to predefined time boxes for Scrum meetings can be supported by a common tool from DT workshops: the *Time Timer*. It is a large round analog timer that visually displays the passage of time as well as the remaining time using a continuously decreasing red disk. After a developer experienced the Time Timer for

Table 5 Mapping of Scrum meetings and Design Thinking techniques based on the data from our study

Practical experiences with Scrum meetings and applicable Design Thinking techniques	
General	• Strict time boxing with a TimeTimer
Product Backlog Refinement	• Storyboards • Establishing and visualizing Personas • Brainstorming to unpack features • Sketching / 30 seconds sketch
Sprint Planning	• Storyboards • Establishing and visualizing Personas • Otherwise not so useful
Daily Scrum	• Nothing mentioned
Sprint Review	• Usability Testing • Feedback Capture Grid • Cooperative idea generation
Retrospective	• Retrospective as DT workshops • 3 Ls • Clustering • Brainstorming • Storyboards

Table 6 Description of concrete DT methods and artifacts observed and mentioned in our studies

Method	Description
Time Timer	The Time Timer, a common tool to timebox DT workshops, is a clock that visually displays remaining time. *(facilitation)*
Storyboards	Linear sequence of quickly drawn illustrations which describe a story. *(knowledge organisation, knowledge condensation, prototyping)*
Personas	Fictional characters created for reference, based on user research, which represent key user types. *(knowledge condensation)*
30 Seconds Sketch	Effective technique to visualize and discuss a large quantity of ideas through several 30 second rounds of sketching. *(idea generation, prototyping)*
Usability Testing	Technique involving end-users in testing how easy a product can be used. *(testing)*
Feedback Capture Grid	Structured feedback collection using a 2x2 grid and the categories *what worked*, *what needs to change*, *new ideas* and *open questions*. *(feedback)*
3 Ls	Structured feedback collection using the categories of *liked*, *learned* and *lacked*. *(feedback)*
Brainstorming	Spontaneously generating thoughts in relation to a specific promp, alone or in a group. *(unpacking, idea generation)*
Clustering	Organizing findings from research into named clusters and visualizing relations between them. *(knowledge organization)*

the first time in a DT workshop he explained: "*That was a real aha moment. If we could have these for our meetings that would be great.*"

7.2 Product Backlog Refinement

One of the SMs from our study, who is also a developer, explained how he plans to introduce knowledge condensation techniques, to better describe features and achieve shared understanding when discussing backlog items: "*We will start using Storyboards for larger features, explaining why the user needs this feature and how it is connected to the solution. Something we usually ignore, but with such Storyboards, it can be done. Storyboards can serve as the vision for this feature. Additionally, I would like to have Personas in the backlog for orientation and then sort the backlog around these Personas and stories for orientation instead of the 4 levels of items we have now. Like this, we can achieve shared understanding early on.*" In a similar notion, another developer mentioned that he would like to have Personas around all the time to remember the users: "*We should print these Personas and have them always near, so we never forget who is the user of the things we do.*" A developer explained how his team uses a quick form of brainstorming to discuss features and achieve a shared understanding: "*If a task still has aspects that need clarification it can be interesting to do a short design thinking session. In the simplest case, we just do a short brainstorming to discuss what the feature means for each of us.*"

Concerning idea generation techniques during refinement, developers saw potential in using ideation techniques, but not in all cases. One developer stated: "*With sketching, [he talked about the 30 s sketch technique]]. It is easier to communicate one's ideas, and seeing them drawn makes it easier to modify and build upon them.* However, another developer noted: *It depends on the level of refinement. If all that remains are technical challenges, DT does not really help. If aspects of the feature are not clear yet then DT is a really interesting approach to refinement.*"

7.3 Sprint Planning

In our study, the process of planning the next items to work on was rarely mentioned in conjunction with DT techniques. As described in 7.2, one interviewee wanted to use storyboards and personas whenever discussing backlog items. In contrast, other study participants pointed out that the Scrum planning meeting relied on facts, such as the velocity of the team, the amount of work they forecast could be completed in the upcoming development iteration and the priority of work items. As one developer explained: *"For planning it's not a fit I think. It's about prioritizing and the capacity of the team; it's just not a creative process."*

7.4 Daily Scrum

The Daily Scrum meeting, as a regular check-in and team synchronization meeting, was not explicitly mentioned by interviewees as requiring the support of DT techniques. However, Daily Scrum meetings should ideally be short and rely on strict time boxes. As described in Sect. 6.1 the Time Timer can help teams with time keeping especially during this meeting.

7.5 Sprint Review

The use of testing techniques during review meetings to collect feedback from customers and stakeholders were mentioned several times in our studies. For example, one interviewee noted: *"Evaluation techniques—not traditional ones, where you ask questions and they answer, but, for example, you just give the system to the user and see how he interacts with it—that might help in getting better feedback.*"

Furthermore, feedback techniques were mentioned as being useful. One developer explained how a feedback capture grid not only helped to collect feedback but also led to cooperative ideation with the customer:*"I think it was good, and it helps to share different point of views. You can see what is liked and disliked, and the criticisms make you ask for questions and the questions give you some ideas. So there is a cycle for generating ideas."*

7.6 Sprint Retrospective

One of the SMs from our studies facilitates his Retrospectives as DT workshops: "*Running a feedback session with 30 people is chaos, it would be a whole day retro with nothing achieved. For large groups, more organization is needed so we collected feedback with the 3 Ls in the group, quickly clustered and then split up into smaller groups for the next steps. That's where DT really makes sense.*" As next steps, the smaller groups in this retrospective would brainstorm solutions and prototype them in a storyboard or sketch. A developer from his team describes his first experience with such a DT-Retrospective as follows: "*That big retro was a very nice experience. It was totally different from what we are used to. Normally, we just sit in a room, but here I find it was very interactive and I liked that people were open. I felt that most of the conclusions from that meeting are on their way now, so it really had a good impact on our team.*"

8 Discussion

In this chapter, we analyzed the integration of DT techniques into the meeting activities of the Scrum software development method from two perspectives.

First, we analyzed the five central Scrum meetings and reviewed their goals, requirements, and outcomes along with categories of DT techniques and the tasks they support. From the outcomes of this step, we derived a mapping of Scrum meetings and applicable DT techniques. In this theoretical mapping, we showed links between at least two categories of DT techniques with each of the five investigated Scrum meetings. We found the most matches between categories of DT techniques and the Retrospective, the Product Backlog Refinement, as well as the Sprint Planning meetings. This finding is in line with related research and practice for Scrum activities in meetings, which also concentrate on Retrospectives, Product Backlog Refinement and Sprint Planning.

In a second step, we analyzed data from two previous studies in which the topic of *DT techniques supporting Scrum meetings* emerged. This meta-analysis provided concrete examples from practice of how DT techniques support the facilitation of Scrum meetings. In our case study, we found examples of DT techniques supporting four out of the five Scrum meetings. For the remaining meeting, the Daily Scrum we could not observe any specific methods and the participants did not specifically mention the meeting in our interviews. For the Sprint Planning, some study participants specifically mentioned that they do not consider DT techniques useful for this event, while others considered artifacts from knowledge condensation within the backlog as useful to establish vision and shared understanding.

Comparing the two created mappings, we find several of our theoretical links confirmed by concrete examples from practice. Table 7 provides an overview of which entries could be confirmed and which entries could not be confirmed within this study. However, some of the categories of DT techniques were not mentioned or observed in our studies. For example, the anticipated general applicability of DT warm-up techniques for all Scrum meetings could not be confirmed as part of this study. This might be explain able by the perceptions of participants. Some of the better known warm-ups and activities, especially those that include elements of playfulness, are easily perceived as childish games or “fluff” by professionals (Kolko, 2018; Verity & McCullagh, 2012).

A noticeable difference between the mappings can be seen for the Sprint Planning meeting. While the meeting requirements, goals, and outcomes match with several DT activities, the data from our studies only supports one of these matches: the use of knowledge condensation techniques to explain the vision of a feature and support shared understanding. Additionally, participants even mentioned that they do not think DT activities are useful in this meeting. A possible explanation for this discrepancy might be found in the different ways planning meetings are facilitated. While in some teams Sprint Planning and Backlog Refinement are held as one meeting or as two meetings directly following each other, other teams separate the two meetings by a number of days. Thus, a Sprint Planning meeting which directly follows a Backlog Refinement probably focuses on planning and prioritizing action items and does not require further establishment of shared understanding. For a Sprint Planning that does not include or follow a Backlog Refinement the theoretically useful artifacts and techniques might be more relevant.

Another interesting finding was the mention of the 3Ls as a DT technique. This technique is commonly known as a retrospective activity for the data gathering stage, albeit as the 4Ls.[1] From our experience, this activity is less common in the DT community. While it is hard to say in which area activities or techniques were first introduced and used, we believe this is an indicator that the toolkit of Scrum meetings and the DT toolkit are compatible and seem to have merged in the eyes of the interviewee.

9 Conclusion

This study is based on two previous studies investigating DT use in software development teams but did not specifically focus on DT techniques for Scrum meetings.

[1] 4Ls Retro activity https://retromat.org/de/?id=78.

Table 7 Mapping of Scrum meetings and Design Thinking techniques highlighting our evaluation results. (light gray—mappings that could be confirmed with our data)

Comparison of theory and practice of Scrum meetings and applicable Design Thinking techniques	
General	• Warm-Ups • Facilitation techniques (timeboxing with TimeTimer)
Product Backlog Refinement	• Artifacts from knowledge organization techniques • Artifacts from knowledge consolidation techniques (persona, storyboard) • Artifacts from prototyping techniques • Unpacking techniques (brainstorming) • Idea generation techniques (30 seconds sketch) • Low-fidelity prototyping techniques (sketching) • Voting techniques
Sprint Planning	• Artifacts from knowledge organization techniques • Artifacts from knowledge consolidation techniques (persona, storyboard) • Artifacts from prototyping techniques • Unpacking techniques • Idea generation techniques and their artifacts • Voting techniques
Daily Scrum	• Unpacking techniques • Knowledge sharing techniques • Establishing and visualizing meeting rules
Sprint Review	• Testing techniques (usability testing) • Feedback techniques (feedback capture grid)
Retrospective	• Unpacking techniques • Knowledge sharing techniques • Feedback techniques (3 Ls) • Knowledge organization techniques (clustering) • Knowledge consolidation techniques • Idea generation techniques (brainstorming) • Low fidelity prototyping techniques (storyboards) • Retrospective as DT workshops

Accordingly, these previous studies featured an adequate sample of participating teams but provided only a small selection of concrete examples for DT techniques used in Scrum meetings. Further studies explicitly focused on this topic, can confirm additional entries in our mapping or can add additional new entries and concrete examples. Additionally, in this study, we investigated the five central Scrum meetings, which are generally applied in professional software development contexts. Several additional meetings exist, especially in different scaled versions of Scrum. One such example is the *Scrum of Scrums* meeting, introduced by one of the originators of the Scrum method (Sutherland, 2001), which provides a meeting with members of multiple development teams. Future work should, therefore, investigate the use of DT techniques in these scaled Scrum meetings.

Despite these limitations, our results indicate that integrating DT methods into Scrum meetings is a promising area of research with practical implications. Several of the DT activities that can be added to the toolbox of Scrum activities promote collaboration and a shared understanding within the development team. They help achieve meeting goals and prevent common Scrum meeting issues. As added activities to the Scrum toolkit, they provide a wider variety of methods for each Scrum meeting, promote diversity, and prevent monotonous meetings. Finally, development teams are confronted with additional use cases for DT techniques they might already be familiar with from DT workshops or DT project phases. Thus Scrum team members have the opportunity to further practice these DT techniques. However, not all Scrum meetings can or should be supported by DT in the same way, i.e., Sprint Planning and Daily Scrum can only marginally integrate DT techniques, while the Retrospective and the Product Backlog Refinement provide several opportunities to make use of DT techniques.

This research complements existing efforts to integrate DT and agile software development methods. Our results provide a starting point for practitioners to leverage DT's advantages, not only in an initial DT phase but also during regular development activities, in a similar fashion as proposed in the ad hoc DT model or the toolbox model. The mapping developed in this chapter presents an initial guideline on how to integrate individual DT techniques into existing Scrum meetings. It provides motivation and ideas on enhancing Scrum meeting routines with further techniques for meeting facilitators and teams: either identifying additional use cases for already familiar techniques or by pointing out new techniques worth investigating.

References

Akif, R., & Majeed, H. (2012). Issues and Challenges in Scrum Implementation. *International Journal of Scientific and Engineering Research, 3*(8), 1–4.

Baldauf, C. (2018). *Retromat - run great agile retrospectives!* Leanpub.com.

Brenner, W., Uebernickel, F., Abrell, T. (2016). Design Thinking as Mindset, Process, and Toolbox. In W.Brenner, F. Uebernickel (Eds.).*Design Thinking for Innovation*, (3–21). Springer International Publishing.

Brown, T. (2008). Design Thinking. *Harvard Business Review, 86*(6), 84.

Carleton, T., Cockayne, W., Tahvanainen, A. J. (2013). *Playbook for Strategic Foresight and Innovation - A hands-on guide for modeling, designing, and leading your company's next radical innovation.* Stanford.

Carlgren, L., Elmquist, M., & Rauth, I. (2014). Design Thinking: Exploring Values and Effects from an Innovation Capability Perspective. *The Design Journal, 17*(3), 403–423.

Cho, J. (2008). Issues and Challenges of Agile Software Development with Scrum. *Issues in Information System*, *IX*(2), 188–195.

Conboy, K., Coyle, S., Wang, X., & Pikkarainen, M. (Jul 2011). People Over Process: Key People Challenges in Agile Development. *IEEE Software, 28*(4), 48–57.

Deemer, P., Benefield, G., Larman, C., Vodde, B. (2012). *The Scrum Primer*. Version 2. available at:https://www.infoq.com/minibooks/Scrum_Primer/

Dobrigkeit, F., de Paula, D. (2019). Design Thinking in Practice: Understanding Manifestations of Design Thinking in Software Engineering. In *Proceedings of the 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE '19)*, (1059–1069). ACM Press.

Dobrigkeit, F., de Paula, D., Uflacker, M. (2018). InnoDev - A Software Development Methodology Integrating Design Thinking, Scrum and Lean Startup. In *Design Thinking Research - Looking Further: Design Thinking Beyond Solution-Fixation,* Understanding Innovation, 99–228. Springer.

Dobrigkeit, F., Pajak, P., de Paula, D., Uflacker, M. (2020). DT@IT Toolbox: Design Thinking Tools to Support Everyday Software Development. In *Design Thinking Research - Investigating Design Team Performance*, Understanding Innovation. Springer.

dos Santos Braz, R., Merlin, J. R., Trindade, D. F. G., Ribeiro, C. E., Sgarbi, E. M., de Sordi Junior, F. (2019). Design Thinking and Scrum in Software Requirements Elicitation: A Case Study. In: Marcus A., Wang W. (eds). Design, User Experience, and Usability. Design Philosophy and Theory. HCII 2019. Lecture Notes in Computer Science, Vol 11583. Springer

Esther, D., Larsen, D. (2006). *Agile Retrospectives: Making Good Teams Great*. Pragmatic Bookshelf.

Glomann, L. (2018). Introducing 'Human-Centered Agile Workflow' (HCAW) – An Agile Conception and Development Process Model. In T. Ahram and C. Falcão (eds.). Advances in Usability and User Experience, Advances in Intelligent Systems and Computing, Vol. 607, 646–655, Springer International Publishing.

Gray, D., Brown, S., Macanufo, J. (2010). *Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers*. O'Reilly Media, Inc., 1st edition.

Grossman-Kahn, B., Rosensweig, R. (2012). Skip the Silver Bullet: Driving Innovation through Small Bets and Diverse Practices. *Leading Innovation through Design: Proceedings of the DMI 2012 International Research Conference*, 815–829.

Gurusamy, K., Srinivasaraghavan, N., Adikari, S. (2016). An Integrated Framework for Design Thinking and Agile Methods for Digital Transformation. In A. Marcus (Ed.), *Design, User Experience, and Usability: Design Thinking and Methods,* Vol. 9746, 34–42. Springer International Publishing.

Haugen, N. C. (2006). An Empirical Study of Using Planning Poker for User Story Estimation. In *AGILE 2006 (AGILE'06)*, *2006*, 23–34. IEEE.

Hildenbrand, T., Meyer, J. (2012). Intertwining Lean and Design Thinking: Software Product Development from Empathy to Shipment. In A. Maedche, A. Botzenhardt, L. Neer (Eds.), *Software for People: Fundamentals, Trends and Best Practices*, Management for Professionals, 217–237. Springer.

Hohmann, L. (2006). *Innovation Games: Creating Breakthrough Products Through Collaborative Play*. Addison-Wesley.

IDEO. *IDEO Method Cards: 51 Ways to Inspire Design*. 2003.

Jovanović, M., Mesquida, A-L., Radaković, N., & Mas, A. (2016). Agile Retrospective Games for Different Team Development Phases. *Journal of Universal Computer Science, 22*(12), 1489–1508.

Kerth, N. L. (2000). The Ritual of Retrospectives: How to Maximize Group Learning by Understanding Past Projects. *Software Testing & Quality Engineering, 2*(5), 53–57.

Kniberg, H. (2015). *Scrum and XP from the Trenches - How we do Scrum*. InfoQ, 2nd edition.

Kolko, J. (Apr 2018). The Divisiveness of Design Thinking. *Interactions, 25*(3), 28–34.

Kua, P. (2013). *The Retrospective Handbook: A Guide for Agile Teams*. Leanpub.com.

Kumar, V. (2012). *101 Design Methods: A Structured Approach for Driving Innovation in Your Organization.* Wiley.

Lindberg, T., Köppen, E., Rauth, I., Meinel, C. (2012). On the Perception, Adoption and Implementation of Design Thinking in the IT Industry. In *Design Thinking Research - Studying Co-Creation in Practice*, Understanding Innovation, 229–240. Springer.

Martin, B., Hanington, B. (2012). *Universal Methods of Design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions.* Rockport Publishers.

Meyer, B. (2014). *Agile! The Good, the Hype and the Ugly*. (1st ed.). Springer International Publishing.

Ng, Y. Y., Skrodzki, J, Wawryk, M. (2020). Playing the Sprint Retrospective: A Replication Study. In *Advances in Agile and User-Centred Software Engineering,* Lecture Notes in Business Information Processing, 133–141,Springer International Publishing..

Pedersen, T. Ø. (2016). *UX Toolbox for Software Developers: Methods and Training*. Dissertation, Aalborg University.

Pereira, J. C., de F. S. M. Russo, R. (2018). Design Thinking Integrated in Agile Software Development: A Systematic Literature Review. *Procedia Computer Science*, *138*, 775–782.

Przybyłek, A., Kotecka, D. (2017). Making agile retrospectives more awesome. In 2017 Federated Conference on Computer Science and Information Systems (FedCSIS), 1211–1216. IEEE.

Ralph, P. (May 2013). Possible Core Theories for Software Engineering. In *2013 2nd SEMAT workshop on a general theory of software engineering (GTSE)*, 35–38. IEEE.

Rubin, K. S. (2012). *Essential Scrum: A practical guide to the most popular Agile process.* Addison-Wesley

Schwaber, K., Sutherland, J. (2017). *The scrum guide - The definitive guide to scrum: The rules of the game.* scrumguides.org.

Sohaib, O., Solanki, H., Dhaliwa, N., Hussain, W., Asif, M. (2018). Integrating design thinking into extreme programming. *Journal of Ambient Intelligence and Humanized Computing*, 10, 2497–2504.

Sutherland, J. (2001). Inventing and Reinventing SCRUM in Five Companies. *Cutter IT Journal, 14*, 5–11.

Sutherland, J., Schwaber, K. (2007). *The Scrum Papers: Nut, Bolts, and Origins of an Agile Framework*, Scrum, Inc.

Thoring, K., Müller, R. M. (2011). Understanding design thinking: A process model based on method engineering. In *Proceedings of E&PDE 2011, the 13th International Conference on Engineering and Product Design Education, London, UK, 08.-09.09. 2011*, 493–498.

Verity, J., McCullagh, K. (2012). Design Thinking. In *The New Strategic Landscape*, Vol. 37, 52–68. Palgrave Macmillan.

Wölbling, A., Krämer, K., Buss, C. N., Dribbisch, K., LoBue, P., Taherivand, A. (2012). Design Thinking: An Innovative Concept for Developing User-Centered Software. In *Software for People - Fundamentals, Trends and Best Practices,* Management for Professionals, 121–136. Springer.

Ximenes, B. H., Alves, I. N., Araújo, C. C. (2015). Software Project Management Combining Agile, Lean Startup and Design Thinking. In A. Marcus, (Ed.).*Design, User Experience, and Usability: design discourse*, Vol. 9186, 356–367. Springer International Publishing.

Examining the Neurocognitive Basis of Applied Creativity in Entrepreneurs and Managers

Jan Auernhammer, Neeraj Sonalkar, Hua Xie, Katerina Monlux, Jennifer Bruno, and Manish Saggar

Abstract Creativity and innovation lie at the heart of design thinking. Understanding how creativity is manifested in the brain is perhaps one of the most significant challenges of the twenty-first century. However, most of the previous research in understanding the neural basis of creativity has been limited to examining the neurocognitive basis of creative potential (or ability). Little is known about applied creativity in business (i.e., entrepreneurial creativity). At its core, entrepreneurship requires harnessing the power of a creative idea through discovery, evaluation, organization, and management to a final product/business. A design thinking mindset has been considered an essential component of successful entrepreneurship. In this study, we aim to better understand the cognitive basis of applied creativity in business by examining differences in entrepreneurs' neuropsychological profile as compared to that of matched (on age, gender, and socioeconomic status) administrators/managers. Our ongoing work could have a broad impact on the fields of academia and industry. The knowledge generated from this project may provide a cornerstone in how we train (or design) the next generation of entrepreneurs and managers.

J. Auernhammer · N. Sonalkar
Center for Design Research, Stanford University, Building 560, 424 Panama Mall, Stanford, CA 94305, USA
e-mail: jan.auernhammer@stanford.edu

N. Sonalkar
e-mail: sonalkar@stanford.edu

H. Xie · K. Monlux · J. Bruno · M. Saggar (✉)
Psychiatry and Behavioral Sciences—Center for Interdisciplinary Brain Sciences Research, 401 Quarry Rd, St 1356, Stanford, CA 94305, USA
e-mail: saggar@stanford.edu

H. Xie
e-mail: huaxie@stanford.edu

K. Monlux
e-mail: kmonlux@stanford.edu

J. Bruno
e-mail: jenbruno@stanford.edu

C. Meinel and L. Leifer (eds.), *Design Thinking Research*, Understanding Innovation,
https://doi.org/10.1007/978-3-030-76324-4_18

1 Introduction

Creativity and innovation lie at the heart of design thinking. Understanding how different cognitive processes facilitate creativity is perhaps one of the most significant challenges of the twenty-first century. A large amount of cognitive neuroscience research has been done to understand how we produce creative ideas and the basis of the creative process itself (Dietrich & Kanso, 2010). In our recent works, we have provided evidence to suggest that creative capacity can be enhanced using targeted design thinking-based training (Bott et al., 2013; Kienitz et al., 2014; Saggar et al., 2017). Furthermore, brain development and upbringing can affect creative potential in young children (Saggar et al., 2019), and creative teams have higher synchrony in the brain areas related to the Theory of Mind (Xie et al., 2020). However, most of the previous research in understanding the neurocognitive basis of creativity has been limited to examining the neural basis of creative potential—with very little known about applied creativity.

Applied creativity, sometimes also referred to as innovation, as suggested by (Sheppard et al., 2013), is very different from mere creative potential. One has to combine the intrinsic factors (creative potential and confidence, etc.) with other extrinsic factors (need-finding, communication, building a design solution, etc.) to push the idea to its fruition and sustain it. Although, researchers have studied the cognitive basis of applied creativity in the domains of music and art, e.g., (Csikszentmihalyi, 1997; Limb, 2010; Loui et al., 2011; Zaidel, 2014). Very little is known about the cognitive basis of applied creativity in business, particularly during business formation. We also know little about the cognitive basis of applied creativity in entrepreneurs, who often use a design thinking mindset to bring breakthrough innovations to society.

Here, we use the definition of (Eckhardt & Shane, 2010) to define entrepreneurship,

> *"... as the discovery, evaluation, and exploitation of future goods and services ... [by] ... creation or identification of new ends and means previously undetected or unutilized by market participants."*

At its core, entrepreneurship requires harnessing the power of a creative idea through discovery, evaluation, organization, and management and bringing it to a final product/business. A design mindset has been considered as an essential component of successful entrepreneurship (Duening, 2010).

In the project funded by the Hasso Plattner Design Thinking Research Program (HPDTRP) for the year 2019–2020, we aimed at studying the neurocognitive basis of applied creativity in business by examining the differences in brain structure and cognitive function of entrepreneurs as compared to those of matched (on age, gender, and socioeconomic status) administrators/managers. Given the geographical environment, we are situated in (i.e., Silicon Valley and Stanford University)—we were confident in terms of tapping individuals at different levels of experience in entrepreneurship and administration.

At the beginning of 2020, due to the COVID-19 pandemic and associated restrictions on in-person data collection, we pivoted to only collecting neuropsychological data online. Although this pivot provided an exciting vantage point for measuring cognitive differences in entrepreneurs and managers' cognitive differences remotely, we could not measure originally proposed neural differences using functional magnetic resonance imaging (fMRI).

In the current chapter, to establish a background for this work, we first provide definitions of entrepreneurs and managers, how we measured success in each profession, and previous work in this area. We then present our innovative research design for collecting neuropsychological assessments online (via video conferencing and surveys). Lastly, we describe plans for data analytics and the future impact on design thinking research.

2 Background and Hypothesis

2.1 Identifying Entrepreneurs and Managers

One of the main challenges was to appropriately identify entrepreneurs and managers while also considering their characteristics, context, and success. This sub-section outlines the development of the measurement instrument for categorizing and assessing entrepreneur vs. manager characteristics in study subjects.

An entrepreneur is typically a person who establishes, organizes, and operates a new business (Laureiro-Martínez et al., 2014). There are many different types of individuals who establish, organize, and operate a business, such as self-employed individuals, owner-operators, lifestyle entrepreneurs, and large business founders. For this study, we focused on entrepreneurs who have established at least one business with two or more people. Such individuals require both creating a new business and collaborating with others. Similarly, managers are people who are responsible for controlling and administrating parts of the organization. However, managers can manage a project team or be responsible for a specific product. In this study, we focused on managers who are responsible for managing two or more people. Such individuals require organizational behavior for people and specific business aspects.

The specific selection of entrepreneurs and managers gives us the opportunity to examine the differences in their neurocognitive characteristics. An entrepreneur takes risks and creatively establishes a business, while a manager organizes given resources analytically. However, real-world conditions do not permit a clear separation of these two groups as individuals so one can be an *entrepreneur and manager simultaneously* or may have been one or the other in their past. Therefore, identifying individuals who are only in one group is essential. This examination of study subjects allows the identification of an individual who is clearly either an entrepreneur or a manager. Tables 1 and 2 show the survey designed to identify where each *individual lies on the spectrum* between entrepreneurs and managers. This information was used to

Table 1 Survey to identify entrepreneurial quotient

Entrepreneur		
El	How many businesses (with 2 or more people employed) have you started?	– ———number of (co-)founded businesses
E2a	Please elaborate on your motivations for starting a business.	– —————
E2b	Check all the motivation that apply	☐ Working with a team ☐ Secure career ☐ Variety of careers ☐ Role in contribution to community welfare ☐ High prestige ☐ Meaningful work ☐ Desire for self-determination ☐ Opportunity to innovate ☐ Desire for responsibility ☐ Challenging task ☐ Economic gain
E3	In the last five years, how many years have you been self-employed, founder of a company with 2 or more people and/or being employed as a manager?	(a) Self-employed———(years) (b)Founder———(years) (c)Manager———(years)
E4	In the last five years, did you work for two jobs at the same time? For how long? (e.g., (c) employed and (a) self-employed for 2 years)	
E5	If you were employed in the last five years, what was your most recent position?	Position:——— Active:———(months) Supervision:———(No of employees)
E6	How many other people are there between you and the CEO in your current position? (e.g. CEO = 0; CTO = −1; Senior Vice President = −2)	Your level counting down from CEO:——— Total level in company:———
E7	Do you consider yourself as a successful entrepreneur?	(not successful)—(very successful) ☐ ☐ ☐ ☐ ☐
E8	Who would be your ideal role model of an entrepreneur you want to be?	

quantify an individual's position on that spectrum via the entrepreneurial-managerial quotient.

Table 2 Survey to identify managerial quotient

Manager		
Ml	How many managerial positions (responsible for managing people) have you had in your career?	– ————number of managerial positions held
M2a	Please elaborate on your motivations for a managerial career.	
M2b	Check all the motivation that apply.	☐ Working with a team ☐ Secure career ☐ Variety of careers ☐ Role in contribution to community welfare ☐ High prestige ☐ Meaningful work ☐ Desire for self-determination ☐ Opportunity to innovate ☐ Desire for responsibility ☐ Challenging task ☐ Economic gain
M3	In the last five years, how many years have you been self-employed, founder of a company with 2 or more people and/or being employed as a manager?	(a) Self-employed———(years) (b) Founder———(years) (c) Manager———(years)
M4	In the last five years, did you work for two jobs at the same time? For how long? (e.g., (c) employed and (a) self-employed for 2 years)	– ———————
M5	If you were employed in the last five years, what was your most recent position?	Position:——— Active:———(months) Supervision:———(No of employees)
M6	How many other people are there between you and the CEO in your current position? (e.g., CEO = 0; CTO = −1; Senior Vice President = −2)	Your level counting down from CEO:——— Total level in company:———
M7	Do you consider yourself as a successful entrepreneur?	(not successful)—(very successful) ☐ ☐ ☐ ☐ ☐
M8	Who would be your ideal role model of an entrepreneur you want to be?	

2.2 Assessing Entrepreneurial and Managerial Success

The study also examined the motivation and success within the two groups (entrepreneurs and managers). The motivation for being an entrepreneur and manager could differ, as could the objective for success. Different examples of motivation for entrepreneurship could be, for instance, the "4-h work week," wealth, and/or making the world a better place. This motivation will lead to different decisions. For this reason, the study identified each individual's personal motivation and success in being an entrepreneur or manager.

Entrepreneurs perceive entrepreneurial success as the presence of both personal and macro-level variables, such as the need for achievement, locus of control, self-leadership, profitability, and innovation (Fisher, Maritz, & Lobo, 2014). A standard macro-level measurement for entrepreneurial success in business and management academic discourse is the organizational performance, measured using factors such as size, growth, and profitability (Combs et al., 2005; Eisenhardt and Schoonhoven 1990; Unger et al., 2011). This assessment provides the organizational context in which both entrepreneurs and managers operate; personal success includes income, compensation, share value, and the number of employees supervised. Table 3 shows the survey designed to identify the motivation and success of entrepreneurs and managers.

2.3 Previous Work in Identifying Neurocognitive Differences in Entrepreneurs Versus Managers

Very few studies have attempted to understand the neurocognitive differences between entrepreneurs and managers. Further, researchers have just begun to tackle this question using neuroimaging. For example, (Laureiro-Martínez et al., 2014) focused on cognitive reasoning and compared decision-making efficiency between entrepreneurs and managers. This study used a 4-armed bandit task to assess the exploitative vs. explorative decision-making behavior in entrepreneurs and managers. In this gambling task, the participants chose one among four slot machines with different gains to ultimately maximize payoffs within a given number of trials. They found that entrepreneurs were more efficient in decision-making than managers, as entrepreneurs obtained the same profit much faster. The researchers also found explorative behavior (i.e., switching slot machines) elicited stronger activation in the dorsal sector of the anterior cingulate cortex and the locus coeruleus, while exploitative behavior evoked stronger activation in the medial prefrontal cortex and hippocampus. However, no group effect (entrepreneurs vs. managers) nor interaction effect between choice type and the group was found at the whole-brain level.

Besides decision-making efficiency, another critical aspect of entrepreneurial cognition is *affected*. An fMRI study by (Halko et al., 2017) recently investigated whether entrepreneurs' emotional/brain responses in viewing their own firms resemble those of parents toward their own children. The participants viewed the pictures of their own children/venture logos vs. those of other familiar children/logos. They found entrepreneurs who self-rated as closely attached to their venture shared a similar suppression of activity in the posterior cingulate cortex, temporoparietal junction, and dorsomedial prefrontal cortex as did fathers when viewing pictures of their logos/children. Their result suggested a similar brain structure responsible for reward and emotional processing underly entrepreneurial and parental love.

Table 3 Entrepreneurial-managerial success factor scale

	Record ID	ID:————
Personal		
1	What are your top three personal success indicators?	first———— second———— third————
la	Growth in the first personal success factors (last 5 years)?	(not even close)-(achieved) ☐ ☐ ☐ ☐ ☐
lb	Growth in the second personal success factors (last 5 years)?	(not even close) — (achieved) ☐ ☐ ☐ ☐ ☐
lc	Growth in the third personal success factors (last 5 years)?	(not even close)—(achieved) ☐ ☐ ☐ ☐ ☐
Organizational		
2	What is the name of your company?	– ———
3	How many employees work currently in your company?	– ———(No of employees)
4	What is the current sales revenue of your company?	– ———(US$ sales revenue)
5	In comparison to companies in the same sector, how successful would you rate your company?	(lower end)—(market leader) ☐ ☐ ☐ ☐ ☐
Individual		
6	Identify your personal annual income area (including stock options).	☐ < 70,000
		☐ 70,001–100,000
		☐ 100,001–130,000
		☐ 130,001–160,000
		☐ 160,001–200,000
		☐ 200,000–300,000
		☐ 300,001–400,000
		☐ 400,001–500,000
		☐ 500,001–1M
		☐ 1M–2M
		☐ >$2M
		☐ ($ values)
8	How many people do you supervise?	(No of employees)
9a	In comparison to individuals with the same or similar role within your company, how successful would you rate yourself?	(bottom 5%)–(top 5%) ☐ ☐ ☐ ☐ ☐

(continued)

Table 3 (continued)

	Record ID	ID:————
9b	In comparison to individuals with the same or similar role across similar companies, how successful would you rate yourself?	(lower end)—(very successful) ☐ ☐ ☐ ☐ ☐
10	What are your current company shares worth?	☐ < 100,000 ☐ 100,001-500,000 ☐500,000–1M ☐1M-10M ☐10M-50M ☐50M-100M ☐> 100M ☐($ values)

A risk-taking propensity is also an ideal starting point for the neuroscientific query of entrepreneurial cognition (Krueger & Welpe, 2014). Risk propensity supports the business foundation, and a significantly higher risk propensity is found in entrepreneurs as compared with managers in numerous studies; for a meta-analysis, see (Stewart & Roth, 2001).

In sum, recent previous research has focused on decision efficiency, emotional reasoning, and risk propensity in entrepreneurs and managers.

3 Research Design

We aimed at taking a comparative cognitive approach in understanding how applied creativity differs between entrepreneurs and managers. We specifically focused on differentiating between the two groups based on cognitive abilities (attention, cognitive control, working memory, etc.), social processing (understanding others), negative valence (fear, anxiety, etc.), positive valance (reward responsiveness, etc.), and sensorimotor systems. As noted earlier, we directly compared the two groups of entrepreneurs and managers, and we also treated the dichotomy as an entrepreneurial/managerial spectrum performing a parametric analysis.

3.1 Overall Approach and Hypothesis

To examine differences in brain functioning between entrepreneurs and managers, we aimed at using a novel approach to mine all five major domains of human cognition. Researchers have now taken a similar approach by trying to find biological markers of mental functioning in health and disease (Cuthbert & Insel, 2013). Specifically, we aimed at examining the following five major domains of human functioning to

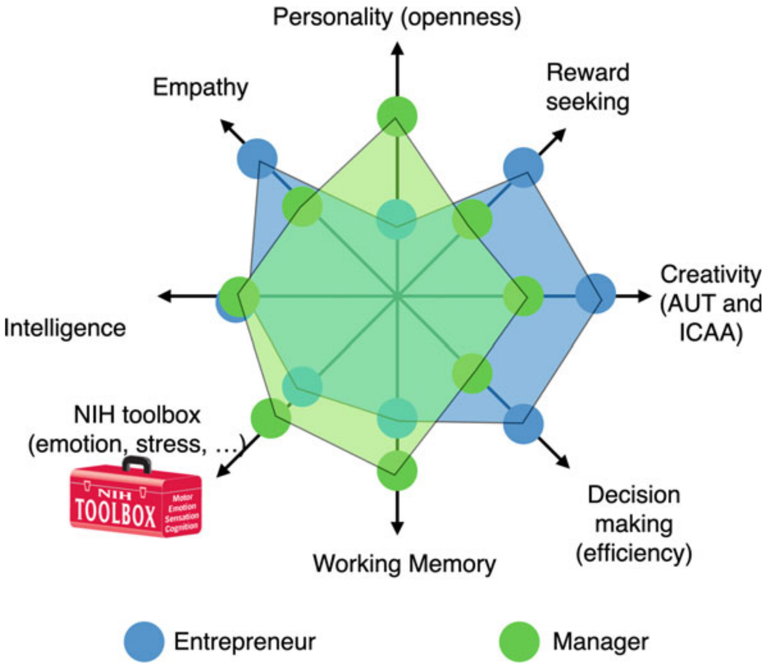


Fig. 1 Spider chart showing sample neurocognitive profiles for a *hypothetical* entrepreneur and manager. By assessing multiple dimensions, we aimed at capturing a holistic cognitive profile for each participant

construct a neurocognitive profile for each participant, as shown in Fig. 1. Such a novel approach will allow us to perform a comprehensive examination of how entrepreneurs differ from managers regarding overall and domain-specific brain functioning.

a. *Negative valence*: Negative valence systems are primarily responsible for responses to aversive situations or contexts, such as fear, anxiety, and loss. Our working hypothesis was that entrepreneurs' brain responses would differ from managers in the sense that entrepreneurs would be less responsive to aversive situations.
b. *Positive valence*: Positive valence systems are primarily responsible for responses to positive motivational situations or contexts, such as reward-seeking, consummatory behavior, and reward/habit learning. For this dimension, our working hypothesis was that the entrepreneurs would seek higher rewards as compared to managers.
c. *Cognitive systems*: This dimension includes a series of sub-components, such as attention, cognitive control, perception, working memory, and creative potential. Here, our working hypothesis was that entrepreneurs would putatively outperform managers with respect to cognitive subsystems of perception, attention, and creativity. However, managers would outperform entrepreneurs with respect

to cognitive systems that require higher response inhibition and impulse control (e.g., cognitive control).

d. *Social processing*: Systems for social processes mediate responses in interpersonal settings of various types, including perception and the interpretation of others' actions. Here, our working hypothesis was that entrepreneurs' neural responses to social stimuli would be higher than that of managers, given that entrepreneurs are anecdotally required to be persuasive and empathetic.

e. *Sensorimotor systems*: Sensorimotor systems are primarily responsible for controlling and executing motor behaviors and their refinement during learning and development. Here, our working hypothesis was that entrepreneurs' neural responses to sensorimotor stimuli might elucidate differences in low-level motor planning.

3.2 Participant Recruitment

Participants were recruited by word of mouth, on the Stanford Business listserv, LinkedIn, and Craigslist using targeted postings. All study flyers and recruitment emails were pre-approved by Stanford University's Institutional Review Board. Potential participants were invited to fill out a brief screening form with minimal demographics information, and, if they met the inclusion criteria, they were invited to complete a series of online assessments.

Inclusion criteria were split into two categories, managers and entrepreneurs. In order to qualify as a manager, potential participants would need to be between the ages of 18 and 45 years, a person responsible for managing a part of a company or organization, and currently managing a team of over two employees. To qualify as an entrepreneur, potential participants would also need to be between 18 and 45 years old. In addition, they also needed to be the founder of a company and that company needed to have more than two employees.

Potential participant responses were reviewed by study personnel, and eligible participants were contacted by email for enrollment. Interested participants were invited to complete a series of online surveys and assessments delivered over secure video conferencing.

3.3 Data Collection

Due to the COVID-19 pandemic, all study procedures took place remotely. Data were collected using HIPAA compliant video conferencing as well as an encrypted survey database. Participants first completed a brief online survey and were later invited to complete additional assessments administered over video application and self-paced surveys. Assessments administered over video included intelligence assessment (WASI-II subscales) as well as the working memory sub-domain of the NIH Toolbox application.

3.3.1 Survey Administration

Surveys that were self-administered online included:

- A brief screening form
- Minimal risk consent form
- Demographics form
- Toronto empathy questionnaire
- Entrepreneur-manager quotient
- Melbourne decision-making questionnaire
- NEO-FFI personality scale
- Inventory of creative activities and achievements
- Reward responsiveness questionnaire
- General self-efficacy scale
- Alternative uses test
- NIH toolbox subscales (emotion, friendship, perceived stress, and emotional support).

These assessments took roughly an hour to complete, and participants took breaks as needed.

The Toronto empathy questionnaire, based on (Spreng et al., 2009), is a 16-question brief self-report measure to assess empathy. TEQ assesses empathy as an emotional process, and the individual answers questions pertaining to empathy with the categories of "Never, rarely, sometimes, often, and always."

The entrepreneur-manager quotient was developed specifically for this study and was a self-assessment in which the individual would provide information pertaining to managerial and entrepreneurial qualities (see Table 1).

The Melbourne decision-making questionnaire, as outlined by (Mann et al., 1997), is a 31 question self-report questionnaire. This inventory was designed to measure tendencies to use three major coping patterns: vigilance, hypervigilance, and defensive avoidance (procrastination, buck-passing, and rationalization).

The NEO-five factor inventory (NEO-FFI), as described by (Whiteman et al., 2001), has 60 items broken up into 12 per domain type. This is derived from the original 240 items, and individuals are invited to rate whether they strongly disagree, disagree, are neutral, agree, and strongly agree with the question items.

The inventory of creative activities and achievements (ICAA), based on (Diedrich et al., 2018), is a broad assessment developed to tease apart individual differences in creativity. Individuals filling out the ICAA indicate scales and frequencies of participating in creative activities and the level of achievement across eight creative domains.

The reward responsiveness scale (RR), as outlined by (Berg et al., 2010), is a 14 item self-assessment used to determine an individual's level of responsiveness to rewards. Each question is rated as a strong disagreement, mild disagreement, mild agreement, and strong agreement to different reward categories.

The general self-efficacy scale (GSE), as expressed by (Schwarzer & Jerusalem, 1995), is a 10 item self-rating scale where each question assesses successful coping mechanisms and measures of success. Individuals indicate on a 4-point scale whether each statement is not at all true, hardly true, moderately true, or exactly true for them.

The alternative uses task (AUT), as outlined by (Guilford, 1967), asks the individual to list as many alternatives uses for a list of standardized items as they can think of in 2 min. Each use must differ from the other uses and be different from the standard use to count into the score, which assesses creativity.

Finally, measures from the National Institute of Health Toolbox, such as emotions, friendship, perceived stress, and emotional support scales, as discussed by (Heaton et al., 2014), were administered.

3.3.2 Video Conference Administration

To administer intelligence assessment, WASI-II was employed over video conferencing. The vocabulary section was administered with the assessor reading off the words and pointing to them using a cursor. Next, the matrix reasoning section was completed. Participants were invited to take as many breaks as needed and to sit in a part of their home that was comfortable and free of distractions. After the WASI-II administration, a short break was typically taken while the assessor set up the NIH toolbox application for the working memory subtest. The application was opened on the iPad with the screen broadcast over Zoom to run the NIH toolbox assessment. A Bluetooth keyboard, as well as a Bluetooth speaker, were connected to the iPad for administration, and the assessor read off the instructions while at the same time handling the keyboard to manually enter the participants' responses.

4 Challenges and Opportunities

Here, we briefly outline the challenges our team faced last year to accomplish the goals of this project.

4.1 Identifying and Characterizing Entrepreneurs and Managers

As already discussed in Sect. 2.1, one of the main challenges was to appropriately identify entrepreneurs and managers while considering their characteristics, context, and success. To address this challenge, we developed a new psychometric instrument—entrepreneur-manager quotient (EnMgQ), which takes a parametric approach instead of a dichotomous one. Further, we also developed an instrument to consider individual differences in terms of success, i.e., how do individuals measure their own success. Lastly, we aim to use a parametric modulation analysis approach for all data analytics, where instead of dichotomy, we plan to use the degree of EnMgQ as an anchor to study individual differences in cognitive abilities.

4.2 Converting In-Person Data Collection Tools to Online-Only Data Collection

Due to the COVID-19 pandemic, no data were collected in-person. All the behavioral and neuropsychological assessments were performed online using surveys and video-conferencing options. The pandemic posed an immense challenge for both recruitment and data collection. Technically, we had to shift entire data collection to an online-only format that resulted in additional personnel time and costs in setting up online data collection, training, piloting, and equipment purchase. Further, all surveys were set up electronically in a secure environment for data collection. Lastly, we attempted to keep the burden for participants to a minimum. Participants only needed a reliable Wi-Fi connection, a device that could connect to the Internet, chargers, and headphones. Instructions were sent to the participants prior to starting remote sessions.

4.3 Data Analytics

Given the large amounts of multidimensional data that has been collected, we plan to use both traditional statistical tools and recently developed unsupervised tools, such as topological data analysis (TDA; (Saggar et al., 2018; Geniesse et al., 2019)) to generate data-driven insights.

5 Future Work and Impact

5.1 Implications for Design Thinking Research

Our work's primary implication is to situate applied distinctions such as "entrepreneur" and "manager" either at the individual or at the social level. We hypothesized that if we were to find significant differences in how managers and entrepreneurs think, the results would indicate how professional distinctions such as manager and entrepreneur map onto cognitive and neural traits. However, if significant group differences are not observed, then the results would indicate that these professional distinctions are meaningful only in the social and organizational context of use and that they do not necessarily characterize cognitive differences at the individual level. Previous work provides evidence for the existence of an "entrepreneurial mindset" (Davis et al., 2016; Haynie et al., 2010; Mauer et al., 2017). However, it may turn out that the entrepreneurial mindset is not limited to entrepreneurs. Today corporate environments, especially in high-tech companies, may necessitate managers to acquire an "entrepreneurial mindset," which corresponds with a design thinking mindset, to successfully deal with the complexity and uncertainty inherent in every profession. Thus, our results could hold direct implications for studying entrepreneurship as a phenomenon and its role in modern society.

5.2 Academic Impact

The development of a neurocognitive profile could serve as a benchmark diagnostic for measuring the development of entrepreneurial creativity. Researchers could utilize neural markers to test various theories, frameworks, and even tools to develop applied creativity in practice, which can impact how we develop and test curricula in design education and entrepreneurial education.

5.3 Industry Impact

Innovation and creativity are recognized as a source of strategic advantage in industry. However, the tools for fostering such applied creativity are based on best practices and hype, which do not translate into success across context differences. Characterizing neurocognitive profiles has the potential to provide industry with a person-specific and context-dependent neural marker for both developing and evaluating the effectiveness of design thinking and creativity applications. This has special significance to the world of venture investing, which could greatly benefit from the use of targeted tools for improving entrepreneurial creativity.

5.4 Community Impact

Mapping out the neurocognitive profile for various levels of creativity in the real world could not only help individuals optimize their skills but could also help create novel design thinking-based training paradigms that are optimal in transitioning people from one profile to the next. The design thinking research and practice communities could use such diagnostics to evolve a NeuroDesign paradigm that closely ties design behavior and human biology for a real-world impact.

References

Bott, N., Quintin, E.M., Saggar, M., Kienitz, E., Royalty, A., Hong, D., Liu, N., Chien, Y.H., Hawthorne, G., Reiss, A.L. (2013). "Creativity training enhances self-directed attention and information processing." In *Poster presented at the annual meeting of the cognitive neuroscience society.*

Combs, James G., T. Russell Crook, and Christopher L. Shook. 2005. "The Dimensionality of Organizational Performance and Its Implications for Strategic Management Research." In . https://doi.org/10.1016/s1479-8387(05)02011-4.

Csikszentmihalyi, M. (1997). *Creativity: Flow and the psychology of discovery and invention.* HarperCollins Publications.

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine, 11*(1), 126. https://doi.org/10.1186/1741-7015-11-126

Davis, Mark H., Jennifer A. Hall, and Pamela S. Mayer. 2016. "Developing a New Measure of Entrepreneurial Mindset: Reliability, Validity, and Implications for Practitioners." *Consulting Psychology Journal* 68 (1). https://doi.org/10.1037/cpb0000045.

Diedrich, Jennifer, Emanuel Jauk, Paul J. Silvia, Jeffrey M. Gredlein, Aljoscha C. Neubauer, and Mathias Benedek. 2018. "Assessment of Real-Life Creativity: The Inventory of Creative Activities and Achievements (ICAA)." *Psychology of Aesthetics, Creativity, and the Arts* 12 (3). https://doi.org/10.1037/aca0000137.

Dietrich, A., & Kanso, R. (2010). A Review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychological Bulletin, 136*(5), 822–848. https://doi.org/10.1037/a0019749

Duening, Thomas N. 2010. "Five Minds for the Entrepreneurial Future: Cognitive Skills as the Intellectual Foundation for next Generation Entrepreneurship Curricula." *Journal of Entrepreneurship* 19 (1). https://doi.org/10.1177/097135570901900101.

Eckhardt, J. T., & Shane, S. (2010). An Update to the individual-opportunity nexus. *Handbook of Entrepreneurship Research.* https://doi.org/10.1007/978-1-4419-1191-9_3

Eisenhardt, Kathleen M., and Claudia Bird Schoonhoven. 1990. "Organizational Growth: Linking Founding Team, Strategy, Environment, and Growth Among U.S. Semiconductor Ventures, 1978–1988." *Administrative Science Quarterly* 35 (3). https://doi.org/10.2307/2393315.

Fisher, Rosemary, Alex Maritz, and Antonio Lobo. 2014. "Evaluating Entrepreneurs' Perception of Success: Development of a Measurement Scale." *International Journal of Entrepreneurial Behaviour and Research* 20 (5). https://doi.org/10.1108/IJEBR-10-2013-0157.

Geniesse, Caleb, Olaf Sporns, Giovanni Petri, and Manish Saggar. 2019. "Generating Dynamical Neuroimaging Spatiotemporal Representations (DyNeuSR) Using Topological Data Analysis." *Network Neuroscience* 3 (3). https://doi.org/10.1162/netn_a_00093.

Guilford, J P. 1967. "The Nature of Human Intelligence." http://psycnet.apa.org/psycinfo/1967-35015-000.

Halko, Marja Liisa, Tom Lahti, Kaisa Hytönen, and Iiro P. Jääskeläinen. 2017. "Entrepreneurial and Parental Love—Are They the Same?" *Human Brain Mapping* 38 (6). https://doi.org/10.1002/hbm.23562.

Haynie, J. Michael, Dean Shepherd, Elaine Mosakowski, and P. Christopher Earley. 2010. "A Situated Metacognitive Model of the Entrepreneurial Mindset." *Journal of Business Venturing* 25 (2). https://doi.org/10.1016/j.jbusvent.2008.10.001.

Heaton, Robert K., Natacha Akshoomoff, David Tulsky, Dan Mungas, Sandra Weintraub, Sureyya Dikmen, Jennifer Beaumont, et al. 2014. "Reliability and Validity of Composite Scores from the NIH Toolbox Cognition Battery in Adults." *Journal of the International Neuropsychological Society* 20 (6). https://doi.org/10.1017/S1355617714000241.

Kienitz, Eliza, Eve-Marie E.-M. Quintin, Manish Saggar, Nicholas T N.T. Bott, Adam Royalty, Daniel Wei-Chen D.W.C. Hong, ning liu, Y.-H. Yin-hsuan Chien, Grace Hawthorne, and Allan L Reiss. 2014. "Targeted Intervention to Increase Creative Capacity and Performance: A Randomized Controlled Pilot Study." *Thinking Skills and Creativity* 13 (0): 57–66. https://doi.org/10.1016/j.tsc.2014.03.002

Krueger, Norris, and Isabell Welpe. 2014. "Neuroentrepreneurship: What Can Entrepreneurship Learn from Neuroscience?" In *Annals of Entrepreneurship Education and Pedagogy - 2014*. https://doi.org/10.4337/9781783471454.00011.

Laureiro-Martínez, Daniella, Nicola Canessa, Stefano Brusoni, Maurizio Zollo, Todd Hare, Federica Alemanno, and Stefano F. Cappa. 2014. "Frontopolar Cortex and Decision-Making Efficiency: Comparing Brain Activity of Experts with Different Professional Background during an Exploration-Exploitation Task." *Frontiers in Human Neuroscience* 7 (JAN). https://doi.org/10.3389/fnhum.2013.00927.

Limb, C. J. (2010). Neural substrates of spontaneous musical improvisation. *The Journal of the Acoustical Society of America, 127*(3), 1950. https://doi.org/10.1121/1.3384930

Loui, P., Charles Li, H., Hohmann, A., & Schlaug, G. (2011). Enhanced cortical connectivity in absolute pitch musicians: a model for local hyperconnectivity. *Journal of Cognitive Neuroscience, 23*(4), 1015–1026. https://doi.org/10.1162/jocn.2010.21500

Mann, Leon, Paul Burnett, Mark Radford, and Steve Ford. 1997. "The Melbourne Decision Making Questionnaire: An Instrument for Measuring Patterns for Coping with Decisional Conflict." *Journal of Behavioral Decision Making* 10 (1). https://doi.org/10.1002/(SICI)1099-0771(199703)10:1<1::AID-BDM242>3.0.CO;2-X.

Mauer, René, Helle Neergaard, and Anne Kirketerp Linstad. 2017. "Self-Efficacy: Conditioning the Entrepreneurial Mindset." In . https://doi.org/10.1007/978-3-319-45544-0_.

Saggar, Manish, Eve-Marie E.-M. Marie Quintin, Nicholas T. N.T. Bott, Eliza Kienitz, Y.-H. Yin-hsuan Hsuan Chien, Daniel W-C D.W.-C. W.C. Hong, Ning liu, Adam Royalty, Grace Hawthorne, and Allan L. Reiss. 2017. "Changes in Brain Activation Associated with Spontaneous Improvization and Figural Creativity After Design-Thinking-Based Training: A Longitudinal FMRI Study." *Cerebral Cortex (New York, N.Y. : 1991)* 27 (7): 3542–52. https://doi.org/10.1093/cercor/bhw171.

Saggar, Manish, Olaf Sporns, Javier Gonzalez-Castillo, Peter A. P.A. Bandettini, Gunnar Carlsson, Gary Glover, and Allan L. Reiss. 2018. "Towards a New Approach to Reveal Dynamical Organization of the Brain Using Topological Data Analysis." *Nature Communications* 9 (1). https://doi.org/10.1038/s41467-018-03664-4.

Saggar, M., Xie, H., Beaty, R. E., Stankov, A. D., Schreier, M., & Reiss, A. L. (2019). Creativity slumps and bumps: examining the neurobehavioral basis of creativity development during middle childhood. *NeuroImage, 196*(August), 94–101. https://doi.org/10.1016/j.neuroimage.2019.03.080

Schwarzer, R, and M Jerusalem. 1995. "Generalized Self-Efficacy Scale." *Measures in Health Psychology: A User's Portfolio. ...*.

Sheppard, Sheri, Phil Weilerstein, Tina Seelig, Tom Byers, Tina Seelig, Sheri Sheppard, Phil Weilerstein, Tina Seelig, and Tom Byers. 2013. "Entrepreneurship: Its Role in Engineering Education." *The Bridge*, no. Summer.

Spreng, R. Nathan, Margaret C. McKinnon, Raymond A. Mar, and Brian Levine. 2009. "The Toronto Empathy Questionnaire: Scale Development and Initial Validation of a Factor-Analytic Solution to Multiple Empathy Measures." *Journal of Personality Assessment* 91 (1). https://doi.org/10.1080/00223890802484381.

Stewart, W. H., & Roth, P. L. (2001). Risk propensity differences between entrepreneurs and managers: a meta-analytic review. *Journal of Applied Psychology*. https://doi.org/10.1037/0021-9010.86.1.145

Unger, Jens M., Andreas Rauch, Michael Frese, and Nina Rosenbusch. 2011. "Human Capital and Entrepreneurial Success: A Meta-Analytical Review." *Journal of Business Venturing* 26 (3). https://doi.org/10.1016/j.jbusvent.2009.09.004.

Van den Berg, I., Franken, I.H.A., Muris, P. (2010). "A new scale for measuring reward responsiveness." *Frontiers in Psychology 1*(DEC). https://doi.org/10.3389/fpsyg.2010.00239.

Whiteman, M. C., A. Bedford, E. Grant, F. G.R. Fowkes, and I. J. Deary. 2001. "The Five-Factor Model (NEO-FFI) and The Personality Deviance Scales-Revised (PDS-R): Going around in Interpersonal Circles." *Personality and Individual Differences* 31 (2). https://doi.org/10.1016/S0191-8869(00)00134-3.

Xie, Hua, Iliana I. Karipidis, Amber Howell, Meredith Schreier, Kristen E. Sheau, Mai K. Manchanda, Rafi Ayub, et al. 2020. "Finding the Neural Correlates of Collaboration Using a Three-Person FMRI Hyperscanning Paradigm." *Proceedings of the National Academy of Sciences of the United States of America* 117 (37). https://doi.org/10.1073/pnas.1917407117.

Zaidel, D. W. (2014). Creativity, brain, and art: biological and neurological considerations. *Frontiers in Human Neuroscience, 8*, 389. https://doi.org/10.3389/fnhum.2014.00389