113 學年度第 2 學期設計學研究所資格考試題

考科：**設計文獻評論**

1.本考科分為二大部份：第一部份試題及第二部份試題，請從各部份試題中，任選二題作答，共 4 題，每題 25 分。
2.分別說明各篇論文的**主旨、綱要**，並**加以評論**。

【第一部份試題】 設計文獻評論（一）：選擇兩題回答，每題 25 分，總分 50 分。
(請在答案卷上作答，並標示題號)

---

1. Stone, K. (2023). **Reparative game creation: Designing for and with psychosocial disability.** *Design Issues*, 39(1), 14-26.
2. Christensen, B. A. (2023). **Revisiting the Experience Machine: A Philosophical Note on the Limits of Experience Design.** *Design Issues*, 39(2), 14-26.
3. Sieghart, S. (2023). **The Influence of Fonts on the Reading Performance in Easy-to-Read Texts: A Legibility Study with 145 Participants.** *Design Issues*, 39(3), 30-44.
4. Vink, J. (2023). **Embodied, Everyday Systemic Design-A Pragmatist Perspective.** *Design Issues*, 39(4), 35-48.

【第二部份試題】 設計文獻評論（二）：選擇兩題回答，每題 25 分，總分 50 分。
(請在答案卷上作答，並標示題號)

---

1. Cooper, L. (2023). **Constructing accounts of decision-making in sustainable design: A discursive psychology analysis.** *Design Studies*, *84*, 101158.
2. Das, M., Ostrowski, A. K., Ben-David, S., Roeder, G. J., Kimura, K., D'Ignazio, C., ... & Verma, A. (2023). **Auditing design justice: The impact of social movements on design pedagogy at a technology institution.** *Design Studies*, *86*, 101183.
3. Kwon, E., Rao, V., & Goucher-Lambert, K. (2023). **Understanding inspiration: Insights into how designers discover inspirational stimuli using an AI-enabled platform.** *Design Studies*, *88*, 101202.
4. Cash, P., Daalhuizen, J., & Hekkert, P. (2023). **Evaluating the efficacy and effectiveness of design methods: A systematic review and assessment framework.** *Design Studies*, *88*, 101204.

Reparative Game Creation: Designing For and With Psychosocial Disability

Kara Stone

Introduction

The relationship between psychosocial disability and videogames commonly appears in two places: games' positive or negative effect on individuals' behavior, and games' representation of mental illness.[1] In pop culture, videogames have been associated with violence, antisocial behavior, and addiction in their players. In the field of game studies, games are positioned as vessels for positive change, socialization, and self-soothing for players.[2] Rather than focus on the affect of players or the representation of illness in various games, this article focuses on videogame design and its possible addition to a paradigm shift that aligns media with practices of amelioration, healing, and rest. I propose a design framework called reparative game creation, a process of creating interactive media focused on healing, emotional acceptance, and accessibility for the psychosocially disabled. It is informed by disability studies, affect theory, anti-capitalist thought, and artist scholarship on research creation or critical practice. The goal of this work is to recognize the current massification of psychosocial disability, mental illness, and debility and to orient the design and use of interactive media toward healing and care, rather than toward a slow debilitation as is the current convention. Though much of game design and game studies focuses on the end product or the player experience, this article instead focuses on the process of game design, and as such it does not analyze specific games but proposes new ways of creating games informed by psychosocial disability. It is intended to be useful for game creators, artists, and designers interested in creating for and with disabilities and those wanting to rethink dominant paradigms of creation.

Gaming culture and dominant game design are in need of repair. The design of most commercial games and many independent games use a compelling and compulsive cycle of work and reward and the feeling of productivity to entice the player to play more. Most games follow a cycle of frustration and satisfaction,

1 The accessibility of videogame hardware such as controllers is also a site of discourse between games and disabilities, though it is related to physical more than psychosocial disabilities.

2 Brendan Keogh, "Games Evangelists and Naysayers," *The Conversation*, March 20, 2014, http://theconversation.com/games-evangelists-and-naysayers-25006.

Design*Issues:* Volume 39, Number 1 Winter 2023 https://doi.org/10.1162/desi_a_00703

increasing in difficulty as the game progresses. Katherine Isbister states that playing a game activates "reward-related mesolimbic neural circuits—parts of the brain associated with motivation and reward."[3] Brie Code, designer of app game *#SelfCare*, also describes gaming in terms of biochemical psychology: "Game design theory is based on an adrenaline/dopamine response to stress. Game designers aim to stress the player, and then give the player opportunities to win a challenge."[4] This system of play makes players feel as if they have accomplished something even if in reality all they have done is to sit in front of a screen for sixty hours and disassociate from their body and feelings. In itself, this is not necessarily "negative," but the intense desire to feel productive is an internalization of current neoliberal capitalist culture, where one's worth or morals are determined by productivity. Many games are designed to tap into this cultural desire, creating a cycle of work and reward are closely planned out to keep the player playing for as long as possible.[5]

Cultures surrounding gaming are often exclusionary particularly to women and people of color.[6] In North America, white cisgendered men are the dominant demographic of game developers. Employees at more and more studios, including mainstream and indie, have been coming forward with experiences of toxic workplaces. The games industry notoriously relies on crunch, a mode of work that demands overly long hours and weekend work, prompted often by appealing to the employees' passion for videogames or through guilt. These crushing workplaces make game development inaccessible to many, including disabled people who are unable and/or unwilling to work in ways that actively damage their lives. Because of dominant design paradigms and exclusionary workplaces and cultures, I see a turn toward reparative game creation as a provocative and beneficial framework for exploring alternate possibilities in the videogame realm.

The following section lists the principles of reparative game creation—direct and simplified practical guides to creating games oriented toward healing and repair. These principles include how a reparative game should operate and consider the process of making reparatively as equally (if not more) important. I start with them to give a general sense of the key elements of this framework. The rest of the article lays out the foundational theories of reparative game creation, the reparative position, and healing as it relates to disability. Reparative game creation is part practical design, part analysis, and part utopian dreaming of an alternate media landscape, one that helps sustain life rather than drain it.

3 Katherine Isbister, *How Games Move Us: Emotion by Design* (Cambridge, MA: MIT Press, 2017), 3.

4 Brie Code, "About: How Do You Feel?," *TRU LUV*, https://web.archive.org/web/20190302171843/http://truluv.ai/about-tru-luv/ (accessed August 8, 2022).

5 See Natasha Dow Schüll, *Addiction by Design: Machine Gambling in Las Vegas* (New Jersey: Princeton University Press, 2012) for detail on how casino games are designed to keep the player playing.

6 See TreaAndrea M. Russworm and Jennifer Malkowski, *Gaming Representation: Race, Gender, and Sexuality in Video Games* (Indiana: Indiana University Press, 2017)

The Principles of Reparative Game Creation

These principles are not rules but guiding intentions for designing games with and about psychosocial disability that I have found useful in my own design process.

Repair

- Games are no more healing than any other art form.
- Games will never cure anyone. Cure is not the goal. Make work that changes the media landscape, adding to a culture oriented toward making life more livable for those with psychosocial disabilities.
- Do not only represent psychosocial disability but design for it.
- No score, no winning, no losing, no good paths, no bad paths, no one right way to play, no punishment, no reward.
- Be emotionally challenging, not mechanically challenging.
- Describing some feelings as "good" and some feelings as "bad" are judgments. They might ease classification but are not inherent or true.
- Feelings are a political resource.
- Everything is an experiment and an exploration.

Care

- Recognize not only the labor of the time spent working but also the preparation and recovery.
- Do not expect to work like a machine (consistent, always productive, not emotionally affected by mood, physical experience, environment, food, weather, past experiences, brain fog, pain, injustice, neurochemicals, microbiome, our partner's mood, noise, etc.). We are not machines. Capitalism has set up unrealistic and unattainable work expectations. These expectations actively damage and drain us. Resist wanting and trying to live up to them.
- Know our cycles of work. Know that some are in our control, and some are out of our control.
- Rest without guilt.
- Art does not need suffering to be made.
- Affect cannot be quantified. Foster emotional self-reflection rather than aiming to funnel the players into a specific emotion.
- Do not overwork, and do not ask players to overplay.

- Be cautious—and grateful—of how much time is asked of players.
- Remind ourselves and players that there is life beyond the screen.
- Make space for self-reflection in the game: emotional choices, diary entries, mood trackers, pauses, time-outs, direct questions, and surveys.

Share

- Be vulnerable. Be radically open during the creation stage. Edit after it is already created if necessary.
- Ask others for help. There is beauty in interdependence.
- The process of making is about coming to understanding, not demonstrating understanding.
- A good game is not determined by a game one wants to play all the time. Meaningful experiences last beyond the playtime.
- Perfection does not exist. Nothing is ever perfect, only an expression of that moment in time's preferences, desires, expectations, conditions, and experiences.
- Feelings are powerful and should not be avoided, but consider how some may affect potential players. Will it make the game inaccessible to certain groups?
- Consider the audience as someone in the disabled communities, experiencing psychosocial disability and many forms of debilitation, not an audience we must explain our experiences to.

Psychosocial Disability

This article deviates from previous research on videogames and mental illness that align with medical institutions and medicalized models of health, instead coming from lived experience, process-based art, and critical disability studies. There are other terms to describe the same or similar communities, such as consumer/survivor/ex-patient (c/s/x), mental illness, neurodivergence, and mental disabilities. None of these are necessarily wrong or right; I choose the term "psychosocial disability" because it positions one's debilitation in both the inner psychological and the outer social contexts and puts it in relation to the broader category of disability; contained within that is the field of disability studies, activism, legal status, and cultural marginalization. Although anti-psychiatry, Mad Pride, and c/s/x each have specific histories and meanings, often viewed as distinct from disability, generative concepts are produced when they are considered together, such as critiques of normality,

reconfiguring health and illness, and desiring experiences often viewed and undesirable.[7] As such, much of the theory that is foundational to this article is affect theory and disability studies, rather than psychoanalysis, psychiatry, or psychology. Disability studies offers a way of understanding the broader context of accessibility, debility (a form of understanding how populations are worn out, decapacitated, and disenfranchised without attaching the binary of disabled/abled[8]), and culture and how they relate to mental health.

Disability scholar Lennard Davis argued that psychoanalysis is a "eugenics of the mind—creating the concepts of normal sexuality, normal function, and then contrasting them with the perverse, abnormal, pathological, and even criminal."[9] Psychoanalysis and psychiatry are not the only therapies and care the field of clinical psychology offers, however, they are the most common in pop-cultural understandings of mental health and most integrated into scholarship in the humanities. Still, most clinical psychologies operate on the categorization basis of "excessive" affect or unwanted behaviors. In following an affect theory– and disability studies–informed model of health and wellness, videogames are less a tool for systematically decreasing psychosocial disability, as if the goal is to overcome psychosocial disability and shed all negative feelings, but instead positions game design as a potential avenue for interactive practices that build toward creating a more livable life by providing moments of self-reflexivity, connection, bodily sensation, and acceptance.

The Reparative Position

The term "reparative" is inspired by Eve Kosofsky Sedgwick's theorizing of paranoid and reparative positions. In previous work, I laid out the foundations for reparative game creation as I thought through my process for making the videogame *Ritual of the Moon* (2019), and how these two positions, reparative and paranoid, might not be so oppositional.[10] Briefly, *Ritual of the Moon* is a twenty-eight-day long videogame that the player plays for three minutes a day, following a sparse narrative and engaging in reflective activities. It was the beginning of my thinking about combining Sedgwick's theories with game design. Sedgwick, a preeminent queer theorist, details paranoid reading, the most common form of critique in academia, a mode of searching for and then revealing hidden violence, oppression, or wrong ideas.[11] The vigilance, the demand for perfection, the constant search for flaws or limitations, all prevent one from feeling other experiences or transformations—especially ones that are considered naive, such as surprise, openness, and celebration.[12] She argues that this is the most dominant form of

7 For a historical account of the relationship between mental illness and disability, see Bradley Lewis, "A Mad Fight: Psychiatry and Disability Activism," in *The Disability Studies Reader*, 4th ed., Ed. Lennard J. Davis (New York: Routledge, 2013), 116.

8 Jasbir K. Puar, *The Right to Maim: Debility, Capacity, Disability* (Durham, NC: Duke University Press, 2017), xv.

9 Lennard Davis, "Normality, Power, and Culture," in *The Disability Reader*, 4th ed., Ed. Lennard J. Davis. (New York: Routledge, 2013), 8; https://doi.org/10.4324/9780203077887.

10 Kara Stone, "Time and Reparative Game Design: Queerness, Disability, and Affect," *Game Studies: The International Journal of Computer Game Research* 18, no. 3 (2018); Eve Kosofsky Sedgwick, *Touching Feeling: Affect, Pedagogy, Performativity* (Durham, NC: Duke University Press, 2003).

11 Ibid., 222.

12 Ibid.

creation in academia, and she then moves on to theorize "reparative reading." Reparative reading would be a mode of analysis and creation that is not solely invested in pointing out more insidious forms of oppression but looks toward vulnerability, survival, affect, risks, and sustenance. Sedgwick uses words like "multiplicity, surprise, rich divergence, consolation, creativity, and love" to describe it.[13] It does not have all the answers right away. It is vulnerable. It doesn't default to searching for mistakes or comparisons. Opening oneself for healing and repair means opening oneself to risk, including so-called negative feelings. Being reparative can mean oscillating between good and bad affects and experiences, deeming them as not necessarily clearly divisible from each other, and approaching bad feelings as not necessarily an experience to avoid. Reparative game creation is committed to the speculative imagining of what research creation may look like when it is dedicated to survival, love, and risk.

For the purposes of this article, the reparative position or practice, reparative *reading* is being refigured to reparative *game creation*. In the context of videogames, reparative reading might at first be thought of as akin to reparative *playing*, which is certainly possible, but Sedgwick uses reading not only as the act of taking in information but also as the response to it—how academics write and talk in communication and discussion with each other. Reparative game creation focuses on the process of creation, specifically from a psychosocially disabled point of view, and ways that creators can incorporate disabled and reparative practices into process and design. Instead of a mood of suspicion, interrogation, guardedness, and desire to expose others, the reparative response takes risks in proposing, modeling, or exploring alternative modes of being and understanding—strategies very common in research creation and critical practice.[14] Reparative game creation is in line with research creation that focuses on the creative process as research.

Healing

Healing, as I use it in this article, has a specific connotation in relation to disability and social justice. Disability justice activist Leah Lakshmi Piepzna-Samarasinha describes healing justice, a movement and term created by queer and trans Black, Indigenous, and people of color, in which Piepzna-Samarasinha is a major figure. She asks, "What do you think 'healing' is? Do you think that it means becoming as close to able-bodied as possible? Do you think it is always sad or terrible to be sick or disabled? Do you think everybody wants to be able-bodied and neurotypical, and would choose it if they could?"[15] Questioning expectations of the

13 Heather Love, "Truth and Consequences: On Paranoid Reading and Reparative Reading," *Criticism* 52, no. 2 (2010): 235–41.

14 I use the phrase "research creation," which is common in Canada. It has a similar meaning to "practice-based," "practice-led," "creative practice," "practice as research," and "arts-based research."

15 Leah Lakshmi Piepzna-Samarasinha, *Care Work: Dreaming Disability Justice* (Vancouver: Arsenal Pulp Press, 2018), 94.

desirability of able-bodiedness and able-mindedness is integral for artists and academics interested in social activist games focused on disability to combat potential ableism. Piepzna-Samarasinha knows it is hard to get good healing, healing that is affordable, a practitioner that listens to the person's needs, and one who disavows their own expectations of what the person wants. She communicates that "most sick and disabled people I know approach healing wanting specific things—less pain, less anxiety, more flexibility—but not usually to become able-bodied."[16] Healing is done to make life more livable and to have the opportunity to thrive, not to overcome or eradicate disability.

Healing can be easily dismissed—seen as not important to political futures, reserved for people with money, or an individual responsibility that is not worthy of being addressed in collective organizing. Because it is often dismissed, the burnout rate in community organizations, academia, and the game industry is incredibly high. Cara Page of the Kindred Southern Healing Justice Collective proposed that "Our movements themselves need to be healing, or there is no point to them."[17] Care work and anti-ableist practices are not to be side-lined but are central to political movements—not individual practices that we do to get ready to do the "real work" again. Piepzna-Samarasinha puts it this way: "It doesn't have to be either healing or organizing: it's both . . . Healing justice is not a spa vacation where we recover from organizing and then throw ourselves back into the grind,"[18] as much of the rhetoric around self-care insinuates.

For Piepzna-Samarasinha, like Ann Cvetkovich and other queer affect theorists, feelings are essential to political organizing. Many movements involve grieving and come out of a need for grief to be publicly recognized: "grief and trauma are not a distraction from the struggle."[19] Healing and reparative work does not come from pushing down negative feelings and bringing up positive feelings, but from giving time and energy to both. Reparative game creation is not trying to make players feel better, to totally absolve them of anxiety, depression, mania, diagnosis, pain, and more, but to reorient media practices and play toward supporting life, sustaining energy, rejuvenating each other, and creating a life one considers worth living.

The Need for Reparative Media

While disability justice advocates are not necessarily against all uses of the medical model of mental illness, the individual nature of common psychological care limits therapy's cultural effect. Individualized therapy is not enough to tackle the world's massive

16 Ibid., 103.

17 Quoted in ibid., 100.

18 Ibid., 107.

19 Ibid., 108.

amount of psychosocial disability and debilitation. Kazdin and Blase highlight the inefficiency of the individualized delivery system of talk therapy.[20] Although there are many techniques of therapy, such as dialectical behavioral therapy, psychoanalytic therapy, cognitive behavioral therapy, eye movement desensitization and reprocessing therapy, which are among the most popular, they almost all share the same model of delivery: administered to one person at a time or up to a small family unit. Kazdin and Blasé point out that there are not enough psychologists to address the massive population in need of care; they estimate that 25 percent of the population has at least one psychiatric disorder in a given year. In the United States, that would be approximately 75 million people.[21] Even if the current workforce of psychologists were to double, it would have "little discernible impact given the number of individuals requiring services," especially those who live in remote areas or do not have adequate access to health care.[22] Kazdin and Blase advocate for an opening up of the delivery model from intense one-to-one sessions to less intensive but more massive scale. A treatment can look very different than what we would think of as therapy. Technology is a significant portion of this: the internet, phone calls, text messages, and apps are all cited as having the potential to deliver therapy at larger and more economical scales. Although there are many reasons to be skeptical of technological programs designed for therapy, Kazdin and Blase offer optimism for the possibilities—and necessity—of using technology to provide widespread therapeutic programs. Technology may never be a substitute for a real person, but "technology is not competing with a real person. It is directed toward the goal of reducing the burden of mental illness, and in this regard it can make a contribution to a portfolio of delivery models that is without peer."[23] They even write that "the creativity of video games may increasingly be applied to treatment of preventive regimens and be made readily available."[24] The work of Kazdin and Blase showcase how the processes of care for psychosocial disability need to be broadened and that many forms of media, communication, and organizing can be oriented toward healing.

Measuring a Reparative Experience

How can a reparative experience be measured, if at all? The majority of the reparative game creation principles are focused on the creative process, not on the player experience. This is a radically different turn from the majority of popular game design tactics, which are player-centric. Such tactics determine the success of a game's design through player feedback: if the players' values changed, if they understood the messaging, if they had an engaging play experience,

20 Alan E. Kazdin and Stacey L. Blasé, "Rebooting Psychotherapy Research and Practice to Reduce the Burden of Mental Illness," *Perspectives on Psychological Science* 6, no. 1 (January 2011): 21–37.

21 Ibid., 23.

22 Ibid., 26.

23 Ibid., 32. Kazdin and Blase define the burden of mental illness as the "personal, social, and monetary costs associated with impairment," not the popular conception of burden as something someone else has to unwillingly care for" (21).

24 Ibid., 32.

and if it enabled some "change" to take place in the player's resulting actions. This is not to foreclose the idea of reparative play—which I believe is possible—but to problematize player-centricity and the emphasis on measurable outcomes of art.

If one were to measure a reparative experience to quantify it, study it, or use the information in next designs, what exactly would be measured? To measure emotion or affect, psychology and psychiatry use patient self-reporting or individual assessments, but often this tracks only the short term. In the field of human–computer interaction, affect has been quantified through biometrics such as heart rate and facial micro-expressions, which do not universally map onto experiences of emotion. Even if I could construct a long-term clinical psychology-style quantitative study on the effects of playing a reparative game, what markers would I look for? A person feeling "better" is not the goal, nor is a cure or eradication of bad feelings. For example, the measurement could not simply look for lessened anxiety or heightened joy as evidence of efficacy. Instead, the goal for reparative game creation is to participate in a movement of media that function as tools for care, connection, reflection, and radical acceptance of emotion. Again, I do not wish to disavow or renounce player experience as a metric or evaluative tool, and it may well become a part of reparative game creation but finding a mode of measurement while staying in line with disability studies, affect theory, and demedicalization of affect is not the goal of this article.

Others are also hesitant about the demand for quantification of the social effects of game design. Paolo Pedercini has criticized the measurement model of games-for-change, saying that "the presumption is that social change can be measured in the same way you can measure the calories burned by playing an exercise game."[25] Demanding everything be quantifiable is a tool of capitalism: "If you can measure something, you can rationalize it, you can optimize it, you can sell it." It is not that any focus on outcome is bad, but that by directing all the focus on "measurable goals we narrow our action. We favor individual change versus systemic and long-term change. We target burning calories without addressing food politics and food justice. We try to impose prepackaged behavior protocols rather than facilitating critical thought."[26] The common types of change that games are designed for are small-scale and individual-focused, rather than the grander political overhaul Pedercini supports. Instead of making sure a game produces quantifiable outcomes, Pedercini argues for the liberatory potential of game design and how it can teach criticality and demystify the medium. For these reasons, reparative game creation focuses on the process of

25 Paolo Pedercini, "Making Games in a F****d Up World," presented at Games for Change Festival (2014).

26 Ibid.

creating games rather than the games' effect on individuals. Furthermore, reparative game creation aims to strike a balance between self-change and world change. Making a game reparatively can be one transformational tool for self-reflection and healing journeys. Players may relate to the game and find it validating or transformational experiences through play. Still, a single reparative game—or a few, or ten—will not have the massive effect needed to transform the structural and systemic forces of oppression that disable and debilitate people. Reparative games call for transforming the media landscape, just one step of what needs to be a mass movement toward media that is healing, inclusive, and sustainable.

A Summary of the Reparative Game Creation Principles

I have synthesized the list of principles that began this article into five categories that summarize their themes and goals: games are not cures, design for psychosocial disability, a reparative process, no addictive cycles, and anticapitalism and speculative utopia.

Games Cannot Be Cures

Cure is a complicated issue in disability studies. For the social model of disability, there is a difference between impairment and disability; the former is the bodymind-based limitation and the latter is the effects of society. Disability studies scholar Alison Kafer summarizes it as such: "People with impairments are disabled by their environments; or, to put it differently, impairments aren't disabling, social and architectural barriers are."[27] Focusing on cure may address impairment but not the socially disabling culture. Cure might mean eradication of a certain way of being in the world, to make "normal," and for disability to be undesirable rather than a unique, variant part of life. Disability is so often written out of the future, as something utopia will "fix" through eradication. This idea showcases how deeply set cultural imaginings of disability as undesirable are, so the position of the social model can feel empowering. On the other hand, there has recently been more conversation around cure from scholars and activists with chronic illness and pain. While some disabled people distance themselves from being considered unhealthy, sick, or suffering, others are embracing those terms while taking seriously the social and political realities of impairment.[28] The suffering the unhealthy or sick disabled experience is real; some could be eliminated by social justice, but not all of it. To eradicate anxiety completely, to get rid of depression entirely, is undesirable and unrealistic. Of course, these feelings can do harm. Instead of a cure through eradication of any negative feeling, we need to find ways to live alongside them, mitigate their harm, and

27 Alison Kafer, *Feminist, Queer, Crip* (Bloomington: Indiana University Press, 2013), 7.

28 See Susan Wendell, "Unhealthy Disabled: Treating Chronic Illnesses as Disabilities," *Hypatia* 16, no. 4 (2001): 17–33; and Johanna Hedva "Sick Woman Theory," https://johannahedva.com/SickWomanTheory_Hedva_2020.pdf.

find the political and reparative potential in them. Reparative game creation, informed by affect theory and many forms of psychotherapy, transforms these affects such as anxiety and depression into something livable at the very least, and then beyond that to being informative, community building, and politically provocative. In imagining a utopian future, let us not imagine one without psychosocial disability; let's instead imagine and work toward a future without compulsory able-bodiedness and able-mindedness, where psychosocially disabled people have control over their time, where they are seen as valuable rather than detrimental to society and productivity, and where their feelings are respected and normative expectations of emotions, care, relationships, and work are expanded and flexible.

Design for Psychosocial Disability, Not Just Represent It

The focus of reparative game creation is to consider how healing and disabled practices can be incorporated into the process of making, with an intention toward psychosocially disabled creators. A common practice in making games about marginalization and oppression is to incorporate marginalized characters at the level of character representation. Reparative game creation is about the structure of a game's design, the feelings it fosters, and the process of creating the game, not about representation of disabled and mentally ill characters. In fact, a reparative game may have no characters at all. Another common design mode for games designed for social change is to depict oppression or lived experience, such as disability, but direct the game toward audiences who do not experience that marginalization as a way of fostering empathy or social change. Reparative game creation instead is geared toward psychosocially disabled and debilitated creators designing for psychosocially disabled and debilitated audiences. Although creators might want to express our experiences with psychosocial disability or debility, it is not necessarily to inform an unfamiliar audience but instead allows us to come together, share experiences, and let each other know *I feel this way, too*. The audience is statistically quite likely to have experienced mental illness and disability.[29] Therefore, instead of expressing ourselves through videogames intended for nondisabled people, reparative game creation is self-expression toward community. This brings up questions related to the games design, such as is the game meant to bring about anxiety? How will that affect people with high levels of anxiety? Does it use addictive patterns? Whom does that benefit? This is not to say all games should be soothing or calming; I believe all feelings can be reparative, with reflection and intention. As creators we can try

29 In 2018, the Centers for Disease Control and Prevention estimates that one in five people in the United States have a mental or physical disability. Centers for Disease Control and Prevention, "CDC: 1 in 4 US Adults Live with a Disability," August 16, 2018, https://www.cdc.gov/media/releases/2018/p0816-disability.html.

to be conscious of what feelings the piece is bringing up, who will be feeling those feelings, what they are meant to do in the world through the imagining, designing, creating, prototyping, and playtesting stages.

The Process of Creation Itself Can Be Reparative, Not Just the Product

The process of making can be a reparative experience for the creator. This isn't specific to games; there is a long history of process-based art and art therapy. In game design, this takes away the player as the ultimate focus, which is the dominant paradigm, while not necessarily writing players out entirely. As such, it is not as player-centric as many game design heuristics and models. Games do not have to be released or widely played to be of value. Yet the process of making can feel destructive. It is the primary way making art is positioned: as our "blood, sweat, and tears," draining energy, needing suffering and self-criticism, seeming pointless, time-consuming, and even physically harmful like crunch. Reparative game creation foretells that the process can also be healing in self-reflection, feeling masterful, experimenting, articulating and communicating ideas and feelings to others, and in different forms of self-expression. The process of creation can enable understanding and refiguring the value of productivity. We need to find our own strategies that help guide us to one side rather than the other. Suffering is not a necessity of art. Psychosocial disability is not a necessity of art. Many of the reparative game creation principles are oriented toward the making process, not just the end result, in hopes of marrying the design process to the outcome.

No Addictive Cycles of Frustration and Reward

Most videogames are designed to perpetuate a cycle of frustration and momentary gratification, work and reward, and obsessive behavior to bring out addictive behavior in players. To be addicted to a videogame is not likely to be the designer's actual goal; however, game studios benefit when their players continue playing—and playing and playing. Players keep paying subscriptions, buying downloadable content, writing walkthroughs on how to 100% a game, (unpaid labor), stream (unpaid advertising), buy the sequel, and more. For most videogame designers and studios, the only reason to stop playing a game is when that game is completed, regardless of it being 1 hour or 200 hours. Within reparative game creation, growing bored is a fine affective outcome. To stop playing because the player feels satisfied is great. Rather than keep them playing, the goal is to create self-reflection on their emotional state. Cycles of frustration and reward, and increasing difficulty, are so deeply

embedded in game studies' definitions of games and the cultural consciousness around games that to design outside of it registers as not-a-game. Reparative game creation imagines what games could be if they did not have to include winning, losing, increased difficulty, score, good paths, bad paths, punishment, or reward.

Anticapitalism and Speculative Utopia

Capitalism is a major factor in mass debilitation. It determines how we make art, how we believe we should work, the materials we use, what platforms we distribute on, and which companies profit from our work. One cannot make a reparative game that requires a new console of which its production destroys the environment, uses stolen labor, or upholds abusive work expectations. A reparative game is not designed to be addictive to extract money from the player. It would not be possible for a videogame corporation like Ubisoft to make reparative games because of their carbon footprint, their capitalist business structure, and the burnout rate of their nonunionized employees. I intend reparative game creation to be useful for independent artists and small studios to inspire a paradigm shift in game design. It is utopian and speculative; not becoming stagnant in the belief that what is now will always be. Imagining and working toward utopia is not about being perfectionistic or unrealistic, but a refusal of doomed futures, settling for the status quo, and the stagnation or detriment of healing.

Conclusion

This article offers reparative game creation as a form of research creation where the process of making art is a source of coming to critical information, and the art product is a form of making research. It was not my goal to cement the ideas and principles in perpetuity; rather, the principles are intended to be ever-changing and nonprescriptive as we become mindful of what strategies work best when and for what. It resists measurement—at least measurements that reify able-minded-centricity and props up demands to erase so-called negative feelings. This article outlines what reparative game creation is and its foundational principles to give a practical framework for those interested in designing more restorative, reparative, and different games. It aims to reconfigure Sedgwick's reparative position into reparative game creation, a mode of making games that aims to promote repair, healing, and spaces for feelings not common in videogames. "Reparative" does not mean curing or becoming able-bodied or able-minded, but to support and encourage a vibrant life rather than sap and deplete it. Creating and sharing videogames that support psychosocial disability is one way to move toward a more healing and caring media landscape.

Revisiting the Experience Machine: A Philosophical Note on the Limits of Experience Design

Bo A. Christensen

Introduction

Robert Nozick's example of the experience machine is probably one of the most discussed thought experiments in the history of twentieth-century philosophy. It is a small section in a chapter called "Moral Constraint and the State" in Nozick's *Anarchy, State, and Utopia*.[1] It also is part of the genus of interesting philosophical thought experiments, exemplified by Hilary Putnam's brains in a vat and René Descartes's evil demon.[2] Here is Nozick's description:

> Suppose there was an experience machine that would give you any experience you desired. Superduper neuropsychologists could stimulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book. All the time you would be floating in a tank, with electrodes attached to your brain. Should you plug into this machine for life, preprogramming your life's experiences? If you are worried about missing out on desirable experiences, we can suppose that business enterprises have researched thoroughly the lives of many others. You can pick and choose from their large library or smorgasbord of such experiences, selecting your life's experiences for, say, the next two years. After two years have passed, you would have ten minutes or ten hours out of the tank, to select the experiences of your next two years. Of course, while in the tank you won't know that you're there; you'll think it's actually happening. Others can also plug in to have the experiences they want, so there is no need to stay unplugged to serve them. (Ignore problems such as who will service the machines if everybody plugs in). Would you plug in?[3]

By replacing the neuropsychologists with designers and the brain with experience in the quote, would the experience machine not become the perfect design model, fulfilling people's experiences, needs, and desires? The experience machine thought experiment is

1 Robert Nozick, *Anarchy, State, and Utopia* (London: Blackwell Publishers, 1974).

2 See Hilary Putnam, *Reason, Truth, and History* (Cambridge: Cambridge University Press, 1981); and John Cottingham, René Descartes: Meditations on First Philosophy with Selections from the Objections and Replies (Cambridge: Cambridge University Press, 1996).

3 Nozick, *Anarchy*, 42–43.

https://doi.org/10.1162/desi_a_00714

important because it seems to bracket specific theoretical and practical discussions about experiences and instead asks a fundamental question: If designers could make—and people had the opportunity to choose—the perfect experience, fulfilling all the needs of their lives, would they care to do it? Nozick's answer is presumably no—the fulfillment of experiences and needs is but one aspect of people's commitment and caring about their lives within a world that matters. A broader sense of reality is not capturable by a sole focus on experiences as the fulfillment of pleasure and needs.

In this article, I present a philosophical thought experiment—or a kind of speculative design fiction[4]—using Robert Nozick's idea of an experience machine to discuss the relation between experience and design. The discussion is predominantly philosophical and not empirical. Hence, I try to lay out the argumentative ramifications of Nozick's thought experiment for a general idea of experience design. I use Marc Hassenzahl's thoughts on experience design as background, or case.[5] Although I mostly agree with him, I address a level of existential or lived commitment that serves as a significant background of how the world matters to us, emphasizing how we care about it.[6]

This approach may strike a note of recognition in readers who are familiar with the experience-centered design of Peter Wright and John McCarthy, as well as the human-centered design approach associated with Richard Buchanan and Ian Hargraves.[7] However, differences also may be noted between these pragmatic-oriented and human-centered approaches and the notion of commitment I pursue here. To contextualize Hassenzahl's approach, I first present some trajectories in the development of experience design based on Jodi Forlizzi and Katja Batterbee, as well as Wright and McCarthy.[8] I then explain Hassenzahl's version of experience design as the means to bring these trajectories together. I then engage with Nozick's experience machine, both by discussing previous interpretations and by laying out a new interpretation. In light of this new interpretation, I discuss Hassenzahl's version of experience design and the possible consequences when experience design includes the related human-centered design.

The conclusion and results of this article are twofold. First, Nozick's experience machine directs us toward a central aspect of experience related to existential and lived commitments, which are not capturable by experience design as presented by Hassenzahl and others. A commitment characterized by a decentered human perspective is as important as the human-centered one. Second, the article reveals that using a philosophical design fiction—like Nozick's experience machine—offers a methodological tool for questioning or deconstructing certain ingrained conceptions related to experience design.[9]

4 Anthony Dunne and Fiona Raby, *Speculative Everything: Design, Fiction, and Social Dreaming* (Cambridge MA: MIT Press, 2013).

5 Marc Hassenzahl, *Experience Design* (California: Morgan & Claypool, 2010).

6 John Haugeland, *Having Thought* (Harvard: Harvard University Press, 1998); and Marc Lance, "Life is Not a Box-Score: Lived Normativity, Abstract Evaluation, and the Is/Ought Distinction," in *Meaning Without Representation: Essays on Truth, Expression, Normativity, and Naturalism*, ed. Steven Gross et al. (Oxford: Oxford University Press, 2015), 279–306.

7 See, respectively, Peter Wright and John McCarthy, *Experience-Centered Design* (California: Morgan & Claypool, 2010); Richard Buchanan, "Human Dignity and Human Rights: Thoughts on the Principles of Human-Centered Design," *Design Issues* 17, no. 3 (Summer 2001): 35–39; and Ian Hargraves, "Care and Capacities of Human-Centered Design," *Design Issues* 34, no. 3 (Summer 2018): 76–88.

8 Jodi Forlizzi and Katja Battarbee, "Understanding Experience in Interactive Systems," in *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (New York: ACM Press, 2004), 261–8; and Wright and McCarthy, *Experience-Centered Design*.

9 Rupert Read and Bo Allesøe Christensen, "Why 'Swampman' Would Not Even Get As Far As Thinking It Was Davidson: On the Spatio-Temporal Basis of Davidson's Conjuring Trick," *Philosophical Investigations* 42, no. 4 (2019): 350–66.

Versions of Experience Design

As Forlizzi and Battarbee emphasized, the turning toward experiences in interaction design was a result of moving past a myopic focus on functionality in design research.[10] Turning to experience discloses different relevant aspects of experience—physical, sensual, cognitive, emotional, and aesthetic—that result from the interaction between people and products. Forlizzi and Battarbee divided experience design into three categories: product-, user-, and interaction-centered.[11] The product-centered approach, which is similar to the functional approach, focuses on the criteria for establishing a good, usable design. In Nozick's example, this functionality is simply assumed: The machine is, as a premise, taken to work as it should. The user-centered aspect is about understanding people and what they would find relevant in interacting with the product, and Forlizzi and Battarbee mentioned Hassenzahl as an example of a design researcher who espouses this perspective; with his focus on broadening traditional goal- and task-based thinking to include fun and action-oriented modes of behavior. In Nozick's experience machine example, this approach is tantamount to providing people with the opportunity to select a diversity of preferable experiences within the machine. The interaction-centered aspect focuses on the role of products in bridging designer and user. Wright and McCarthy's pragmatic (and phenomenological) approach to the qualitative aspects of experience, inspired by John Dewey, is mentioned here. Regarding Nozick's example, this third aspect will provide a genuine first-person experience of being embodied within the machine, including a pragmatic sense of trial and error and thereby an experience of learning and gaining knowledge.

Florizzi and Battarbee expanded on this third aspect of interaction by describing three elements of experiential interactions.[12] The first is a sense of fluency—that is, interactions are more or less automatic, need no attention, and allow us to concentrate on the results, similar to riding a bicycle and enjoying the weather. The second element of interaction is a sense of cognition. Here, the focus is on the product at hand and can result in knowledge, confusion, or the making of mistakes if the user does not know the product or is learning to use it. Imagine that when choosing the experiences, the user wants the experience of learning how to shoot a bow. For this experience to be as realistic as possible, some effort would need to go along with the learning process. Experience designers might ask users how steep a learning curve they want. Whether they opt for a slow or quick learning uptake, the experience needs to be accompanied by a sense of change in the user, of having learned something.

10 Forlizzi and Battarbee, *Understanding Experience*.

11 Ibid., 262.

12 Ibid.

The third element in interaction is expression: "In expressive interaction, users may change, modify, or personalize, investing effort in creating a better fit between person and product."[13] Here, the interactions come with narratives, such as learning to shoot a bow for participating in a competition or for hunting. Of course, these three aspects of experience can also be tied to the broadly Deweyan understanding of experience. As in Dewey's understanding, they are related to our constant stream of experiencing continuous events while we are conscious; to the possibility of singling out one particular experience with a beginning and end, like that of a good meal experience; and to understanding experience and experiencing as shaped by social contexts.[14] The experience machine thus epitomizes the ideal experience design by providing a framework for all the different kinds of interactions, and hence from where different kinds of experiences spring.

Hassenzahl would agree with much of these descriptions—especially interactive products having the power to shape what we feel, think, and do, thus influencing our experience.[15] However, he emphasized that a deeper study of experiences has been neglected in the practice of designing interactive products. He undertook such a study by combining the three experiential interactions just described with goal-oriented behavior. For Hassenzahl, experiences involve "…abstract *be-goals*, which provide meaning, motivation, and emotion to an activity; *do-goals*, which capture concrete, desired outcomes of activities and plans to achieve those outcomes; and *motor-goals*, which regulate activities on an operational level—grabbing, dragging, pressing buttons and so forth…."[16]

Thus, experience design ought to take all three levels of goals into account, addressing them in the design process. Aligning all these goals, Hassenzahl proposed a core for all meaningful and positive experiences—namely, pleasure is the fulfillment of psychological needs, and the role of products is to be instrumental—"that is, able to create or shape the experience as desired."[17]

Hassenzahl's proposal of experience design is understandable within Nozick's experience machine, since it incorporates be-, do-, and motor-goals in creating the most pleasurable experience fulfilling the psychological needs of an individual. Furthermore, as Nozick and Hassenzahl claimed, the latter in a true utilitarianist spirit, negative experiences can be part of this if they serve pleasure as a more valuable end.[18] Hence, postponing the satisfaction of some needs might serve as a final higher pleasure, such as experiencing strain as part of a learning process. Thus, the experience machine seems, within a generous interpretation, to capture the entire motivation of doing experience design, especially Hassenzahl's version.

13 Ibid.

14 Ibid., 263.

15 Hassenzahl, *Experience Design*, 8.

16 Ibid., 29 (italics in the original).

17 Ibid., 57.

18 Ibid., 31.

Now, as will be argued below, this comes with a challenge as well—namely, that understanding experiences as the fulfillment of needs fails to capture all the putative users' experience, indicating a sense of experience where care and commitment matter.[19]

A New Interpretation I

The experience machine has been interpreted as approximating the dichotomy between those who argue for and those who argue against pleasure—hedonism—as the primary goal of human experience. One side of the debate claims that Nozick's argument does not necessarily undermine hedonism[20]; for example, "they usually offer some account as to why people's alleged preference for reality ends up supporting, rather than conflicting with, their favored version of hedonism."[21] The other side believes that Nozick's argument is spot on: Hedonism is flawed, and if people were to choose between pleasurable experiences and a broader experiential reality, they would choose the latter.[22]

However, I see another way of interpreting the experience machine: I suggest that Nozick's argument is neither for nor against hedonism, nor is it designed to conclude that people prefer living in contact with reality. Instead, in my interpretation, Nozick reminds us that pleasure cannot be all that matters to people. In this sense, Andrew Moore places Nozick in the "not only" department of objections against hedonism because people value many things besides pleasure.[23] However, Moore fails to consider the things that underscore the *practical manner in which* living a life matters to people. Elaborating on Nozick here, I suggest that this practical manner in which people care could be depicted as a sort of "existential commitment" (using Haugeland's words).[24]

Supporting this argument is Felipe De Brigard's claim that the way the experience machine has been interpreted—namely, as either for or against pleasure as the goal of experiences—is not the way forward.[25] However, the experimental "testing" of Nozick's thought experiment, as in De Brigard's own case, is probably not the only way forward either. For example, Y. Michael Barilan claims what might be interpreted as the *opposite* result of De Brigard, albeit within a different experimental setting.[26] Barilan's experimental results were based on the observation that many terminal patients do not wish for terminal sedation or euthanasia, despite suffering terribly. Instead, terminal patients insist on "…the 'right to die,' linking death itself, not merely the absence of suffering, to the values of human dignity and self-determination."[27] In the practical manner already described, and almost as with some sort of ethical guidance, Barilan suggests that "[t]he ultimate goal of palliative care is expected not to be limited to dealings with physical suffering

19 See Joseph Rouse, *How Scientific Practices Matter: Reclaiming Philosophical Naturalism* (Chicago: University of Chicago Press 2003); Andrew Sayer, *Why Things Matter to People: Social Science, Values and Ethical Life* (Cambridge: Cambridge University Press, 2011); and Derek Parfit, *On What Matters* (Oxford: Oxford University Press, 2011).

20 See, e.g., Matthew Silverstein, "In Defense of Happiness: A Response to the Experience Machine," *Social Theory and Practice* 26 (2000): 279–300; and Torbjörn Tännsjö, "Narrow Hedonism," *Journal of Happiness Studies* 8 (2007): 79–98.

21 Felipe De Brigard, "If You Like It, Does It Matter If It's Real?" *Philosophical Psychology* 23, no 1 (2010): 44.

22 See, e.g., Will Kymlicka, *Contemporary Political Philosophy: An Introduction* (Oxford: Oxford University Press, 1990); John Lemos, "Sober and Wilson and Nozick and the Experience Machine," *Philosophia* 29, no 1–4 (2002): 401–09; John Lemos, "Psychological Hedonism, Evolutionary Biology, and the Experience Machine," *Philosophy of the Social Sciences* 34 (2004): 506–26; and Daniel M. Hausman, "Hedonism and Welfare Economics," *Economics and Philosophy*, 26 (2010): 321–44.

23 Andrew Moore, "Hedonism," *Stanford Encyclopedia of Philosophy (SEP)* (2004), http://plato.stanford.edu/entries/hedonism/ (accessed May 21, 2021).

24 Haugeland, *Having Thought*.

25 De Brigard, "If You Like It," 53.

26 Y. Michael Barilan, "Nozick's Experience Machine and Palliative Care: Revisiting Hedonism," *Medical Health Care and Philosophy* 12 (2009): 399–407.

27 Barilan, "Nozick's Experience Machine," 403.

(to avoidant hedonic goals), but to sustain a vision of good death in respect for persons and their autonomy."[28] Thus, in some cases, living a dignified life matters more than the absence of discomfort.

Despite their differences, De Brigard and Barilan both confirm Nozick's objective: that in different experiential situations, different things seem to matter. This difference *might* involve a craving for pleasure, or the possibility of choosing a dignified ending of one's life. Nozick thus claims that "[p]erhaps what we desire is to live (an active verb) ourselves, in contact with reality. (And this, machines cannot do *for* us.)"[29]

Thus, if there *is* a primary desire for something, and if our experiences do have one specific goal, as Hassenzahl claims, then it is for living *per se*. Experience machines cannot do this—that is, they cannot live our lives for us. As a tentative understanding in terms of experience design, this statement probably implies that *living* cannot be exhausted by aiming at designing for (user) experiences only; instead, we need to consider the broader context of life in which these experiences are embedded. The conjecture here is that living is dependent on a sense of caring—a human concern not characterizable *per se* as the fulfillment of needs. This follows from the incentive behind De Brigard's claim that "… many things we value we just don't value as commodities."[30] Things and people matter, then, in ways that are unimaginable for a machine to handle or a design to fulfill, at least for a first impression.

A New Interpretation II: What Matters

In several places where he discusses the experience machine, Nozick uses the term "matters," which Matthew Silverstein claims is highly significant.[31] For some unknown reason, Silverstein does not discuss Nozick's early work (1974) but only a later reformulation of the experience machine (1989).[32] However, the concept of mattering plays a role in Nozick's later and early work. One sense is rhetorical, as in "*What else can matter to us other than how our lives feel from the inside?*"[33] In another sense, "matters" exemplify cases, where someone or "… something matters to us, in addition to experience…."[34] What could Nozick mean here? One suggestion is that situations occur where our response to what matters cannot be understood, invoking experience as desire or fulfillment of pleasure or needs. Instead, experience is *caring about what matters*, and Nozick provided three examples of argument: the argument from acting, the argument from being somebody, and the argument from being self-defeating.

The argument from acting says that, in most cases, people want to do things and not just experience the doing of them.[35] This distinction is difficult to draw because doing anything without experiencing it at some level is hard; Nozick's idea may be that most

28 Ibid., 406.

29 Nozick, *Anarchy*, 45.

30 De Brigard, "If You Like It," 54.

31 Silverstein, "In Defense of Happiness," 286.

32 That is, he focuses on Robert Nozick, *The Examined Life* (London: Simon & Schuster, 1989), rather than on Nozick's Anarchy.

33 Nozick, *Anarchy*, 43 (italics in the original).

34 Ibid., 44.

35 Ibid., 43.

people, when they are involved in activities (and not just highly specialized kinds of activities), presuppose or are concerned about the correctness of the activity. This "correctness" involves criteria beyond the experience of just doing the activity, making it matter in a way surpassing the experience itself. Thus, the old saying, "learning by doing," not only is correct but turns out to be imperative. To illustrate, I do not simply want the experience of cycling; instead, I want to *do* it—to know what about it matters as an activity. In learning these matters, I am holding myself responsible to standards (e.g., traffic rules, health issues, fixing of bikes, etc.) and not depending on my experience alone.

Nozick's argument from being somebody claims that, plugging into the machine, you are not really anybody. "There is no answer to the question of what a person is like…" when that person has been plugged into the machine a long time. Hence, "…plugging into the machine is a kind of suicide."[36] Presumably, Nozick is suggesting that what makes a person a person cannot be confined to how this person experiences herself. Being a person means being recognized as a specific person—that is, by other people *telling* and *answering* (for) who you are. Now, a whole crowd of people for whom one matters can be part of the machine, but they are a result of one's experience only—that is, they are selected from the "smorgasbord of experiences," in Nozick's memorable phrase. Thus, the suicide to which Nozick alludes may amount to an inability to recognize oneself within the machine because all the criteria needed to do so (i.e., to recognize the self as distinguished from other people) are established by reference only to one's own (chosen) experiences.

The argument of being self-defeating claims that plugging into the machine limits us to a human-made reality: "…to a world no deeper or more important than that which people can construct."[37] This argument can be divided into two parts. First, it amounts to a certain dilemma: Eventually, it will occur that the reality the machine's user is about to choose is created by his or her selection of it. But how can users then be sure that the experiences they are selecting best fulfill their needs? The chosen experiences are supposed to determine the experience of pleasure in the machine, thus presupposing that the users' predilections will remain unchanged, or will change according to what they have planned throughout the period they stay in the machine. However, it is entirely possible that those who plug into the machine cut themselves off from obtaining more happiness outside the machine—either by chance, by forgetting, or simply by not knowing about something or someone that could make them happier outside the

36 Ibid.

37 Ibid.

machine. The outcome is that a user's happiness also depends on a world not of her making and that what she chooses could actually be self-defeating.

The second part of the argument of being self-defeating questions whether choices in a human-made reality are actually choices made in a person's self-interest. This argument can be clarified using Derek Parfit's examples of choices, which are "directly individually self-defeating."[38] Rephrasing Parfit's description,[39] we might say that a choice (C) is indirectly individually self-defeating when it is true that, if someone tries to achieve his aims by doing C, these aims will be, overall, worse achieved. The aim of plugging into the machine is to achieve pleasure as a fulfillment of needs. Could the user, in plugging in, actually be worse off or further from achieving that aim? If so, then plugging into the machine would be self-defeating. To illustrate, suppose the user is standing in front of the smorgasbord of experiences, selecting all the experiences he desires. He is meticulous, making sure that all the things he knows will fulfill his needs will be part of the experiences. He cannot select any experience that is too unfamiliar because he cannot know whether it would give pleasure. In selecting experiences in the vicinity of what he already knows is pleasurable, his happiness is, in a certain sense, limited by past experiences of what makes him happy. The possibility thus exists that one or more experiences from the smorgasbord might make him happy, altogether surpassing the chosen experiences in the machine, but they are not chosen because they are not known. Plugging into the machine would therefore not be in the user's self-interest because he faces the risk of being worse off than not choosing to plug in.

All three arguments—the last clearer than the other two—suggest that considering plugging into the machine, the user realizes that he or she is committed in ways exceeding the frame of understanding the machine offers. One discovers a world that matters—a world that is not reducible to the experience of it, and the caring for this world sometimes implies adopting a less human-centered perspective. As noted previously, this view seems close to Wright and McCarthy's claim:

> The question and the challenge for experience-centered design researchers and practitioners is to find the means of harnessing the potential of experience-centered design to bring about change that really matters to people and to society and to open up a space where designers can have the opportunity, once again, to help shape society towards some of the bigger issues of significance in our lives both individually and as a society.[40]

38 Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984),1–54.

39 Ibid., 5.

40 Wright and McCarthy, *Experience-Centered Design*, 73.

The argument I present here agrees with the challenge presented for experience design. However, Wright and McCarthy's notion of what really matters to people is vague and fails to clarify how we are to understand it. Furthermore, it allows for opposite perspectives to be equally valid, indicating the need to address different levels of "mattering." For example, to a smoker, cigarettes matter; to his doctor, his stopping smoking is what matters. Obviously, the doctor is right regarding what *really* matters, but can a general distinction be made that makes these levels of mattering explicit?

A better approach would be the human-centered design (HCD) proposed by Hargraves.[41] Here, the mattering of people—involving care—serves as a condition. Experience in this case "…is not simply the impact of a product, nor a class of design products. Experience at its core is concerned with human life and living, and design's participation in its honoring, cultivation, shaping and reshaping."[42] Taking his departure from IDEO'S understanding of HCD, Hargraves claims that experience thus implies that capacities and opportunities for exercising human power to make a difference in the world become important. However, he also warns—rightly—against this notion of opportunity becoming a sole demonstration of what designers can do, while failing to provide us with an understanding of what thereby is "…to be centered by the human, or the person, or the patient in design."[43] Asking instead *why* designers should do what they do, Hargraves follows Buchanan, claiming that the answer to this "why" turns not on usability or user experiences, but on strengthening human dignity. Design should support the dignity of human lives in different social, economic, political, and cultural circumstances.[44]

I agree with the incentive behind Hargrave's point while also asserting that human dignity is essentially a contested concept. It can be used to argue both for abortion, protecting the mother, and against abortion, protecting the right of the fetus. Furthermore, and more worrisome in our Anthropocene age, it might block the insight of developing an approach to design that is human decentered or post-human[45]—where the actions of humans must be contextualized, put into perspective, or seen as needing to be naturally limited by a world where other species and the limits to natural resources is center as well.[46] In the following section, I explore how Haugeland (and Lance) can help, pointing to a notion of commitment capable of incorporating a caring for both human-centered and -decentered matters.

Existential Commitment

It was claimed above that there was a significant way of interpreting Nozick's experience machine without understanding it as an argument either for or against hedonism, as it is often interpreted.[47] Nozick's arguments from acting, from being somebody, and from

41 Hargraves, "Care and Capacities."

42 Ibid., 81.

43 Ibid., 86.

44 Ibid., 87.

45 Bo Allesøe Christensen, "Why Do We Care About Post-Humanism? A Critical Note," in *Geografiska Annaler. Series B. Human Geography* 96, no. 1 (2014): 23–35.

46 A decentered human perspective does not imply a neutral perspective or a god's eye view. Rather, it means that caring for what matters comes with the possibility, or necessity, of setting aside a predominant focus on human needs, desires, values, or experiences. Hence, a focus on sustainability might mean giving up on certain human needs or desires, thereby decentering the human perspective.

47 De Brigard, *If You Like It*; Barilan, *Nozick's Experience Machine*.

the self-defeating argument(s) were all three taken to show that things, persons, or circumstances could matter through their importance, exceeding our possible experience of them. Relating this insight to experience design indicates that experience is more than Hassenzahl's fulfillment of needs or a human-centered perspective. It involves holding oneself responsible for and caring about matters not of one's making, and in terms of the three arguments, it involves doing so in practice, as a social relation and as a complex uncontrollable context. Furthermore, note that the subject matter of the three arguments can be related to but also *expand* Hassenzahl's three goals that makeup experience: motor (i.e., the concrete experiencing), doing (i.e., pursuing one's goals as dependent on other people), and being (a complex context functioning as a resource of meaning). This expansion can be elucidated by turning to American philosopher John Haugeland's distinction between deontic and existential commitments.

Haugeland notes that most of our engagements with our surroundings are accompanied by a commitment to some *what* and some *who* of this engagement. Take the experience of playing a game of chess. Partaking in a game of chess commits you to certain rules of moving the pieces in distinct ways, to not using your opponent's pieces, and, upon winning or losing, to behave in a way suitable for the character of chess. Hence, playing chess commits you, in a specific sense, to the rules, with the pieces and the board, toward the opponent, and perhaps some audience—all things, manners, and persons that matter when being engaged in a chess game. Refusing to abide by the rules is not necessarily tantamount to not playing; but what one is playing, then, is just not chess.

Haugeland divides commitments into two categories: deontic and existential. A deontic commitment is an obligation or duty, a way that one is supposed to behave: "Making a commitment to the rules of the game means, in this sense, undertaking an obligation to play by the rules—say, by entering into (or implying) an agreement. Someone who fails to abide by such a commitment is corrected, or, if incorrigible, rejected as a player."[48] However, the existential commitment is more than an obligation; it is a dedicated way of living, "…a determination to maintain and carry on."[49] This is a resolute commitment, "…not 'to' other players or people, or even to oneself, but rather to an ongoing, concrete game, project, or life…. [I]t is a *way*, a *style*, a *mode* of playing, working or living… that relies and is prepared to insist on that which is constitutive of its own possibility, the conditions of its intelligibility."[50] Hence, marriage and a monastic vow are expressions of deontic commitment, whereas love and faith are existential commitments. The latter is not an obligation in the sense of just doing one's duty to other people. Instead, it is a manifestation of the resoluteness of taking responsibility for the conditions of the possibility of a way, style, or mode of living.

48 Haugeland, *Having Thought*, 341.

49 Ibid.

50 Ibid.

This commitment is not just a matter of following the rules implied by the obligations; it is a concern for the possibility of this way of living (including the *what* and the *who*) to which one is committed. It is a consideration not for the experience in itself but for *what* and *who* make this experience possible.

The two kinds of commitments entail two kinds of responsibilities, according to Haugeland. Being responsible in a deontic sense is being responsible only for the character and consequences of one's *own* behavior. Thus, the responsibility that the action of plugging into the experience machine entails—namely, what a person is committing to by being in the machine—revolves around no other than this person.

In this regard, when Hassenzahl wrote about the normative powers of design and referred to designers who make overly wasteful cars, thus disregarding their responsibility by claiming that this design is what consumers want, they lack deontic responsibility.[51] As Hassenzahl phrased it, the designer must "…[take] on the responsibility for the particular way of fulfilling a need she or he just devised. Each product is a proposition, and we cannot escape the fact that it has the power to change how people feel, think, and act. To do this consciously is important."[52] These designers fail to undertake the obligation of playing by the rules because they fail to consider the normative powers of design.

In contrast, existential responsibility is a responsibility encompassing the conditions for which a given commitment depends. Therefore, it entails "…responsiveness to the constituted phenomena, in particular with regard to their compliance with the standards in accord with which they are constituted."[53] To use the chess example, a chess player is committed not only to following the rules but also to observing whether the opponent is playing by the rules. However, this commitment is not tantamount to "…an agreement to play by the rules, on pain of being rejected, but rather [is] an involved insistent *way* of responding and playing, so of finding things and dealing with them, on pain of 'giving up the game.'"[54] Most of Haugeland's examples are taken from a natural scientific context, where responsibility is tied to scientists' dealings with incompatibilities disclosed through their work. Existential responsibility means responding to this incompatibility in a non-conformist fashion; it is unacceptable and thus requires a willingness to disregard previous scientific practices, replacing them with new ones so that the incompatibilities do not show up again. Nevertheless, existential responsibility works in the case of the car designers, and in more mundane cases as well.

Returning to the car designers in Hassenzahl's example, the *what* would indicate thinking about what transportation as a way of life would mean. It would have to expand the dimension of user

51 Hassenzahl, *Experience Design*, 73–74.

52 Ibid., 74.

53 Haugeland, *Having Thought*, 342.

54 Ibid.

experience to incorporate knowledge (authoritative criteria) from researchers in the sciences. The commitment is to working with and learning from new and developing materials, investigating alternative practices of transportation, developing new infrastructures, and being concerned with sustainable economies. The *who* would incorporate a social dimension—that is, that we answer to other people in the present and future. This commitment minimally would involve thinking about how the issue of transportation is a matter of connecting personal and public values and interests and would include caring as the developing of an ethical awareness that incorporates guidelines, such as the precautionary principle or a cradle-to-cradle dimension in our design practices.

Existential Commitment and Design

This perspective on experience broadens the sense of normativity in the normative powers of design, to which Hassenzahl referred. Rather than merely considering the effect of design on people, we move toward inquiring into the overall conditions and possibility of this normativity and include a decentered human perspective in Hargraves's human-centered design. Therefore, in experience design, when particular activities are suggested through an interactive product, hopefully leading to a good and appropriate experience of need fulfillment, such suggestions will be expanded by elaborating on the appropriateness of the interactive product in relation to a way of living.[55] This view connects the designer and consumer in a mutual commitment as partakers and caretakers—not of the car business as such, but of transportation and its consequences on our way of living. The designer and the consumer, therefore, both have a responsibility to consider the fulfillment of needs, the design, and the desire for an experience, as constrained by a commitment to a way of living as mattering more than the fulfillment of a particular need or a pure human-centered perspective. In congruence with Nozick's interpretation, this perspective is not for or against understanding experience as the satisfaction of desires or needs; instead, it claims experience as embedded within a broader context of commitment and responsibility—that is, within a life that matters. What humans experience matters to them; they *care* about it in more ways than just centering on their needs and values.

To add a final facet to this perspective, the focus on caring and mattering reveals that experience design already is connected to political and sociological discussions as part of the normative power of design. Likewise, and as a result, it is connected with different strands of critical design and existential design, as well as post-phenomenologically–inspired design methodology.[56]

55 Hassenzahl, *Experience Design*, 74.

56 See Carl Mitcham, "Dasein Versus Design: The Problematics of Turning Making into Thinking," *International Journal of Technology and Design Education* 1, no 1 (2001): 27–36; Anthony Dunne and Fiona Raby, *Speculative Everything: Design, Fiction, and Social Dreaming* (Cambridge, MA: MIT Press, 2013); and Carl DiSalvo, *Adversarial Design* (Cambridge, MA: MIT Press, 2015). See also Merlijn Smits et al., "Values that Matter: A New Method to Design and Assess Moral Mediation of Technology," *Design Issues* 38 (Winter 2022): 39–54.

Closing

The aim of this study was to offer a theoretical discussion of the notion of experience, departing from a reinterpretation of Nozick's concept of an experience machine as a philosophical design fiction and connecting to different notions of experience design. Nozick's experience machine implies that understanding experience as the fulfillment of needs, as pleasure, or as too human centered is limited. These notions of experience fail to consider the commitments and responsibilities related to the way experience is embedded in and responsive to a complex reality that is not of its making. We explored this notion of experience using John Haugeland's idea of caring as a form of existential commitment, and we tentatively sketched the consequences of using it in relation to design in a very general fashion. This philosophical design fiction has allowed us to question the overall human centeredness of "experience" as used in experience design, pointing us toward a sense of the decenteredness involved when experience means caring for what matters.

The Influence of Fonts on the Reading Performance in Easy-to-Read Texts: A Legibility Study with 145 Participants

Sabina Sieghart

What Is Easy-to-Read Language?

Easy-to-read language is used as a tool for barrier-free communication and is becoming increasingly important. Bureaucratic offices, institutions, and companies have been legally obliged to make information freely accessible since the United Nations Convention on the Rights of Persons with Disabilities came into force in 2009.[1] The concept of *easy-to-read language* aims to provide people with cognitive disabilities with information in a form that they can understand, thereby enabling them to participate in our literate society. For this purpose, the texts are linguistically simplified and visually preprocessed. The practice is characterized by rigidly defined linguistic and typographic rules.[2]

The most widespread German rulebook requests a sans-serif font and recommends the Arial font in at least a type size of 14 point as a good example.[3] However, this recommendation contradicts the findings of general legibility research.[4] The font recommendation is based on the practice of the German *easy-to-read language*, reflecting the possibilities of lay designers who only had MS Word system fonts at their disposal.[5] These recommendations, which were not supported by research, were repeatedly and unquestioningly implemented and cited in several other guidelines. In fact, no study thus far has examined the effect of the fonts on reading with the target group. The study on which we report in this article aims to close this gap and to provide a scientific basis for making informed font selection decisions.

How Important Is Font Selection for a Text's Understanding?

Thanks to recent scientific findings, we know that reading is a complex process for which our brain has no predisposition.[6] Visual areas must be linked to speech areas to connect the visual input of a written word with meaning and pronunciation. The first step in the reading process is to *decipher* or *decode* what we are seeing. The visual information must first be absorbed before it can be provided with linguistic meaning in a complex process. Only then do other

1 https://www.behindertenrechtskonvention.info/inklusion-3693/ (accessed December 13, 2022).

2 Bettina Zurstrassen, "Inklusion durch Leichte Sprache? Eine kritische Einschätzung," [Inclusion Through Easy-to-read Language? A Critical Assessment] in *Didaktik der inklusiven politischen Bildung* [Didactics of Inclusive Civic Education], ed. Christoph Dönges, Wolfram Hilpert, and Bettina Zurstrassen (Bonn: Bundeszentrale für politische Bildung, 2015), 126–38.

3 *Bundesministerium für Arbeit und Soziales* [German Federal Ministry of Labour and Social Affairs], ed., Leichte Sprache. Ein Ratgeber [Easy-to-read language: A guidebook] (Bonn: Publikationsversand der Bundesregierung, 2014), 52–56.

4 See Stanislas Dehaene, *Reading in the Brain: The New Science of How We Read* (New York: Penguin and Wolf, 2009); and Maryanne Wolff, *Reader, Come Home: The Reading Brain in a Digital World* (New York: Harper, 2018).

5 Sabina Sieghart, "Leichte Sprache, Typografie & Angemessenheit" [Easy-to-read language, typography & appropriateness] in *Lesbar. Typografie in der Wissensvermittlung* [Readable: Typography in knowledge transfer], ed. Rudolf Paulus Gorbach and Ulrike Borinski (Zürich: Triest Verlag, 2019), 124–32.

6 Antonia Cornelius, *Buchstaben im Kopf. Was Kreative über das Lesen wissen sollten, um Leselust zu gestalten* [Letters in the head: What creatives should know about reading to create passion for reading] (Mainz: Verlag Hermann Schmidt, 2017), 48.

Design*Issues*: Volume 39, Number 3 Summer 2023 https://doi.org/10.1162/desi_a_00724

Font Arial

Font Thesis TheSerif

Figure 1
Similarity of first three letters of the word *Illustration* in the Arial font. (Other fonts, like Thesis TheSerif, have clearly recognizable letters.)

cognitive processes come into play, enabling the content of the text to be understood. A font that facilitates this first step is described as "legible."[7]

The choice of font is crucial: "Letterforms that correspond to the physiology of the retina and the visual cortex" facilitate reception.[8] Letters must be recognizable and distinguishable and have open forms; they also must have a good amount of stroke contrast.[9] This feature is particularly applicable for people with visual impairments[10]—but it does not apply to the Arial font. For instance, the letters lowercase *l* and capital *I* are not visually different enough (see Figure 1). In addition, the gray value is too dark, and the *a* and *s* are too closed.[11]

Studies of the reading process have shown that the decoding of characters takes place on a purely physical level: the "psycho-motor domain"[12] or the "sensory level."[13] The question is whether the eye apparatus (i.e., the physical prerequisites for decoding) works in the same way for readers of *easy-to-read language* as for non-impaired readers.

The design of the glyphs of a font, as described, can be subsumed under the term, *nano typography*.[14] In the course of the reading process, other typographical factors come into play. Here, we can distinguish between the areas of micro and macro typography. Micro typography describes the fine adjustments in typesetting, such as font size, line spacing (i.e., leading), and line length, as well as forms of typographic emphasis, which can include the use of bold, italics, and small caps, and orthographic decisions, like the use of the correct quotation marks. The resulting white spaces within and between the lines are relevant to the reading process. If they are "physiologically beneficial and semantically clearly organized," the decoding process will be easier.[15]

Macro typography refers to the layout—that is, the arrangement of the text on the medium. Through typical arrangements of patterns, design genres develop, which give the reader visual clues about the content to be expected. A previous study showed

7 Ann Bessemans, "Matilda: A Typeface for Children with Low Vision," (conference presentation, *Typo Day*, Puna, India, February 28, 2014).

8 Martin Tiefenthaler, "Zuerst liest der Körper, dann der Geist [The Body Reads First, Then the Mind] in *Lesbar. Typografie in der Wissensvermittlung* [Readable: Typography in Knowledge Transfer], ed. Rudolf Paulus Gorbach and Ulrike Borinski (Zürich: Triest Verlag, 2019), 28.

9 Cornelius, *Buchstaben im Kopf* [Letters in the Head], 26.

10 Florian Adler, Tim de Gruisbourne, and Claudine Stolzenburg (2017) *leserlich. info – Schritte zu einem inklusiven Kommunikations design* [Steps to an Inclusive Communication Design], *Online-Plattform des Deutschen Blinden- und Sehbehindertenverbands* [online platform of the German Association for the Blind and Visually Impaired], http://www.dbsv.org/leserlich/kapitel/leserlich-und-lesbar.php (accessed December 13, 2022).

11 Ralf Turtschi, "Arial: ein Nekrolog" [Arial: A Necrology], *Publisher – Schweizer Fachmagazin für Publishing und Digitaldruck* [Publisher—Swiss trade magazine for publishing and digital printing] 2 (2005): 38.

12 Robert Waller, "The Typographic Contribution to Language: Towards a Model of Typographic Genres and Their Underlying Structures" (PhD diss., University of Reading, 1987), 24.

13 Ann Bessemans, "Matilda: A Typeface for Children with Low Vision," in *Digital Fonts and Reading*, ed. Mary Dyson and Ching Y. Suen, Series on Language Processing, Pattern Recognition, and Intelligent Systems (Singapore: World Scientific, 2016), 22.

14 The term nano typography is relatively new and was first used jokingly by Karin and Bertram Schmidt-Friedrichs in the preface to their book, Karen Cheng *Anatomie der Buchstaben* [Anatomy of the Letters], (Mainz: Verlag Hermann Schmidt, 2013). See also Johannes Bergerhausen, "Nano, Mikro, Makro," in *Subtext: Typedesign*, ed. Andreas Pawlik and Martin Tiefenthaler (Salenstein: Niggli Verlag, Schweiz, 2017), 433.

that readers of *easy-to-read language* are familiar with the typical design genres and that they use macro-typographical cues to generate meaning.[16]

In summary, nano and micro typography make text legible at a level of detail, while macro typography makes text readable through structuring at a meta-level. However, such legibility does not guarantee that the text will be understood; a book in an unknown language remains incomprehensible, even with legible typography. Ralf Hermann describes this differentiation, which was made as early as 1796 by German linguist Johann Christoph Adelung, and concludes that "Readability is always dependent on the recipient, in contrast to the rather universal definition of legibility."[17] In addition to linguistic and typographical levels, other factors also influence one's comprehension, such as the reader's environment, prior knowledge, and cognitive abilities.

Criticism and Potential for Improvement of Previous Legibility Studies

Legibility studies have been conducted for 150 years.[18] For almost as long, studies have addressed the question of whether fonts are more legible with serifs or without them. Two meta-studies summarize the problems of most such studies. In 1999, typographer Ole Lund counted 72 studies on this topic and analyzed 28 of them. He certifies that almost all studies have serious errors in the structure of the test material and a lack of "internal validity."[19] That is, whether a causal relationship exists between the manipulation of the independent experimental variables and the observed outcomes is not clear. In addition, the "external validity" (here, he means the coherence between the test situation and the target situation) often is not given, either.[20] In 2022, cognitive psychologist John Richardson criticizes the lack of empirical evidence in various studies and the biased approach of some researchers. His conclusion—that "[t]here is no difference in the legibility of serif and sans serif"—is highly questionable for two reasons.[21] First, he treats the results of all studies equally, although he questions the empirical validity of some of them. Second, that he does not seem to have penetrated the complex terrain of typography is striking, and therefore, unlike Lund, he is unable to assess the internal validity of the studies.

Meanwhile, design researcher Ann Bessemans sees one reason for the lack of validity in the fact that these studies were developed either by cognitive scientists with insufficient professional knowledge of typography or by designers with a lack of scientific competence. On the one hand, insufficient test material is created or selected, the wrong terminology is used, and poorly designed fonts

15 Tiefenthaler, "Zuerst liesst der Körper" [The Body Reads First], 28.

16 Sabina Sieghart, "The Influence of Macro Typography on the Comprehensibility of Texts in Easy-to-Read Language: An Empirical Study," *Visible Language* 54, no. 3 (2020): 48–95.

17 Ralf Hermann, "Das Zwiebelschichtenmodell der Lesbarkeit" [The Onion Layer Model of Readability], *Typo Journal*, no. 2 (2010): 3.

18 Sofie Beier, Reading Letters, Designing for Legibility (Amsterdam: BIS Publishers, 2012), 10–20.

19 Ole Lund, "Knowledge Construction in Typography: The Case of Legibility Research and the Legibility of Sans Serif Typefaces" (Master's thesis, University of Reading: 1999), 247.

20 Ibid., 82–83.

21 John T. E. Richardson, *The Legibility of Serif and Sans Serif Typefaces: Reading from Paper and Reading from Screens* (Cambridge: SpringerBriefs in Education, 2022), 130.

are tested. On the other hand, designers tend to present their "findings" as truths that are not scientifically backed. Bessemans also concludes that "many legibility studies focusing on the influence of design, both within cognitive science and within the design world, lack internal and/or external validity."[22] This shortcoming began to change in 2006 in Europe, when the famous typographer, Gerard Unger, became a professor of typography at Leiden University.[23] In this position, he was able to supervise various PhD researchers who conducted legibility experiments in a methodologically sound manner, thus producing a series of valid legibility studies (although not in my field of research with *easy-to-read language*).

Criteria for Valid Legibility Studies

Based on the criticisms of Lund, Richardson, and Bessemans, I identified eight factors that affect the validity of legibility studies. My goal was to provide a set of specifications that would serve as a theoretical foundation—not only for this study but also for future studies. I established the requirements for this study in close dialogue with typographer Albert-Jan Pool, to whom I offer my gratitude for his valuable input on the study's conceptual design. The eight factors are:

Factor 1: For font comparisons, the fonts need to be shown so that their x-heights are the same size. Fonts often have been compared at the same point size. This approach falsifies the results because the optical size of a font is neglected. In mixed-case settings, this size is determined by the x-height.[24] However, the "point" as a unit of measure refers to the body height of the traditional movable metal type. Because a font can have long or short ascenders or descenders, cap heights may differ by up to 20 percent at the same point size; x-heights differ between fonts by up to 40 percent and descenders by up to 30 percent.[25]

Factor 2: Equal contrast ratios are crucial. If completely different fonts are compared with each other, making a valid statement as to which detail is decisive for legibility is impossible. The contrast ratios of the fonts (and with them, the various form principles, stroke thicknesses, and gray values) must be considered.[26] Therefore, a font family with and without serifs should be chosen because they are based on the same letterforms and, thus, contrast ratios are equal.

Factor 3: Test fonts have to be of sufficient quality. So-called "Labo fonts" (laboratory typefaces) are developed exclusively for experimental research purposes and do not meet type design's quality requirements.[27] They sometimes have bizarre letter shapes that deviate too much from the formal principles of typefaces intended for immersive reading.[28]

Factor 4: Typography's complex interrelationships need to be considered. Superficially, only changing one variable and leaving the rest in common seems sensible. However, this approach does

22 Ann Bessemans, "The Gap Between Science and Typography," in *Design for Visual Communication Challenges and Priorities*, ed. Klimis Mastoridis, Niki Sioki, and Mary Dyson (Cambridge: Scholars Publishing, 2019), 20.

23 Ann Bessemans, "Academic Research into Typographic Design at the Beginning of a New Era," in *Visible Language* 56, no 2 (2022), 85.

24 Björn Schumacher, "Leserlichkeit von Druckschriften heute – Möglichkeiten empirischer Leserlichkeitsforschung," [Readability of Printed Matter Today: Possibilities of Empirical Readability Research] in *Lesbar. Typografie in der Wissensvermittlung* [Readable: Typography in Knowledge Transfer], ed. Rudolf Paulus Gorbach and Ulrike Borinski (Zürich: Triest Verlag, 2019), 240.

25 Deutsches Institut für Normung (DIN) [German Institute for Standardisation], *Schriften—Leserlichkeit* [Lettering—Legibility] (Berlin: Beuth Verlag, DIN 1450:2013-04:4).

26 Antonia Cornelius, "Gestaltung von leserlichen Schriften – Was nützt uns die Forschung," [Designing Legible Typefaces: What is the Benefit of Research?] in *Lesbar. Typografie in der Wissensvermittlung* [Readable: Typography in Knowledge Transfer], ed. Rudolf Paulus Gorbach and Ulrike Borinski (Zürich: Triest Verlag, 2019), 234.

27 Bessemans, "The Gap Between Science and Typography," 7.

28 This applies to various Dyslexie fonts.

not account for typography's complex interrelationships. For instance, if the font size is changed, the line spacing and line length must also be adjusted.[29]

Factor 5: The typesetting has to correspond to craftsmanship standards. Test material must exhibit optimal typographic settings on the micro and macro levels and be comprehensible for typographers.[30] For instance, a text block that is set justified but is full of gaps is so detrimental to the flow that it interferes with the effect of the font selection and no longer is measurable.

Factor 6: The test methods have to be appropriate for the problem. There is a big difference between reading signs on the highway and leisurely reading a book, magazine, or newspaper. We are interested in the legibility of text in print. Miles Tinker, who conducted some of the most comprehensive legibility studies, compared various methods of assessment and identified speed-of-reading as the most appropriate method.[31] Dirk Wendt adapted and optimized the test method and enriched it with a statistical evaluation method.[32]

Factor 7: A sufficient number of subjects suitable for the test is required. Participatory testing is not only a requirement for texts in easy-to-read language but also the most appropriate methodological approach.[33] Readers of easy-to-read language are a heterogeneous group with a wide variety of limitations. In addition, the testing groups are often very small (on average, two to five people per office), and they have been testing for years according to specified test catalogs that strictly adhere to existing policies. Therefore, assembling a sufficiently large, unbiased, and heterogeneous test group is important.

Factor 8: The test situation should correspond to the accustomed reading situation. An unfamiliar environment (like a lab situation) creates commotion and stress and is detrimental to our target audience. A foreign language text or dummy text is confusing for the target audience. The test should therefore be carried out in a normal, relaxed reading situation.

Study Structure, Test Material, and Questionnaire

In this study, we compared the Arial font and the Thesis font family (five fonts in total) at identical x-heights (with a comparable contrast ratio and appearance, in terms of size), according to Wendt's optimized Tinker-Speed-of-Reading-Test.[34] We thus are examining two of the specifications from common *easy-to-read language* policies: The specification that a sans serif font like Arial be used and set at a minimum of 14 points. However, the font selection also inevitably affects another aspect: Every piece of writing has a connotation—that is, a meaning that is felt. We wanted to know whether readers of *easy-to-read language* perceive this aspect of the font—the "affective domain."[35]

29 Lund, *Knowledge Construction in Typography*, 61–65.

30 Bessemans, "The Gap Between Science and Typography," 6.

31 Miles A. Tinker, "Influence of Simultaneous Variation in Size of Type, Width of Line, and Leading for Newspaper Type," in *Journal of Applied Psychology* 47, no. 6 (1963): 380–82.

32 Dirk Wendt, "Lesbarkeit von Druckschriften" [Readability of Print] in *Lesen Erkennen* [Reading Recognition], ed. Rudolf Paulus Gorbach (Munich: Typographische Gesellschaft München, 2000), 32–45.

33 Lund, *Knowledge Construction in Typography*, 232.

34 Wendt, "Lesbarkeit von Druckschriften" [Readability of Print], 32–45.

35 Waller, *The Typographic Contribution to Language*, 28–37.

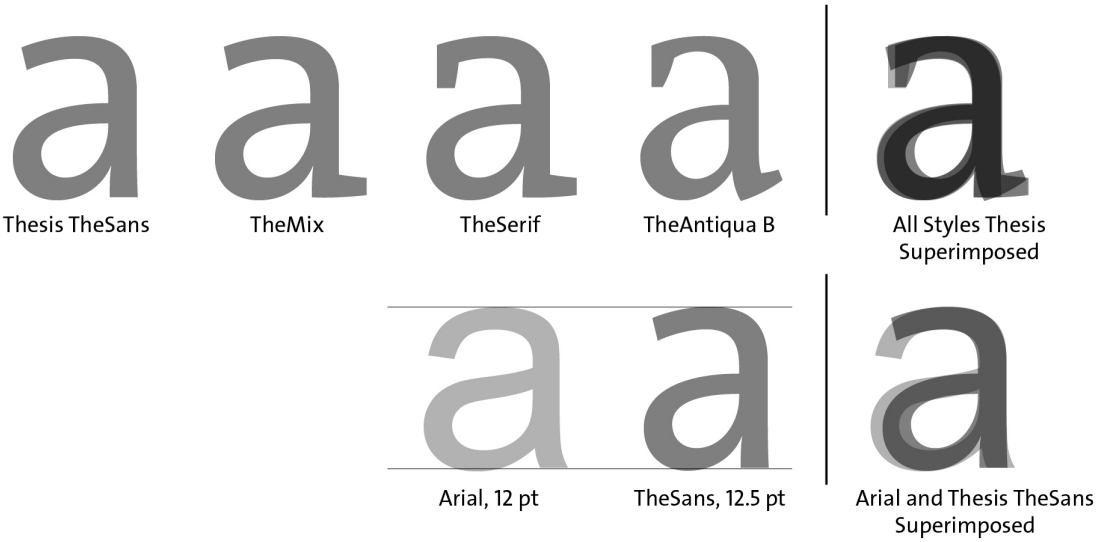

Figure 2
For fair font comparison, the letter forms need to be similar and of equal x-height. Top row: The Thesis font family (from left to right: Thesis TheSans, TheMix, TheSerif, and AntiquaB, and all styles superimposed); bottom row (from left to right): Arial, Thesis TheSans, and both fonts superimposed.

The test material consisted of booklets that used one of the five fonts on each two-page spread. We distributed five different booklet types, and the order in which the fonts appear differed in each one. Thus, we were able to compensate for the expected effect of fatigue when reading. We used five continuous sections of a novel corresponding to the A2 reading level as the text to be read. In this text, we also inserted two names (Illona and Arne) to test critical letter combinations: "l" and "I" often are hardly distinguishable from one another, and "rn" is difficult to read in sans serif fonts that are spaced too tightly or set at too small of a size; in these situations, it can be read mistakenly as an "m."[36]

The five fonts we tested included Arial and Thesis TheSans, TheMix, TheSerif, and TheAntiquaB (see Figure 2). The four different Thesis styles do not differ in the shape of their letterforms, but only in the serifs and stroke-thickness modulation. We chose the Thesis font family from Luc(as) de Groot because it is one of the most comprehensive font families and is considered very legible. In addition, the font family variant called TheAntiquaB is based on the same formal principle but has the typical stroke-thickness modulation of an medium-contrast oldstyle font.[37] These options allowed us to compare not only the individual fonts but also several formal principles:

- The effect of the formal principles in sans serif fonts: Arial vs. TheSans;
- The contribution of the serifs: TheSans vs. TheMix vs. TheSerif; and
- The influence of the modulation of the stroke thickness: TheSerif vs. TheAntiquaB.

In terms of typesetting, we paid attention to optimal typographic design and technically flawless processing.[38]

36 Cornelius, *Buchstaben im Kopf* [Letters in the head], 118.

37 Luc(as) De Groot, https://www.lucas-fonts.com/about/thesis (accessed December 13, 2022).

38 Our line spacing, line length, and balanced left-aligned, ragged-right text corresponded to good typographic practices. A baseline grid ensured that the body text remained in register—that is, the white space between lines was not marred by letters showing through from the other side of the paper. The clear structure of the lines on the page helped readers to find the beginning of the next line more easily. The booklets were in A5 format and were printed on 120-gram off-white uncoated paper, folded, and bound with staples.

The questionnaires were developed iteratively and adapted after the pilot. After each page of text, the subjects answered questions about the respective font and subsequently about their personal reading habits. We noted whether the critical letter combinations in "Ilona" and "Arne" were read correctly, as well as how the test subject felt about the font's size. We repeated this questioning process for all five texts. After the reading test, we asked questions about reading habits and reading levels.

Test Procedure, Pre-Test, and Test Groups

Each person being tested received a test set consisting of a questionnaire and a booklet. The test facilitators briefly explained the test's setup. The point was not to read as quickly as possible but to test the different fonts in a familiar reading mode. The number of characters from the respective fonts that were read in one minute was recorded. To capture this data, the test persons read the text, and a ring tone sounded off after one minute. The reader and test facilitator then marked the last letter as read and filled out the relevant part of the questionnaire.

The test material for the pre-test had an x-height of 2.38 mm (equivalent to 13.5 or 13 pts), which was rated as "large enough" by all five pilot test readers. Therefore, for the main test, we reduced the x-height to 2.25 mm, corresponding to the type size of 12 pts for the Arial font and 12.5 pts for the fonts in the Thesis family.

The main test was carried out in cooperation with capito, a social franchise network specializing in *easy-to-read language* services. In Germany and Austria capito has branches in 20 cities and, of its 565 employees, 324 have disabilities. The philosophy of capito is to select individuals who are representative of the target audiences. The primary occupation of the individuals selected should focus on something other than test-reading, and they should not hold certain concepts about how products in *easy-to-read language* should look. By incorporating the testing into capito's daily business, we were able to provide a familiar test environment. From among capito's disabled employees, 145 test subjects took part in the main test. The execution of the tests was organized by Christian Nolte. The statistical analysis was performed by Claudia Zimmermann.

Results Concerning the Font Style

The following results concern the first specification: The choice of fonts.

Which Font Is Read the Fastest?

Of 143 participants, 27.3 percent read TheSans the fastest, followed by Arial (21.7%), TheAntiqua (18.2%), TheMix (16.8%), and TheSerif (16.1%).[39] Thus, just over half of the readers read one of the serif fonts the fastest (see Figure 3).

39 Though 145 participants took part in the main test, two are excluded from this analysis (for no. 24, two fonts were the fastest; for no. 170, the character specification for one font is missing).

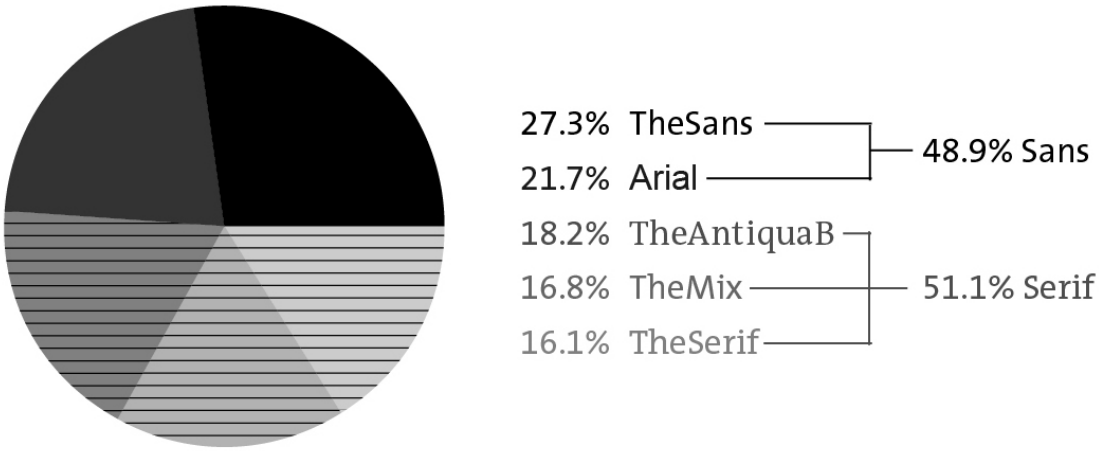

Figure 3
Fonts read the fastest (highest number of characters per reader). All graphs designed by the author.

Which Font Is Read the Fastest on Average?

The mean values of the reading speed for the individual fonts show that TheSans was read the fastest, with an average of 622.14 characters per minute; therefore, it led to a significantly faster read. The mean values of the other four fonts are very close—that is, the reading speeds for the other fonts are roughly the same, although Arial performed the worst (see Figure 4).

Characters per Minute	Arial	TheSans	TheSerif	TheAntiquaB	TheMix
N	145	145	145	145	145
Mean	596.68	622.14	600.37	596.83	601.35

Figure 4
Average characters read per minute by font.

Does the Reading Level Affect the Results?

We determined the test subjects' reading levels first in the questionnaire and then also classified the reading levels through their characters per minute. The results of the two different classifications of the reading level (questionnaire vs. reading speed) do not exhibit a particularly high level of agreement (r_s=0.292, p=0.001).[40] For the evaluation, we decided to use the classification by the number of characters per minute as our basis.[41] Accordingly, among the poor readers (35 participants), the largest proportion read texts set in TheSans the fastest (37.1%). TheMix followed in second place at a clear distance (20.0%). Our largest group was the moderate-level readers (78 participants). The font read fastest by the highest percentage of people in this group was Arial (26.9%), followed by TheSans (23.1%). Among the good readers (30 participants), TheSans (26.7%) was again ahead, with TheMix in second place (23.3%). Thus, the good and poor readers exhibited similar behavior in terms of the fonts they read the fastest, but TheSans's dominance among poor readers is even more evident. Of course, these distributions must be interpreted with great caution because the number of cases for both good

40 Only 35 percent of the readers who classified themselves as reading "very well" on the questionnaire fall into the same category based on their actual reading speed, while 54.7 percent of the "moderately well" readers do. On the questionnaire, only 5 people classified themselves as "not so well," but 33 people had a "poor" reading speed in the tests.

41 The reading speed for each reading level is 0 to 399 characters per minute for poor readers, 400 to 800 characters per minute for moderate readers, and 801 characters or more per minute for good readers. We calculated the average reading speed for all five texts and categorized the participants into one of the three groups based on their results: 24.1 percent were classified as poor readers, 54.5 percent as moderate readers, and 21.4 percent as good readers.

		Fastest Font					
		Arial	TheSans	TheSerif	TheAntiquaB	TheMix	total (N)
Reading Level (Character/Min)	poor	6	13	5	4	7	35
		17.1%	37.1%	14.3%	11.4%	20.0%	
	moderate	21	18	13	16	10	78
		26.9%	23.1%	16.7%	20.5%	12.8%	
	good	4	8	5	6	7	30
		13.3%	26.7%	16.7%	20.0%	23.3%	
	total	31	39	23	26	24	143
		21.7%	27.3%	16.1%	18.2%	16.8%	

Figure 5
Cross-tabulation reading level and reading speed of the five fonts (characters per minute).

and poor readers is very low (see Figure 5). Looking at the number of characters read by the poor reading group (0 to 399 characters per minute), we find that these 35 readers often had read only three lines (or a maximum of nine lines or one paragraph). With so few lines, one of the advantages of serif fonts—their good guidance of the serifs across the line of text—is not so key a consideration.

Do Visual Impairments Affect the Results?

Arial and TheSans (27.8% each) were the fastest-read fonts for participants with visual impairments; TheSans and TheAntiqua (22.6% each) were the fastest-read fonts for those without visual impairments. However, we made a mistake here and formulated our questions incorrectly. We asked: "Does the test person have a visual impairment?" "Does she use a reading aid?" In doing so, we did not consider that a visual aid (e.g., glasses) normally corrects the ametropia and does not provide any information about who was actually affected by impaired vision when the reading test was carried out. We should have asked: "Do you have a visual impairment despite corrective lenses? If yes, what is it?" Therefore, we cannot make any statement about the influence of visual impairments.

Did the Order of the Texts Play a Role in Reading Time?

Because our test group has reading difficulties, we assumed that fatigue would set in over the course of the test and that the texts would be read more slowly toward the end. For this reason, we varied the order of the fonts so that, for example, Arial would be read as the first text for some readers and as the second, third, fourth, or fifth text for other readers. The statistical evaluation showed exactly the opposite of our assumption. The longer the test lasted, the faster the readers read across all fonts (see Figure 6).

Characters per Minute	Text 1	Text 2	Text 3	Text 4	Text 5
Standard Deviation	288.456	229.238	228.928	215.459	235.758
Mean	557.68	595.99	610.20	613.61	627.34

Figure 6
Increased reading speed over the test duration (characters per minute).

Is "Ilona" Misread More Often in the Arial Font Than in Other Fonts?
We inserted proper names in the text to test readers' confusion with like-appearing letters. The letter combination *Il* in the name *Ilona* cannot be deciphered in the Arial font if the reader does not know the name because the letters hardly differ and cannot be clearly assigned to the sound (i.e., "i" or "el"). However, in the styles of the Thesis typeface family, the two letters do differ. The evaluation shows that the percentage of those who misread Ilona is approximately the same across all fonts—between 13.2% (TheAntiquaB) and 15.3% (Arial and TheSans).

We also discovered that the readers from Germany stumbled over the words *Ossiacher See* and often misread the location, while the test readers from Austria had no problem with it. The name of the lake is well-known in Austria but not in Germany.

These two observations indicate that readers of *easy-to-read language* use the lexical way of reading more than previously assumed.

Discussion of the Font Style Results
In summary, we can state that the font Thesis TheSans is read significantly faster than Arial and all other fonts examined in the study. Arial performed worst in the overall comparison, and fonts with serifs were read slightly faster than Arial. We did not get a clear answer to the question about which kind of serifs are good. All three fonts with serifs performed about equally well, as did the fonts with differing stroke contrasts—that is, neither TheSerif and TheMix on the one hand (these fonts have almost the same stroke thickness) nor TheAntiquaB on the other (this font has the typical modulated stroke thicknesses of an old style font) seemed to have had any influence.

Therefore, the test group behaved similarly to experienced readers: They read a (sans serif) font whose design follows the dynamic formal principle (TheSans) faster than a font with a static formal principle (Arial). Even serif fonts with dynamic (humanist) formal principles are read faster than Arial. We can therefore assume that the general principles for reading texts also apply to readers of *easy-to-read language*: Fonts with dynamic formal principles (where the reference to writing with the broad-nibbed pen is still present, giving the letters a diagonal stress and a forward flow that facilitates reading[42]) are more legible—whether they have serifs or not. The shapes of the letters are decisive: They must be recognizable, distinguishable, and open, and they must have good contrast between stroke thicknesses.[43] The Thesis family follows all these criteria.

42 Luc(as) De Groot https://lucasfonts.com/fonts/the-sans/info (accessed December 13, 2022).

43 DIN, Schriften—Leserlichkeit [Lettering—legibility],1450:2013-04, 21.

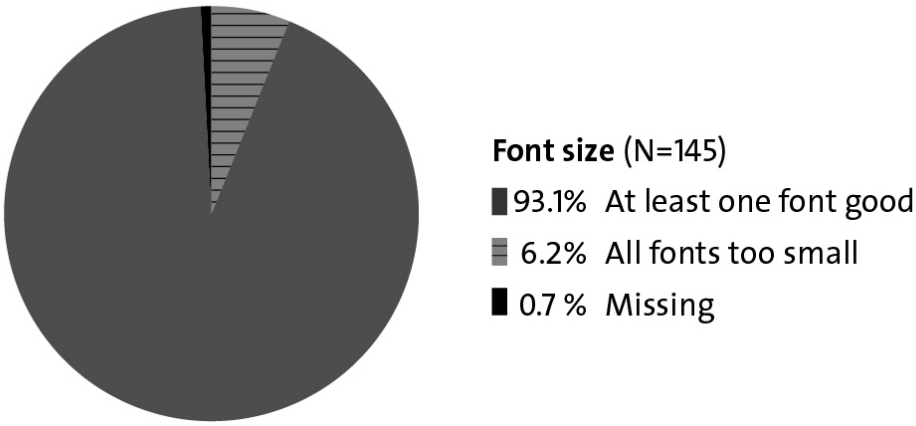

Figure 7
Participants' ratings of font size (all fonts displayed in equal size, 2.25 mm x-height).

This result is surprising in its clarity because, in legibility research, the assumption is that what is read most is read most quickly.[44] Because the test persons have primarily read Arial until now, their speed in reading it should actually be superior to their reading of other fonts. The comments from the test subjects also show that serif fonts in particular are unfamiliar, but that they do not affect reading performance.

We also were surprised that the participants read the texts faster the longer the test took. We assumed that the effort of reading would cause readers to slow down. We offer two possible explanations here. First, the content of the text was exciting, and the participants enjoyed taking part in the study. This perspective could lead to participants' happily reading "further." However, readers with learning difficulties are probably underestimated in terms of the amount they read. If the content of a text is exciting and the micro- and macro-typographical settings are optimal, five pages of text can be mastered without any problems. Second, test administrators reported that some participants diagnosed with mental illness were very nervous and afraid of doing something wrong. This nervousness decreased throughout the test and could explain why participants read texts faster the longer the readers read them. In terms of text volume and reading time, readers of *easy-to-read language* reacted like all readers: "Reading motivation, regardless of the target group, is primarily determined by the ergonomically correct treatment of their bodies."[45] In the test material, the visual apparatus was served ergonomically by appropriate typography.

Results of the Font Size

The following results concern the second specification: The font size.

Which Font Size Is Ideal for a Text Meant for Immersive Reading?

The evaluation showed that only nine people (6.3%) found all of the fonts too small (see Figure 7). Another 36 people (25.5%) said that one of the five fonts was too small but that at least one font's size was "good." One person found all fonts too large.

44 See Wendt, "Lesbarkeit von Druckschriften," 45; and Zuzana Licko, Typeface Design: An Interview with Zuzana Licko," in *Emigre No. 70* (2009): 13.

45 Tiefenthaler, "Zuerst liesst der Körper" [The body reads first], 34.

Font size	Frequency	Percent
One font good + one font too small	36	24.8%
Others	105	72.4%
Missing	4	2.8%
Total	145	100%

Rating/Speed	Frequency	Percent
Font rated good + read faster	9	25%
Font rated too small + read faster	27	75%
Total	36	100%

Figure 8
Discrepancy between reading speed and subjective evaluation of font size.

The evaluation of the questionnaires showed that, of the nine people who found all the fonts too small, two had an appointment with an ophthalmologist (and therefore were probably wearing glasses that were incorrectly adjusted). Of the 36 people who found at least one of the fonts too small, but also found the size of at least one font to be good, 27 participants read the font rated too small faster than a font rated well. In fact, 12 people even read the font rated "too small" the fastest (see Figure 8). These cases are most common with the Arial and TheMix fonts. In many of these cases, the reading speeds are close together and the sample size of 36 people is small, so that results should be interpreted with caution.[46] The cross-tabulation with the reading level shows that the poorer readers more often rated all fonts as being "too small." However, this difference is not equally pronounced in all fonts. Poor readers tended to rate TheAntiqua, Arial, and TheSans as being too small.[47] The influence of visual impairments on the assessment of the font size was surprisingly small: 82 people (56.6%) stated that they had a visual impairment, but most had no problems with the font size.

Discussion of the Font Size

The results for font size come from the individual ratings on the fonts' sizes made by the readers. The vast majority of test participants (93.7%) rated at least one font with an x-height of 2.25 mm (equivalent to 12-pt Arial and 12.5-pt Thesis family) as being large enough. The Arial font was rated as being too small more often than any other font. In terms of size, serif fonts received good ratings more often (if the actual size is the same), so they appear to be larger and more legible. Also interesting is that a font rated too small was often read faster than a font whose size was rated good. Visual impairments also seem to play a lesser role than we had thought. Poor eyesight is probably compensated for by glasses or contact lenses, just as it normally is for the population as a whole.

The recommendation of a minimum font size of 14 pts may need to be revised. An overwhelming proportion of study participants found the smaller font size to be a good thing. For a print product in A5 format, we therefore recommend an x-height of 2.25 mm and a corresponding font size of 12 pts to 12.5 pts.

46 Which fonts were rated as being too small? Assessing the type size by font shows that TheSerif is the least likely to be rated "too small" (13.3%), while Arial is the most likely (20.0%). TheSerif also has the largest share of the "good" rating (86.0%), followed by TheSans (83.2%).

47 For Arial, the "too small" rating ranges from 29.4% (poor readers) to 6.7% (good readers). The chi-square (chi^2) test for this table is not significant. For TheSans, the "too small" rating ranges from 27.3% (poor readers) to 3.3% (good readers). The chi^2 test for this table is significant (chi^2=6.845, df=2, p=0.033). The contingency coefficient indicates a weak association (c=0.215). For TheSerif, the "too small" rating ranges from 23.5% (poor readers) to 3.2% (good readers). The chi^2 test for this table is not significant. For AntiquaB, the "too small" rating ranges from 33.3% (poor readers) to 3.3% (good readers). The requirements for the chi^2 test were not met (too low cell occupancy). For TheMix, the "too small" rating ranges from 23.5% (poor readers) to 9.7% (good readers). The chi^2 test for this table is not significant.

However, legible font sizes depend on various factors. In addition to the font selection, the reading distance and the medium are decisive. The font size calculator of the German Federation of the Blind and Partially Sighted (DBSV) can be of assistance here.[48]

Note also that a smaller font size can reduce the "perceived" amount of text on a page. Because of their many explanations, texts in *easy-to-read language* tend to be too long and thus have a deterrent effect on the readership. When the smaller font size allows for a more airy layout and more white space, the reader then has the feeling of having to cope with a smaller amount of text. To the best of our knowledge, no study has yet been conducted on this effect, but we note an observation on font size from practice: "While they can read the same range of sizes as anyone else, they welcome larger type. They seem to find it less intimidating, perhaps because there is room for less content on each page."[49]

Participants' Subjective Evaluations of the Fonts

Evaluation of the questionnaires gives insight into the participants' subjective evaluation of the font styles.[50] Almost all participants (73%) gave different ratings to the six fonts. The readers noticed clear differences in the fonts and in some cases precisely described their impressions. The additional comments often are surprisingly detailed. For instance, on Text 3 (TheSans), test subject 20 commented: "Very good. Perfect, very comfortable to read! Helped me read fluently!" To Text 4 (TheSerif), this participant commented: "Very good. The font style is even somewhat better than Text 3"; then, about Text 5 (TheAntiqua), the comment read: "Very good. The spacing between letters is even better here!"

Test subject 145 even knew the technical term of line spacing and rated Text 2 (Arial) this way: "Ok, line spacing is too large, does not appear uniform."

Test subject 23c had a clear favorite in Text 1 (TheSerif), commenting that "from its appearance, I like it the best." The subject also negatively associated the Arial font with automated creation (Text 4), commenting "that's nothing—written by a computer." This expresses the fact that the test subject correlates the Arial font with a standardized, impersonal design. Test subject 72b also did not like the Arial font at all, commenting that "...the font stands out in a negative way."

Interestingly, the subjective favorites often do not match the (objective) reading speed. Test subject 20 read TheSans the most quickly and found TheAntiqua to be the best, while test subject 23c read TheAntiquaB significantly faster than the preferred TheSerif, however, test subject 72b read the negatively rated Arial font the slowest. The statistical evaluation shows that fonts that are disliked tend to be read more slowly. However, this effect is not constant.

48 http://www.leserlich.info/tools/schrift-groessenrechner/index.php (accessed December 13, 2022).

49 Robert Waller, "Choosing a Typeface for Reading," Technical Paper 9, (Reading, UK: Simplification Centre, University of Reading, 2011), 3.

50 We asked: How do you like the font? Participants could select one of five options: very good, good, ok, bad, not at all). Individual statements could also be added: "Note down your observations and the test subject's comments on the font."

From the comments, we can see that the serif font is unfamiliar. Nevertheless, it is read faster than the usual Arial font by test subject 81a, who commented "(ok, font is unfamiliar, [I] have no experience with the font)."

Discussion of the Subjective Comments and Font Styles' Connotations

Many people assume that the choice of font does not play a major role beyond the legibility of a text. However, we know from the corporate design that fonts have an appealing quality that contributes to a text's meaning. In contrast, some *easy-to-read language* policies assume that readers of *easy-to-read language* do not perceive this aspect of fonts (i.e., the affective domain[51]) and even propose a "typographical base format." They claim that the target group lacks the "ability to perceive and evaluate the historically differentiated functions of macro- and microstructural typographic properties of texts and the linguistic variance associated with them."[52]

Our results show exactly the opposite: Readers certainly did recognize the different font styles and assigned meaning to them. Their observations were so detailed that it can be assumed that they are particularly close observers of font differences. The practical, participatory work also has shown that these readers compensate for a lack of reading skills through close observation.[53] For practical work with *easy-to-read language*, these findings mean that fonts that are part of a corporate design should be in *easy-to-read language* too—providing that they comply with legibility principles.

Summary and Outlook

Our study shows that people read *easy-to-read language* faster if legible fonts are used and that, in most cases, a font size of approximately 12 pts (x-height 2.25 mm) is considered large enough. The visual apparatus for readers of *easy-to-read language* seems to function in the same way when decoding symbols as it does for non-impaired readers. In addition, many participants perceived the different fonts and applied specific connotations and assessments to them.

This first scientific evaluation of the current German *easy-to-read language* rules regarding font usage shows that the propagated special solution—sans serif fonts, such as Arial, in size 14 pt—is simply wrong. The assumptions underlying this standard had no scientific verification, they neglect findings from design research, and the standards need to be revised.

In practice, legible fonts should be used. The legibility is determined by the form principle of the individual letters. Fonts that follow the dynamic font principle (the Thesis superfamily and in particular the font TheSans) are read faster than a font with a static

51 Waller, *The Typographic Contribution to Language*, 28–45.

52 Ursula Bredel and Christiane Maaß, eds., *Leichte Sprache. Theoretische Grundlagen. Orientierung für die* Praxis [Easy-to-read language. Theoretical foundations. Orientation for practice] (Berlin: Dudenverlag, 2016), 266.

53 Sieghart, "Leichte Sprache, Typografie & Angemessenheit" [Easy-to-read Language, Typography & Appropriateness], 130.

formal principle (Arial). In addition, professional micro- and macrotypography should be applied to optimally support the complex process of reading. There is no such thing as one correct font and font size—but only optimal specifications for the respective application. The function of the text, the content, the reading situation, and the medium also influence the most appropriate font selection. It should also be determined whether it is helpful for the understanding of a message if the sender can be recognized by its corporate typeface. We suggest that determining the requirements of the target audience is part of the participatory process as is taking reading levels and motor or visual impairments into account.

This study looked only at the reading of continuous text in print material. A signage system requires different fonts and scientific testing in a different setup. Another limiting issue is the heterogeneous group of *easy-to-read language* readers and the additional impairments they might have. The visual impairments of our test group were mostly corrected by glasses. However, eye diseases like cataracts, glaucoma, and macular degeneration might require different measures in terms of font selection and type size. One desideratum is to examine the needs of this specific target audience with further studies and larger test groups. Regarding legibility studies, a valid scientific setup, as we discussed in our review of theoretical foundations, is indispensable. The eight factors listed offer a guide for further studies. We also hope to provide scientific background knowledge of typography for other disciplines.

Easy-to-read language is a relatively young tool for barrier-free communication. This study aims to contribute to making texts easier to read for the target group by providing guidance on how to optimize the first step in the reading process, which is to absorb and decipher the letters that we are seeing.

Embodied, Everyday Systemic Design – A Pragmatist Perspective

Josina Vink

Introduction

There is a prominent narrative in design research that as design has evolved from craft into a sophisticated professional practice, it has taken on increasing complexity; based on this narrative, as design continues to evolve and expand toward tackling social systems transformation, there is a need for the integration of systems thinking to aid designers in dealing with the complexity that is otherwise beyond their cognitive limit.[1] This colonial narrative of design's progress emphasizes the centrality of thinking over doing and suggests that hope for social change rests on an extension of the way of reasoning of professional designers.

In this context, systemic design is an emerging design domain that has grown in popularity in recent years. It combines design approaches with systems thinking to address complex societal issues, such as housing, foreign policy, immigration, and environmental issues.[2] Increasingly, systemic design is seen as a professional practice done by systemic designers with their own distinct methods and tools for grappling with complexity.[3] Engagement of diverse stakeholder groups is at the core of systemic design practice and its methods.[4] However, the contemporary emphasis in systemic design remains on what Ezio Manzini designates as "expert design," where the systems change process is led by systemic designers with advanced design skills.[5]

As highlighted by Anne-Marie Willis, "When a question of a philosophical character is posed, such as 'what is design?' the answer is mostly already over-determined by the model of professional design as the model of all designing."[6] By upholding systemic design as primarily a professional practice, the field is positioned in service to capitalism and negates the broad array of intentional shaping that already happens amid social systems. This elitist focus contributes to an active defuturing by design through the erosion of *plurality*—a principle to which the field of systemic design ironically aspires.[7] In recognizing the need for attention to what they call "design by society," rather than more proximate design by expert

1 Kees Dorst, "Design Beyond Design," *She Ji: The Journal of Design, Economics, and Innovation* 5 no. 2 (2019): 117–27, https://doi.org/10.1016/j.sheji.2019.05.001.

2 Birger Sevaldson and Peter Jones, "An Interdiscipline Emerges: Pathways to Systemic Design," *She Ji: The Journal of Design, Economics, and Innovation* 5 no. 2 (2019): 75–84, https://doi.org/10.1016/j.sheji.2019.05.002.

3 Alex Ryan, "A Framework for Systemic Design," *FORMakademisk* 7, no. 4 (2014): 1–14, https://doi.org/10.7577/formakademisk.787; and Mieke Van der Bijl-Brouwer and Bridget Malcolm, "Systemic Design Principles in Social Innovation: A Study of Expert Practices and Design Rationales," *She Ji: The Journal of Design, Economics, and Innovation* 6 no. 3 (2020): 386–407, https://doi.org/10.1016/j.sheji.2020.06.001.

4 Peter Jones and Kristel Van Eal, *Design Journeys Through Complex Systems: Practice Tools for Systemic Design* (Amsterdam, The Netherlands: BIS Publishers, 2021), 18.

5 Ezio Manzini, *Design, When Everybody Designs: Introduction to Design for Social Innovation* (Cambridge, MA: MIT Press, 2015), 37–38.

6 Anne-Marie Willis, "Introduction," in *The Design Philosophy Reader*, ed. Anne-Marie Willis (London: Bloomsbury, 2019): 1–8.

7 On defuturing, see Tony Fry, *A New Design Philosophy: An Introduction to Defuturing* (Sydney, Australia: University of New South Wales Press, 1999). On plurality in systemic design, see Birger Sevaldson and Peter Jones, "An Interdiscipline Emerges: Pathways to Systemic Design," *She Ji: The Journal of Design Economics and Innovation* 5, no. 2 (October 2020): 75–84, DOI:10.1016/j.sheji.2019.05.002.

https://doi.org/10.1162/desi_a_00731

designers, Edward Woodhouse and Jason Patton call for greater care to the broader processes of design, in which myriad persons participate, and the embedded nature of this process within society.[8]

In response, I argue that pragmatist philosophy, which has been foundational to the development of the design discipline (although rarely acknowledged as such[9]), can aid the evolving field of systemic design in more fully understanding the situated, embodied, and everyday nature of systemic design practices. By drawing on the work of classic pragmatists, particularly John Dewey and George Herbert Mead, as well as contemporary feminist pragmatists, including Shannon Sullivan and Patricia Hill Collins, I offer a grounded understanding of everyday systemic design, exemplified by stories of my own lived experience.

A pragmatist perspective on systemic design reveals that the practice of designing social systems is pervasive and mundane. It happens not *to* but *within* and *through* social systems in the everyday. Systemic designing involves an ongoing spiraling of evolving transactions within and between bodies and social systems that are themselves mutually constituted. Recognizing the pervasive nature of systemic design, I call for an emphasis on staying with situated, lived experiences and cultivating bodily habits of careful experimentation in the everyday. I argue that this reframing of systemic design demands a shift in the orientation of the field from crafting expert-driven systems change to cultivating collective reflexivity, so that everyone might continue to shape their own worlds in respectful relation with others.

Need for an Alternative Perspective on Systemic Design

The field of systemic design touts values of pluralism, yet the methods and tools it champions often have a paradoxical relationship with this ambition.[10] These contradictions can be understood through a closer look at one of systemic design's most championed approaches, creating maps that visualize the complexity of systems, such as giga-maps or synthesis maps.[11] These maps offer representations of focal systems using visual language to understand relations across multiple levels and scales. These artifacts are meant to be interpretive and to incorporate different worldviews, but through visual representation, they end up normalizing certain truths and silencing others.[12] The aim is that these system maps are recognized as always incomplete. However, in their attempt to capture supercomplexity, these artifacts can have the effect of perpetuating ontological occupation, as the reality captured in the map contributes to the erasure of other local realities rendered invisible.[13]

8 Edward Woodhouse and Jason W. Patton, "Design by Society: Science and Technology Studies and the Shaping of Design," *Design Issues* 20 no. 3 (Summer 2004): 1–12.

9 Brian Dixon, "Experiments in Experience: Towards an Alignment of Research Through Design and John Dewey's Pragmatism," *Design Issues* 35, no. 2 (Spring 2019): 5–16.

10 Sevaldson and Peter, "An Interdiscipline Emerges," 75–84.

11 See, e.g., Birger Sevaldson, "GIGA-Mapping: Visualization for Complexity and Systems Thinking in Design," in *NORDES Conference Proceedings*, no. 4 (2011): 1–20; and Jeremy Bowes and Peter Jones, "Synthesis Maps: Systemic Design Pedagogy, Narrative and Intervention," in *Relating Systems Thinking and Design RSD5 Symposium Proceedings*, ed. Peter Jones (Toronto: Systemic Design Research Network, 2016): 1–14.

12 Jocelyn Bailey and Lorainne Gamman, "The Power in Maps: Reviewing a 'Youth Violence' Systems Map as Discursive Intervention," in *DRS2022 Conference Proceedings*, ed. Dan Lockton et al., (Bilbao, Spain: DRS Digital Library, 2022), 1–19, https://doi.org/10.21606/drs.2022.563.

13 Josina Vink et al., "Ontological Occupation When Designing for Scale in Healthcare," in *NORDES 2021 Conference Proceedings* (Kolding, Denmark: Nordic Design Research, 2021) 292–301, https:// doi.org/10.21606/nordes.2021.32.

This flat, static "view from above, from nowhere" of systems positions the systemic designer as an all-knowing manipulator, ready to identify the relevant leverage points for change.[14] This view of systemic design is particularly convenient for the commercial design industry as it perpetuates the capitalist business model through which the systemic designer can be billed out in relatively short-term fundable projects. At the same time, the overemphasis on these professional practices of systemic design, using specialized representational frameworks and tools, undermines the embodied, situated understanding and intentional shaping of social systems. The design of social systems is understood more broadly as the conscious creation of possibility in the human worlds we inhabit.[15] However, everyday design by the community itself, which is essential for maintaining plurality in the ongoing negotiation of systems futures, has been largely ignored in contemporary systemic design discourse.[16] Here, it is important to distinguish between the autonomous design of communities—in which the conditions exist for communities to continually change their own norms from within[17]—and participatory design or co-design—in which members of the community are invited into practices of professional systemic designers through episodic projects.[18]

To build this extended understanding of designing in and by communities, pragmatist philosophy offers a grounded, situated, and pluralistic understanding of how social systems are being intentionally shaped by humans. The work of the classical American pragmatist philosophers, taking place roughly between 1870 and 1945, was an intentional departure from dominant philosophical discussions at the time.[19] In this period, the prominent European mode of thinking emphasized the importance of objectivity, upholding the ideal of a singular notion of the truth, and maintained the dualism of mind and body; in contrast, classical American pragmatists were committed to pluralism and relationality in their view of reality, likely at least in part a result of local Indigenous thought that was influencing European philosophy during that time.[20]

Offering a frame for situated knowledge and action, pragmatism highlights lived, bodily experiences as a critical starting place. It recognizes that experiences are an emergent property of humans' active engagement with their social worlds.[21] In doing so, pragmatism brings forward the situated, action-oriented nature of creativity whereby humans intentionally influence their environment through bodily action, while the environment shapes their human subjectivity. This situated, action-oriented view of human creativity amid social systems offers a hopeful means for building a more

14 Donna Haraway, "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective," *Feminist Studies* 14 no. 3 (1988): 575–99.

15 Gary S. Metcalf, *Social Systems Design* (London: Springer, 2014), vii.

16 Arturo Escobar, *Designs for the Pluriverse: Radical Interdependence, Autonomy, and the Making of Different Worlds* (London: Duke University Press, 2018), 166–89.

17 Ibid., 172.

18 Josina Vink et al., "Service Ecosystem Design: Propositions, Process Model, and Future Research Agenda," *Journal of Service Research* 24 no. 2 (2021): 168–86, https://doi.org/10.1177/1094670520952537.

19 John J. Stuhr, *Pragmatism and Classical American Philosophy: Essential Reading and Interpretive Essays* (Oxford: Oxford University Press, 2000), 1–9.

20 Scott L. Pratt, *Native Pragmatism: Rethinking the Roots of American Philosophy* (Indianapolis, IN: Indiana University Press, 2002), xi–xviii.

21 Patricia Hill Collins, "Piecing Together a Genealogical Puzzle: Intersectionality and American Pragmatism," *European Journal of Pragmatism and American Philosophy* 1, sec. 2 (2011): 88–112, https://doi.org/10.4000/ejpap.823.

grounded, pluralistic understanding of systemic design. I argue that a closer look at the writings of classical pragmatists, especially Dewey and Mead, offers valuable insight into the everyday practice of systemic design and that contemporary feminist pragmatists, like Sullivan and Hill Collins, aid in nuancing that understanding and its implications for systemic design.

Transactional Bodies Stewing in Social Systems

For Dewey, human existence is a bodily existence. He refers to bodies as organisms, reflecting the dynamic, living nature of our physicality and a continuity between humans and the more-than-human world.[22] Although the physicality of the body is important for Dewey, he is most interested in bodies as acts of performance or "bodying"; such movements, he emphasizes, are based on predispositions of the corporeal body, or what he refers to as habits.[23] For Dewey, humans' corporeal existence is transactional, meaning that bodies have a co-constitutive relationship with their environment.[24] *Transactional* here can be understood in contrast to *interaction*, which signifies two independently constituted entities engaging in an exchange. The concept of transaction rejects the dualism between the self and the social worlds we inhabit and suggests that they are mutually constituted in a dynamic, evolving relationship.

Dewey reminds us that the skin is an arbitrary boundary of where the body ends and the environment begins. "No creature lives merely under its skin; its subcutaneous organs are means of connection with what lies beyond its bodily frame, and to which, in order to live, it must adjust itself, by accommodation and defense but also by conquest."[25] Dewey acknowledges the inherent interdependence, continuity, and interchanges between living creatures and their surroundings. To aid in this understanding, one can think about how the human organism must ingest other organisms and turn them into parts of itself and, in doing so, can also play a role in cultivating other organisms; if we simply think about our relationship with food, we can understand that the outside world is anything but outside of us.[26] If we reflect more particularly on social systems, a similar relationship can be understood: Human bodies are composed of social systems through their repeated bodily activities, and these habits in turn shape the systems in which humans are embedded.

Shannon Sullivan presents the metaphor of the stew as a way to understand how Dewey's concept of transaction embraces both continuity and difference.[27] The stew is not a fondue, where all ingredients completely melt into each other, nor is it a tossed salad,

22 John Dewey, *Art as Experience* (1934; New York: Perigee Books, 1980), 13–15.

23 Unpacking Dewey's thinking on this, see Shannon Sullivan, *Living Across and Through Skins: Transactional Bodies, Pragmatism, and Feminism* (Indianapolis, IN: Indiana University Press, 2001), 1–40.

24 John Dewey and Arthur F. Bentley, *Knowing and the Known* (Boston: The Beacon Press 1949), 108.

25 Dewey, *Art as Experience*, 13.

26 Lisa Heldke, "It's Chomping All the Way Down: Toward an Ontology of the Human Individual," *The Monist* 101, no. 3 (July 2018): 247–60, https://doi.org/10.1093/monist/ony004.

27 Sullivan, *Living Across and Through Skins*, 1–40.

where different ingredients remain separate together. Rather, says Sullivan, the ingredients stew together and intermingle so that they constitute each other. For example, an onion in the stew becomes a carrot-y onion and a carrot in the stew becomes an onion-y carrot. These ingredients remain distinct from each other, but they permeate and constitute each other in an ongoing back and forth. Humans are the habits that they embody, and these habits become part of the corporeal body and the social systems.

Mead elaborates further on the social and structured nature of the self. For Mead, bodily habits make up the social structures of our systems, and these structures are what inhabit our bodies through our bodily patterns.[28] For Mead, the very idea of the self is itself a social structure that arises through social experience in society.[29] This conceptualization of the embodied, social self locates the human unequivocally in community and highlights that human embodiment is a product of its relations with the social world. In this way, human bodies are open systems that are co-constituted by social systems. Human embodiment, in its repeated enactments, provides the very structure of social systems. Social systems and bodies are entangled and inseparable. However intertwined, distinctions remain between and within bodies in these systems. Aligned with a pragmatist perspective, I offer an example from my own lived experience of the transactional nature of bodies in social systems, entangled and mutually constituted in and through the habits that make up social structures:

> After gym class, I follow to the right exit of the gym and push open the door with the sign "GIRLS' CHANGING ROOM" as I move in unison with the chattery procession. This collective habit of undressing and redressing together with almost half the class validates my gender—no one questions it, not even myself at the time. The boys' changing room was the only other option, and that felt to me totally out of the question. Together, we peel off our baggy t-shirts and shorts, sweaty from the circuits we had completed. That attire is appropriate for gym class, but showing our curves was needed to move through the halls of the high school and not be met with adverted eyes.
>
> Julie gracefully pulls off her shirt, clasps her wire bra on the outside of her sports bra before pulling the sports bra off entirely, carefully replacing fabrics so as not to reveal her nipples. She layers on a striped Roxy tank top. Then after removing her shorts, tugs back on new dark Parasuco jeans, before proceeding to the mirror to fix her hair and make-up.

28 George Herbert Mead, *Mind, Self and Society: From the Standpoint of a Social Behaviourist* (1934; Chicago, IL: The University of Chicago Press, 1972), xxiv–xxv. See also Barbara Simpson, "Pragmatism, Mead and the Practice Turn," *Organizational Studies* 30 no. 12, (2009): 3329–46.

29 Stuhr, *Pragmatism and Classical American Philosophy*, 1–9; and Mead, *Mind, Self and Society*, 140.

More timidly, I turn to face the white wall of the changeroom, scan to make sure no one is looking my way, and frantically, but not too frantically to draw attention, switch my gym shirt for my tank top and Cleveland Cavaliers jersey. I don't change my sports bra because that would mean exposing my not-quite-as-expected body to the scrutinizing looks and whispers of my friends and classmates.

With my street clothes on top, I carefully rearrange the prosthetic breast, which hides in my sports bra, into the proper place for mimicking the female silhouette. I had needed to have something inside the right half of my bra since my left breast started developing, and not the right, in grade six. My mom quickly found me a shoulder pad from one of her old blazers in those early days, until a silicone prosthesis could be obtained from the local mastectomy wear shop. It was never a question that chest symmetry was required. It was such an expectation that I was encouraged to visit the plastic surgeon who recommended that a back muscle be pulled to my front so that I could get a breast implant, which would "make my life so much easier." I halted that operation, unwilling to give up the full rotation of my right arm and jeopardize my baseball throw, just so that my Poland-syndrome-body could conform to images of women I had seen in the magazines and all around me in rural Ontario.

Still, I knew I needed to uphold the illusion of symmetry, so I stowed on the prosthetic breast every morning and kept it hidden under my clothing, even when changing. This choice helped me to walk the halls and perform the habits of little women with confidence, even if I was a bit hesitant. After putting on my own too-tight jeans, I proceeded to the mirror behind Julie. As a sporty girl, I could get away with a jersey and jeans, but that would need to be offset with at least some mascara to look a bit more feminine. Still, my eyes usually ended up with dark smudges at the end of the day because I forgot not to rub them. Just then, I carefully looked ahead in the mirror and moved the applicator from the base of my lashes to the tip, like I saw the other girls do.

This story contextualizes how bodies inhabit social structures and, in doing so, uphold the social structures that give shape to our communities and social systems. Through the routine bodily activities of changing in the girls' changing room, wearing a prosthetic breast daily, diligently putting on mascara, and through the reactions I got from others, I performed what it meant in my

Figure 1
Transactional bodies stewing in social systems. Illustration by Tianqi Li.

community to "be a girl." My identity as a sporty tomboy was not a precursor to my habits but was cultivated in and through these repeated bodily activities. These activities were informed by the regular corporeal performances of others in my social context. At the same time, our habits constituted the gender binary that continued to play out in the local community of my small, rural high school, including through the division of sports teams, in gendered interactions in the hallways, and perhaps most acutely for me, at afterschool school dances held in that same gym. My body is a transactional body constituted by my relations with others and constituting the very structures of my community. Figure 1 shows the stewing of transactional bodies in their social environment, illustrated here in the mundane act of dressing within a high school girls' changing room.

Mutual Transformation of Bodies and Social Systems

For pragmatists, these embodied transactions with one's social context, or "situation" as Dewey often says, provide the opportunity for (mutual) transformation. Dewey makes clear that the human organism and social systems are continually being remade

and reconstituted through ongoing changes in relationships.[30] Where there is friction amid these relations, the opportunity for transformation becomes possible, according to Dewey. When humans meet an aesthetic experience of an unsettling situation, where bodily predispositions do not align with the environment, the situation sparks their doubt and catalyzes the process of inquiry, enabling the possibility that a habit connected with an entrenched social structure can be destabilized.[31] When disruption is experienced, the process of inquiry is initiated, and an existing habit may be deemed inadequate or inappropriate.

Mead elaborates on this consciousness or awareness of social structures that one inhabits, calling it reflexiveness or reflexivity.[32] Through such reflexivity, human organisms can intentionally and creatively adjust themselves through a social act. For Mead, reflexivity is the foundation for the intentional shaping of social systems and a fundamental requirement for social progress.[33] From this perspective, social systems design happens through bodies in reflexive transactions within community. The plasticity of the self can be seen as susceptible to and influenced by change in the environment; however, this plasticity is also the very source of power for intentionally transforming social systems.[34] People's habits can generate friction when they relate to other bodies, and in doing so, they can challenge or transform each other when conflict or friction emerges. Mead reminds us that conflicts are not simply between human individuals but between social structures—sometimes even multiple social structures that one individual human inhabits.[35] He reinforces that these conflicting social structures often necessitate reconstruction of our social systems and that changes to social structures inevitably require changes in the humans that inhabit them.[36]

Amid transactions in a social context, the human organism is not caught in a recursive loop of being influenced by social structures and simply reproducing them. Rather, the relationship is more like that of a spiral, whereby enacting some social structures generates some degree of consciousness, enabling the intentional embodiment of particular social acts that can alter the environment.[37] Human organisms are always in transaction with their social context and, as such, have a direct influence on the very social structures by which they are influenced. However, Sullivan explains that a disruption in one human's habit often does not carry enough force to counter the inertia of institutionalized social structures. "For changes in an individual to result also in changes to an institution, an individual needs the efforts of other individuals attempting to make the same sorts of changes."[38]

30 Dewey, *Art as Experience*, 14.

31 Katarina Wetter-Edman, Josina Vink, and Johan Blomkvist, "Staging Aesthetic Disruption Through Design Methods for Service Innovation," *Design Studies* 55 (March 2018): 5–26, https://doi.org/10.1016/j.destud.2017.11.007.

32 Mead, *Mind, Self and Society*, 134.

33 Ibid., 134.

34 Sullivan, *Living Across and Through Skins*, 94.

35 Mead, *Mind, Self and Society*, 307.

36 Ibid., 309.

37 John Dewey, *Human Nature and Conduct: An Introduction to Social Psychology*, The Modern Library ed. (1922; New York, NY: Random House, 1930), 328.

38 Sullivan, *Living Across and Through Skins,* 109.

Feminist pragmatists are interested in how pragmatism can become a critical social theory and methodology for approaching the social world.[39] Identifying the social movements led by Black women in the United States as visionary pragmatism, Patricia Hill Collins highlights the value of developing a community of inquiry, always testing ideas in a crucible of experience situated in local social contexts: "Everyday life is something that is rooted, grounded, contingent, dynamic, and holistic. It is characterized by infinite opportunities to engage in critical analysis and take action."[40]

Hill Collins advocates for critical reflexivity that informs deliberate action in and by community. Feminist pragmatists see this work as a way of approaching practical structures with continued exploration and relational action; they emphasize the importance of living with, not simply visiting, the social negotiations of diverse communities.[41] In terms of always testing ideas in a crucible of experience, Dewey advocates for the habit of not forming fixed, individual habits but of developing flexible, sensitive habits that stay open for intelligent reconfiguration and mutual transformation in the midst of collectives.[42] The following story from my own lived experience is a window into this process of mutual transformation:

More and more commonly, people began to introduce themselves with their pronouns, along with their names. I dreaded every time such an introduction was about to occur because I could not relate to the pronouns that those around me assumed I might use. My body shuddered when I was in a group that someone called "ladies," and I felt an increasing rub when people referred to me as "her." After being introduced to my roommate's partner, whose preferred pronouns are they and them, and awkwardly stuttering over their pronoun formulation one night in our shared kitchen, I started reading everything I could about gender queerness, transness, and non-binary identities, supported by some life-changing recommendations from my housemate.

I slowly began to recognize that the gender I had been wearing like an itchy sweater, and had taken for granted most of my life, was a social structure that I no longer wanted to enact. With the help of a few friends and my partner at the time, I started testing the use of they/them pronouns. Before this shift, I had felt like I could not possibly inconvenience people by asking them to adjust how they referred to me, but when I heard my friends acknowledge me in this way, I felt so deeply seen and respected. I gained the confidence I had previously lacked to cut my hair off and start making tweaks to my wardrobe. As the months passed, I started to appreciate

39 Hill Collins, "Piecing Together a Genealogical Puzzle," 88–112.

40 Ibid., 108.

41 Danielle Lake, "Pragmatist Feminism as Philosophic Activism: The {R}evolution of Grace Lee Boggs," *The Pluralist* 15 no.1 (Spring 2020): 25–45.

42 Dewey, *Human Nature and Conduct*, 70–74.

the intensely relational nature of gender. To feel more at home in my body, I needed to do something to acknowledge my nonbinary identity among those with whom I interacted every day.

After a tumultuous time of summer reflection, the new school year was upon me, and a fresh batch of master's degree students whom I would soon teach would be in the large auditorium for their orientation. I could not bear the thought of starting the school year without clarifying my preferred pronouns for my colleagues and students. The day before the students were to arrive, I hurriedly crafted an email to my colleagues with cute, comic-like illustrations explaining the gender landscape in which I was roaming around and my pronoun preferences under the subject line "Gender Reveal Party." After a friend's helpful review of the email, I hesitantly hit send.

I received an overwhelming number of supportive replies from my colleagues, but many also expressed worries about messing up my pronouns in their messages. I had some awkward hallway encounters after the email but was also reassured by one colleague who left a treat on my desk with a note that read "no party without cake." Amid this mix of emotions, the time came for the annual staff introductions in the auditorium. I was thankful that the colleague before me introduced herself with her pronouns—even though she was the only one to do so. It made the threshold that felt almost insurmountable just a little bit easier. Next it was my turn. I tried to breathe deeply. After saying my name, I blushed and announced: "I know this might be new for some of you, but I prefer if you would refer to me using they/them pronouns." I tried to be my regular confident self, but my voice cracked as I spoke. The crowd of faces looked at me with friendly but puzzled gazes.

A student came up to me after the introductions to talk through what I meant, and several more asked me for a discussion about gender in the months that followed. I was still mostly referred to using she/her pronouns at school. I corrected people when I had the strength, but I often let it go. Some people did shift how they referred to me, and I know it helped others do the same. Two years later, more folks at the school are using they/them pronouns. I am starting to hear it and other gender-neutral terms more and more in meetings and emails—sometimes even in self-corrections in the moment. Together in our awkwardness, we are muddling through by reconfiguring our habits in our introductions and challenging the social structures associated with gender in our small school community in Oslo.

Figure 2
Mutual transformation of bodies and social systems. Illustration by Tianqi Li.

This story of transactional bodies changing in community offers a window into the spiral of entangled evolution in social systems. By living in community with others, I was confronted with an understanding of gender that I had not yet grappled with. Through the labor of others, I started to build an awareness of the social structures that I had previously taken for granted. In small, safer social settings, I began testing what doing things differently with others, like using different pronouns, felt like in my body and in interactions with others. Informed by my own experience, I then worked intentionally to shape how others referred to me and thought about my gender at my school. Although the system did not change immediately and acts sometimes looked more like reproduction than change, a slow evolution occurred as more people joined these shared social acts of using they/them pronouns and gender-neutral language. This work of shaping social structures was not just an intellectual one; it was an embodied activity that involved increased heart rates, quick breathing, blushing, and strengthening new muscles in mouths. It involved awkward moments of stumbling—for myself and others—and sometimes being painfully subsumed by the forces of the status quo. Yet, this personal narrative also is a messy story of mutual transformation that emerges in and through everyday, embodied systemic design. Figure 2 depicts the mutual transformation of bodies and social systems, exemplified through awkward, and sometimes even painful, processes of stumbling over and shifting the use of pronouns.

Situating Systemic Design in the Everyday

Drawing on the work of classic and feminist pragmatists, I have attempted to explain how social systems are designed by and in transactional bodies through everyday negotiations within communities. By bringing forward my own lived experience of enacting gender, I have shown the embedded nature of systemic design and the ways that systemic design can play out in ordinary, everyday habits, like dressing and introductions. A pragmatist perspective of systemic design highlights bodies as central sites of systemic design. It is particularly important then for systemic design to acknowledge that what is understood as the body is plural, locally situated, and full of incommensurability across cultures and communities.[43]

A pragmatist perspective suggests staying with the plurality of lived bodily experiences in systemic design practices, as opposed to reverting back to extracting and flattening complexity into static representations. The corporeality of human organisms accounts for and navigates situated complexity in ways that are largely underappreciated in more intellectual and rational approaches to systemic design. Greater attention to transactional bodies and lived experiences that cannot be easily translated is needed to embrace the transformative potential of systemic design practices while protecting plurality. Corporeal plasticity and situated creativity offer a hopeful means for engaging in the messy, negotiated, back-and-forth process of the intentional change of social systems by communities.

A pragmatist perspective on systemic design further highlights the importance of reflexivity—an awareness of the social structures that we inhabit—that is cultivated through the experience of unsettling situations. Reflexivity enables human organisms to move beyond simply reproducing social structures in transactional bodies through habit, toward performing intentional social acts and consciously reforming the social structures that influence them. Nurturing collective reflexivity enables humans to navigate and continue to shape their own communities in relationship with others—including and especially by working across difference.[44] Mead argues that collective reflexivity is indeed one of the conditions needed for changing the social structures of a community from within.[45]

This focus on cultivating collective reflexivity allows for a departure from the overemphasis on professional systemic design and moves toward enabling the intentional change of "traditions traditionally" within local communities, connected with Escobar's call for autonomous design.[46] It also resonates with Bela Banathy's

43 Ahmed Ansari, "Plural Bodies, Pluriversal Humans: Questioning the Ontology of 'Body' in Design," *Somatechnics* 10 no. 3 (December 2020): 286–305, ttps://doi.org/10.3366/soma.2020.0324.

44 Lake, "Pragmatist Feminism as Philosophic Activism," 25–45.

45 Mead, *Mind, Self and Society*, 134, 310.

46 Escobar, *Designs for the Pluriverse*, 172–73.

urge for everyone who affects or is affected by the consequences of design to be part of the designing community.[47] Banathy suggests that "[n]obody has the right to design social systems for someone else. It is unethical to design social systems for someone else. Design cannot be legislated, it should not be bought from the expert, and it should not be copied from the design of others."[48] A pragmatist perspective extends Manuela Aguirre's research in systemic design around cultivating co-designing cultures that nurture situated experimentation[49] by highlighting the need to carefully relocate such efforts beyond organizational contexts and into everyday community life.

This more mundane understanding of systemic design, brought forward through pragmatism, does not apply the model of professional systemic design to these more pervasive practices of designing social systems. Rather, recognizing the divergence across bodies and social contexts, it highlights a more generic understanding of systemic design, incorporating vast plurality in its embodiment. I argue that this reframing of systemic design demands a shift in the orientation of the field: from crafting expert-driven systems change to cultivating collective reflexivity so that everyone might continue to shape their own social worlds relationally. This perspective on systemic design reveals a critical means of community self-determination and an alternative approach to direct democracy within social systems.

Furthermore, a pragmatist perspective challenges systemic designers to ground themselves in their own bodies and situated social worlds, as well as to acknowledge that they cannot fully capture and consolidate the bodily experiences of others within social systems. Instead, they might work to nurture collective reflexivity and cultivate bodily habits of careful, relational experimentation in the crucible of experience among a community of inquirers. Learning from the work of Black feminists that are visionary pragmatists, it is crucial that systemic designers apply to this embodied, everyday work an interrogation of power and inequity; to this end, systemic designers need collective principles to guide pragmatic action, rather than seeing any ends as sufficient in this process of ongoing mutual transformation amid plurality.[50] Here, the principles of design justice that promote a community-led design approach, honoring lived experience and local knowledge while seeking liberation, are well aligned with such a pragmatist perspective and offer a hopeful starting point that demands to be further localized.[51] Furthermore, explicit connection to the growing body of work on oppression studies in design can aid systemic

47 Bela H. Banathy, *Designing Social Systems in a Changing World* (New York, NY: Springer Science+Business Media, 1996): 231–34.

48 Ibid., 228.

49 Manuela Aguirre Ulloa, "Transforming Public Organizations into Co-designing Cultures: A Study of Capacity-Building Programs as Learning Ecosystems" (PhD Thesis, Oslo School of Architecture and Design, 2020).

50 Hill Collins, "Piecing Together a Genealogical Puzzle," 88–112.

51 Sasha Costanza-Chock, *Design Justice: Community-Led Practices to Build the Worlds We Need* (Cambridge, MA: MIT Press, 2020): 190–204.

designers in recognizing the oppressive nature of existing social systems and in working consciously and carefully with struggles toward liberation.[52]

Everyday, embodied systemic design is not a design discipline and should not be disciplined, but if appreciated and nurtured, it can contribute to a more relational, pluriversal shaping of social systems. A pragmatist perspective challenges colonial narratives of progress, which responds to ideas of increasing complexity in design with new ways of reasoning. Alternatively, this perspective recognizes the incredible capacity of embodied human organisms to navigate immense complexity in the everyday and, together, to leverage their entangled agency to intentionally and materially adapt the social systems that they inhabit. In this way, banal bodily acts are not simply a history from which design has evolved; instead, they are the very foundation of autonomous social system transformations toward which design increasingly aspires.

Acknowledgments

I thank Katarina Wetter Edman and Danielle Lake for many evolving discussions about pragmatism and design over the years. I also express my appreciation to Amelie Dinh and Giulia Bazoli for their feedback on my personal stories included here.

52 Frederick van Amstel, Rodrigo Freese Gonzatto, and Lesley-Ann Noel. "Introduction to Diseña 22: Design, Oppression, and Liberation," *Diseña* 22, no. 2 (2023): 1–5.

Constructing accounts of decision-making in sustainable design: A discursive psychology analysis

Liz Cooper, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, United Kingdom

Many methods have been developed to help designers make better and more sustainable design decisions. Yet there is limited research on designers' perspectives on design decision-making. In this study, discursive psychology is used to analyse designers' accounts of decision-making. The designers show difficulty in trying to describe decision-making as an identifiable action. Different strategies are used to articulate how decision-making fits into the design process. Accounts of how decisions are made involve constructions of rational decision-making which are then undermined through 'confessions' of intuition. Decisions about sustainability are portrayed as made by other stakeholders, rather than by designers. The findings show that decision-making is a flexible construct that can be used to account for various actions.

Keywords: decision-making, design discourse, psychology of design, reflective practice, sustainability

Many sustainability and circular economy organisations focus on design decisions as being key to product sustainability (Chick & Micklethwaite, 2011; Fairs, 2019; Haug, 2017; Pritchard, 2013). The act of decision-making is therefore highlighted as an important activity for designers. It is commonly argued that the most important decisions related to product sustainability are made during the design process (Devon & van de Poel, 2004). In design literature, decisions are talked about in terms of criteria and choices. For example, designers may take into account sustainability criteria such as longevity, repairability, and supply chain impacts when choosing concepts, materials and assembly methods. Such design decisions are typically made within the constraints of a design brief, either set by clients or internally in a company (Ryd, 2004).

There is a substantial amount of literature providing guidance on how to make more sustainable design decisions (MacAskill & Guthrie, 2013). Most of this guidance takes an engineering design perspective, based on rational decision theory (Faber & Rackwitz, 2004; Schöggl et al., 2017). There is also some

Corresponding author:
Liz Cooper
liz.cooper@ed.ac.uk

www.elsevier.com/locate/destud
0142-694X *Design Studies* **84** (2023) 101158
https://doi.org/10.1016/j.destud.2022.101158

guidance on sustainable design decision-making from a creative design perspective (Buhl et al., 2019; Gould et al., 2019). Observation studies, for example by Cross (2001a) and Akin and Lin (1995), have sought to understand design decision-making through asking designers to 'think aloud' while doing design. However, there is limited research examining designers' own perspectives on design decision-making. Understanding how designers conceptualise decision-making in sustainable design in particular can give insights into how to make sustainable design guidance more appropriate. In this study, designers' accounts of decision-making in sustainable design projects are analysed to identify how the designers themselves characterise design decision-making when reflecting on their work. The findings show that articulating what design decisions are and how they are made can be challenging. The flexibility of the concept of decision-making allows the designers to use different strategies to navigate assumptions about the actions and agency involved.

1 *Contrasting perspectives on decision-making in design*

Different perspectives on how design decisions should be made have emerged and have been subject to much theoretical debate over several decades. While there is unlikely to be such a dichotomy in professional design practice, two contrasting approaches can be identified from the literature. On the one hand, a rational, science-based approach generally advocates a linear design process and the use of decision-support tools (Cross, 2001b; Hazelrigg, 1998). On the other hand, a creative approach embraces intuition and subjectivity as part of the design process (Buchanan, 1992; Cross, 1997). Both perspectives traditionally position the designer as central to design decision-making, although there is increasing recognition of design being a social process, involving extensive collaboration and negotiation among many stakeholders (Devon & van de Poel, 2004; Woodhouse & Patton, 2004). Both literatures include a growing focus on sustainability in design. In reviewing these literatures, we can better understand the assumptions made about the designer's actions related to decision-making and see how expectations are inferred regarding the designer's role in sustainability.

There is a vast body of literature based on a scientific, engineering design perspective, which proposes decision-making in design should be explicit and rational. Design decision-making is framed by engineering design authors as involving identifying options and weighing them up against criteria, often using mathematical formulae (Hatamura, 2006; Jin & Danesh, 2006; Kiker et al., 2005; Schöggl et al., 2017). This reflects classic decision theory which advocates identifying all alternative options and then selecting the optimal one (Kalantari, 2010). Numerous Multi-Criteria Decision Analysis (MCDA) tools are offered that claim to help designers or design teams make better, or often more sustainable, design decisions. Such tools offer mathematical ways of making comparisons between and ranking options against criteria (Kiker

et al., 2005). These can be in the form of computer software or more simple mathematical exercises that can be completed manually (MacAskill & Guthrie, 2013). Most of the literature on MCDA tools in design focuses on explaining how the tools can be used in particular cases and on recommendations for improving their technical functionality. Within these tools, there is an assumption that designers are aware of the point at which they need to make a decision, are able to identify the options, and simply need assistance in comparing options against known criteria in an objective way, thereby removing human subjective judgement (Fuente et al., 2017). Yet this goal of objectivity is criticised for ignoring the inevitable human involvement in the way decisions and criteria are framed by people involved, and in the choice of MCDA method (Smith & Ruiz-mercado, 2014; Steele et al., 2009; Tversky & Kahneman, 1981).

Lots of sustainable design tools are concerned with decisions about materials. For example, the MCDA tool Pugh's matrix compares lists of material properties against requirements and aims to help designers decide what materials to use in a product (Sridhar, 2007). Campaign organisations have also produced guidance on sustainable design decision-making. For example, the Circular Design Guide by the Ellen MacArthur Foundation provides advice on material selection. The guide explains how to research the impacts of different materials and answer a range of questions in order to make an informed decision (Ellen MacArthur Foundation and IDEO, 2018). However, it has been suggested that in reality materials are selected in less formal ways, based on prior experience and knowledge, and possibly aided by informal internet research (Khatib, 2016; Sridhar, 2007). MCDA tools in general are criticised for being overly complex and time-consuming to use (Schöggl et al., 2017; Silva et al., 2009). There is currently a lack of research on understanding the extent to which MCDA type tools are used in sustainable design in industry settings. There is thus an opportunity to explore the extent to which designers themselves report using formal, rational decision-making methods in practice when asked to reflect on their design methods.

There is also a large amount of theoretical literature promoting the appropriateness of creativity and intuition in the design process (Cross, 2001b; Plessner et al., 2011; Schon, 2008; van de Poel, 2015). For example, the 'design thinking' approach is widely advocated, which encourages designers to follow iterative processes involving extensive ideation and prototyping, with a focus on stakeholder engagement and teamwork, in order to solve design problems (Geissdoerfer et al., 2016). Design thinking methods are widely advocated as being useful for achieving sustainability aims, due to their focus on taking into account different stakeholder perspectives in creative ways (Buhl et al., 2019). In creative design, the language of problem-solving is typically chosen over decision-making, since, in this paradigm, design is considered to be about finding new solutions that don't yet exist, rather than choosing between known

options, as in the science-based methods (Ball et al., 2001). Problems are said to be solved through doing design, including the many visual and tactile aspects such as sketching and prototyping (Gumienny et al., 2011) as well as the interplay between the visual and the verbal (Jacobsen et al., 2016).

Nevertheless, some authors have explicitly used the term decision-making when theorising about design thinking, despite the primary focus on problem-solving (Buhl et al., 2019; Gould et al., 2019; Liedtka, 2015). In this literature, the designer is clearly positioned as the one making the decisions, in spite of there being a focus on stakeholder engagement and collaboration in design thinking methods. For example, Buhl et al. (2019) frame the designer as deciding what stakeholder feedback to use and how. Yet the authors focus more on advising the designer on how to gather perspectives to inform decisions, than on the act of making a decision. Other authors propose more specific methods for decision-making when using design thinking methods. For example, Gould et al. (2019) propose combining design thinking with decision support tools for sustainable design. This would involve combining a creative, exploratory approach to design with a rational decision-making method that assumes clear decision points can be identified and that seeks objectivity in decision-making. Relatedly, Liedtka (2015) suggests that design thinking, through its focus on openness, visualisation, and collaboration, can help reduce cognitive biases in design decision-making. Again, there is an attempt to be objective, through reducing biases, albeit through using a creative method. Calabretta et al. (2017) propose that designers are more likely to prefer intuitive approaches, while design managers are likely to expect rational methods, and that both should embrace the contradicting forces of rationality and intuition in decision-making. Based on these different assertions, we can conclude the interplay between objectivity and intuition in design is complex.

1.1 Articulating decision-making in design

Given the extensive theory on how design decisions should be made, some authors have sought to study how they are actually made, often through asking designers to articulate them. Two methods have been used in which participants are asked to articulate what they are doing while designing, including decision-making. The first approach to observing the design process is protocol studies, where designers are asked to work to a design brief in a laboratory setting, and to simultaneously talk out loud about what they are doing and thinking (called 'thinking aloud') (Cross, 2001a; Oxman, 1995). The 'thinking aloud' method reflects a cognitive perspective on design, and so designers may be expected to talk about deciding or choosing as mental processes, which might involve other mental processes such as imagining, evaluating, and perceiving (Cross, 2001a; Dorst & Dijkhuis, 1995; Oxman, 1995). Several historic protocol studies produced findings on how decisions are embedded within a creative design process. For example, Akin and Lin (1995) identified

novel design decisions as associated with multimodal design activity, being more common when someone was drawing, examining, and thinking at same time. In other cases, design teams are observed working to a brief. Cross (1997) found that a 'creative leap' occurred among a design team, indicating a solution was suddenly found. Findings from such protocol studies imply that decision-making in design is done creatively and/or intuitively in the cases studied. However, protocol studies are criticised by some authors since the context of taking part in a laboratory study will influence how participants report what they are thinking, and may influence the design work itself (Dorst, 1995; Lloyd et al., 1995).

Given the difficulty of expecting designers to report what decisions they are making and how in real-time, the second approach involves asking designers to talk about decision-making in design retrospectively, through interviews. Again, the context of taking part in an interview study will influence what participants say (Rapley, 2015), but the retrospective approach gives space and time to the designers to construct their own accounts of decision-making in a project more holistically. However, there has only been limited research involving explicitly asking designers to give accounts of decision-making in specific design projects. One example comes from the designer and researcher Pedgley (2009) who completed interviews with product designers, plus a reflective diary of his own design practice, to try to understand design decision-making with regards to materials choices. The analysis focused on how designers carefully balance stakeholder requirements (for example, those of clients or manufacturers) with their own expertise. Another interview study by Surma-aho et al. (2019), examined empathy among designers. The analysis focused on how different types of empathy were associated with different design priorities. In the extracts presented by Surma-aho et al. (2019), the participants talk explicitly about how empathising with users influences their decision-making. Both studies produce findings that highlight how designers take into account other stakeholders when making decisions. However, what participants said is reported as if it reflects what actually happened in the design process. The nature of reflective accounts as constructive and constructed within a specific interactional context (Rapley, 2012) was not considered.

Additionally, some studies have sought to identify how design decisions are made interactionally in design meetings. Analysing videos and transcripts of such meetings to understand the social aspect of design has become a popular method (Luck, 2012; McDonnell & Lloyd, 2009). Some of these studies have produced findings about particular aspects of decision-making, such as persuading or convincing others, where designers and other stakeholders are seen to be discussing design options collectively (Le Dantec & Do, 2009; McDonnell & Lloyd, 2009; Oak, 2011). This body of work gives insights into the need for designers to develop negotiation skills as well as creative

and/or scientific design skills. Yet this approach can only give insights into decision-making in meetings, and not into decisions that take place in the design studio or elsewhere. There is therefore still a need to gather designers' insights and perspectives on decision-making that takes place in different stages of the design process.

To sum up, in the normative literature on design decision-making, there is a tendency to frame designers as decision-makers, while acknowledging the collaborative nature of design. The engineering design perspective assumes that clear decision points are known and that decisions can be made rationally. The creative design perspective appears to treat some decisions as consciously made, drawing on design experience and stakeholder views, and others as embedded within the creative process of doing design. This literature overall focuses much more on how decisions are made or should be made, rather than on what decisions are made in design. There has only been limited research to date that seeks designers' own accounts of design decision-making across a project. There is therefore an opportunity to take a more detailed approach to analysing how designers construe decision-making in particular projects. In focusing such a study on sustainable design specifically, there is an opportunity to provide new insights to help improve guidance on sustainable design decision-making, to in turn achieve more sustainable products. In this study I analyse designers' accounts of design decision-making related to sustainability to answer the question: *How do sustainability-focused product designers account for what design decisions are and how they are made?*

2 Methods

A discursive psychology (DP) approach has been taken to analyse designers' accounts of decision-making in sustainable design. DP offers a method for analysing talk as actions, and often leads to a respecifying of how psychological concepts are viewed (Edwards, 1999; Potter & Wetherell, 1987). Decision-making is a psychological concept, since it is a term used to refer to mental processes related to evaluating and choosing (Bouyssou, Dubois, Pirlot, & Prade, 2009; Kalantari, 2010). Rather than seeking to access how decision-making happens in people's minds, we can instead use DP to study how designers construe this psychological concept in the ways they talk about it. When designers are asked to give accounts of decision-making in interviews, different questions about decision-making are likely to be met with different ways of accounting for and justifying the design process. Thus, rather than only analysing the content of what designers report, their accounts of the design process can be analysed in terms of how they are constructed in the interactional contexts of specific questions asked (Wiggins, 2017). DP has not been commonly used to look at design contexts before, although the closely related approach of conversation analysis which seeks to understand the nature of

interactions per se, rather than as related to psychological concepts and actions, has become an increasingly popular method to study observations of design practices (Luck, 2012).

Designers' accounts of sustainable design projects were collected using semi-structured interviews, as part of a wider project on psychology and design. Using interviews invites reflective conversation on design practice and enables the generation of interactional accounts specifically about decision-making in design. Reflection on past work is commonly advocated in design education and practice (Schon, 2008), and so designers are likely to be familiar and comfortable with producing detailed accounts of how they carried out their work. Potential participants were contacted via a design email list and a sustainable design group on LinkedIn. Sixteen product designers were recruited to take part in semi-structured video call interviews carried out by the author between July and October 2020 (see appendix A for participant details). Interviews lasted an average of 45 min. Designers were told in advance that they would be asked to give an account of decision-making in a recent design project of their choice. Participants were asked to tell the interviewer about some of the decisions made in the project and then to explain how a particular decision had been made. University research ethics committee approval (School of Philosophy, Psychology and Language Sciences Research Ethics Committee, approval number 324-1920) was given before commencing data collection. This confirmed that appropriate measures were taken to ensure informed consent (using participant information sheets and both email and verbal confirmation), anonymity, and data security in line with British Psychological Society ethics guidelines (Bouyssou et al., 2009).

After watching the video recordings, it was decided that the analysis would focus on the verbal data, without including any multi-modal aspects of interaction such as gestures, since these would not significantly add to the analysis. Basic words-only transcripts of the full dataset were produced. Anonymised transcriptions can be found on the UK Data Service (Cooper, 2021). Analysis involved several iterations of looking closely at the full dataset, making notes on patterns of actions, devices, and sequences related to accounting for design decisions. The context of taking part in an interview and the framing of questions were important aspects of the analysis (Cooper & Burnett, 2006; Rapley, 2012; Wooffitt & Widdicombe, 2006). Attention was paid to how differences in question wordings were associated with different types of responses. Nine lengthy extracts (one to three pages each) where decision-making was discussed were selected for detailed transcription using Jefferson (2004) notations (see appendix B for a guide to transcription symbols), for further in-depth analysis. Patterns were identified regarding different ways being asked to identify specific decisions was treated as problematic, and contradictions in ways of describing how design decisions were made. Once specific findings were identified, the transcripts of the full dataset were then checked again to identify

how often the patterns occurred. Three extracts were shared at a data session with other researchers, who noticed similar phenomena, providing confidence that the analysis is rigorous and justifiable. Four short extracts have been selected for inclusion in this article, which represent clear examples of the patterns of phenomena identified across the wider corpus of extracts analysed.

3 Analysis

The analysis is structured in two parts, with two extracts on identifying what design decisions were made, and two on describing how design decisions were made.

3.1 What were the design decisions?

Here I examine two extracts in which participants indicate that questions about what decisions were made are problematic, since the participants do not report specific decisions. Extract 1 comes from early on in an interview with a designer talking about a trestle table he designed for a university client.

Extract 1, from interview 1 – trestle table

1	I	and the:n=
2		=this might be a little bit (.5) more tricky but could
3		you <briefly> (.7) >tell me about some of< the decisions
4		that you made s:o we=
5		=we can go into more detail afterwards but but what were
6		the particular decisions that you had to make in
7		designing this (.4) product
8	P1	(2.8) e::rm (1.9) well decisions yeah (huhuh)
9		i think designing is always decision-making [e:rm
10	I	[yeah]
11	P1	the whole time (.5) erm (1.9) i mean at the=
12		=at the end of the day you have a (1.2) you have a
13		certain erm e::r there was maybe like a vague i↑dea by
14		the universi↑ty what they ↑wanted and then (.3) it's
15		(1.0) our role as designers to give this idea (1.4)
16		e::rm
17		a form basically to to (.3)
18	I	yep
19	P1	bring it into be:ing (.7) and erm (.5) so (.4) decision
20		making i would say is erm (1.8)
21		e:r(1.0)
22		well i mean there are there are lots of lots of tiny
23		tiny decisions erm at th- i mean at the very beginning
24		erm you you develop maybe a erm a rather abstract idea

25 of of how something could be and then through a
26 iterative process of erm model making and sketching (.5)
27 you erm (.6) you=
28 =yeah you develop this idea but it's erm (.4) oftentimes
29 not so much that you (1.5) erm (1.0) only decide how
30 things should be but (.4) that you erm do something and
31 then the the thing that is in front of you (.5) maybe a
32 drawing or a a model (.6) erm (.4) er speaks back to
33 you=
34 =and erm >in a way< erm s:o (.6) it's erm and then then
35 of course you react to it and e:rm and so (.5) there are
36 hundreds of of decisions detailed decisions

The interviewer's (I) question is prefaced with 'this might be a little bit more tricky' (line 2), which establishes the request to specify decisions made as potentially difficult. This signals to the participant that they may not necessarily be expected to give a straight answer. The question is then formulated as 'could you briefly tell me about some of the decisions you made' (lines 3–4). This is an open question, which allows the participant to select which decisions to focus on. The use of 'some of' portrays a larger number of decisions that the participant can select from. The question is then reformulated, as 'what were the particular decisions that you had to make in designing this product' (lines 5–7). This reformulation makes the request more specific, in saying 'particular decisions', and makes clear the interest in the specific design project, saying 'this product'. The disclaimer that the question might be difficult, and the hesitancy in how to frame the question, indicates possible problems with expecting a clear account of decision-making.

In the designer's (P1) response to this question there are long pauses and an elongated 'erm' in line 8, followed by 'well decisions yeah' and laughter. This both buys time before providing a response and signals that the request to identify decisions is tricky. Pauses and hedging (delaying sounds such as 'erm' which highlight delicateness (Wiggins, 2017)) feature regularly in the extract, which implies that the narrative is in some way difficult to construct. We therefore see that, as indicated in the question, the designer does indeed demonstrate difficulty in formulating a response.

In the rest of the extract, the designer offers a detailed account of the design process. First, he makes a general claim: 'I think design is always decision-making … the whole time' (lines 9–11). This enables the participant to shift the conversation away from identifying specific decisions, as he makes the issue general by using the extreme case formulations 'always' and 'the whole

time'. Such extreme case formulations have been shown to be used to portray something as typical (Edwards, 2000; Pomerantz, 1986). Through this response, he indicates the difficulty of identifying a specific instance of a decision as an action. He then provides an alternative account which reframes his initial portrayal of a continual process, by describing lots of small decisions. He says, 'lots and lots of tiny tiny decisions (lines 22-23) and 'hundreds of decisions detailed decisions' (line 35). The small scale and the large volume of these decisions are highlighted through repetition. This plays down the significance of any specific decisions in the design process, by characterising the decisions as small and commonplace. The features of this response thus work to demonstrate the problems with being asked to specify decisions made.

The participant also offers an account of alternative actions in the general design process. He provides several descriptions of what design involves. The designer's creative role is depicted in saying 'it's our role as designers to give this idea erm a form basically' (lines 14–17), and then adds 'to bring it into being' (lines 17–19). In lines 25–28, he describes design again as a process rather than a series of decisions, this time offering a more detailed description. He says, 'though a iterative process of erm model making and sketching you erm you yeah you develop this idea'. In lines 30–32 he reports 'you erm do something and then the thing that is in front of you maybe a drawing or a model erm er speaks back to you' and then adds in lines 33–34 'and then of course you react to it'. The notions of giving a form to something and bringing it into being portray a practical process of acting and reacting to material things. The metaphor of the drawing or model 'speaking back to you' implies that the agency to do design work lies not only with the designer, but also within the objects themselves. This helps the participant avoid talking specifically about specific decisions he has actively made himself. The account of alternative actions in the design process, which contrast to the notion of making decisions, further portrays difficulty with the assumptions in the original question that specific decisions can be identified.

Furthermore, it is worth noting that throughout this detailed description of the design process, the designer talks about designers as a category. He uses the generalising 'you' pronoun (lines 12, 24, 27, 29, 30, 31, 33 35), which portrays the aspects of the design process he is describing as typical for designers. This also helps him reject the request to give an account of specific decisions he made.

Next, we turn to an extract that shows the participant indicating that the question of what decisions they had to make is problematic in a different way. In extract 2, a product designer who has been talking about designing packaging for shoes for a client, is asked to talk about some of the design decisions.

Extract 2, from interview 9 – shoe packaging

```
1   I     yeah (.) ↑okay (.6) s::o (.3) could you ↑tell me about
2         some of the things that you had to make design
3         decisions a↓bout

4   P9    (.3) mm (hhhh) s::o (1.3) >i think some of the biggest
5         things< (.3) um because this company that we worked
6         with they had never (.7) done any like sustainability
7         initiative pro↑jects be↑fore (.8)

8         it was getting them familiar with the sustainable (.6)
9         materials um (.6) >you know>
```

The extract starts with an open question. The participant (P9) is asked 'could you tell me about some of the things you had to make design decisions about' (lines 1–3). This is a different question to the one asked in extract 1, since rather than asking what the decisions were, the participant is asked what they were about. Nevertheless, the question clearly directly asks the participants about design decisions he had to make himself.

This question is met with pauses and hedging ('mm', 'so') in line 4 which signals difficulty in answering at first. Then, instead of providing an account of what the decisions were, the designer then gives a response that picks up on the question of what the decisions were about. The participant first provides his response, saying 'I think some of the biggest things … ' (lines 4–5). This focuses his account on what he assessed as important. We see that 'decisions' is modified to 'things', which enables him to talk about what the decision was about. He then provides an insertion, used to provide information needed for the subsequent claim to be understood in a particular way. In the insertion he explains that the client company had not worked on sustainable projects before (lines 5–7). The designer then identifies one of 'the biggest things' that a decision was made about as 'getting them familiar with the sustainable materials' (line 8). The designer focuses his response on an aspect important to sustainability, materials choices, thereby orienting to the broader context of the interview, despite sustainability not being in the interviewer's question. The designer thus claims that he had to make a decision about getting the client familiar with sustainable materials. This response does not give any detail about what the decision was, it only highlights this aspect as important, and portrays a challenge regarding the client being less informed about sustainable materials. It can be inferred from this response that the original question, in its focus on design decisions made by the designer, misses the important role of other stakeholders such as clients in decision-making about sustainability. The designer's role is portrayed as working to get the client familiar with more sustainable options, so that the client could make a more sustainable decision.

3.2 How were the decisions made?

In extract 1, when the participant was asked to talk about some of the decisions made, we have seen the participant talk about decisions as embedded within the design process. Thus, no specific decisions are identified. In extract 2, we have seen that when the participant is asked what he had to make decisions about, he highlighted that an important decision about sustainable materials was made not made by him. Despite the difficulty seen across the dataset in talking about specific decisions that the designers themselves made, some participants do go on to identify a decision. We now turn to two extracts where the interviewer asks more direct questions to try to gain more detailed accounts specifically of how these design decisions were made.

The participant in extract 3 has previously talked about a project he undertook to propose a sustainable packaging solution for a ceiling fan to the consumer goods company he works for. The focus was on using a more sustainable material, and he had briefly mentioned having decided between three different material options.

Extract 3, from interview 2 – ceiling fan packaging

I okay (.3) interesting (.6) and so you we:re=

=initially you had these thre::e different material options (1.3) how did you make the final decision (.8)

how did you weigh these options up

P2 (.6) yeah (1.0)

e:rm (.9) s:o (.6) in the end erm (.6) the company is er company caters to the economy er sort of range=

=all the products are based (.3) are very economical

they aren't er really into the premium segment or the luxury segment so cost was always (.5) e:r on the back of my mind that (.5) whatever we do (.5) it's not adding to the product as such you ↑know

I mhm

P2 it might add to the marketing but we had t:o keep in mind that the cost is one of the major factors so that was the first filter (.4)

((36 lines omitted where the participant describes two other factors influencing the decision, the environmental impact of the material and how easy it is to manufacture))

I and was it difficult to compare these different factors and make a decision

P2 (1.7) e:r to be honest i guess erm (2.3) i could have dealt before or you know i could have done a more (.4) er (.4) detailed com[parison

I [mhm]

```
27  P2  of all these (.5) now in retrospect i'm citing these
28      options but (.7) when i was in the whole design process
29      or the kind of process (.3) er ((product name)) (.9)
30      like impressed me so much on the personal front and also
31      the sustainability department people were like 'wow this
32      is really nice'

33  I   yep

34  P2  so i feel we were a bit biased toward towards it (.7) so
35      (.6) that i do realise now when i look at it so we were
36      trying to make er (.4) that option [work

37  I                                      [okay]

38  P2  (.5) because it was s:o erm (.4) sort of er it ticked
39      (.5) all the requirements which we had
```

In the opening lines, the interviewer (I) produces a summary or gist. The gist is used to selectively reproduce elements of the participant's prior talk to frame as a follow up question. The interviewer's question is prefaced with 'initially you had these three different material options' (lines 2–3), where 'three' is elongated. This prefacing works to highlight the idea of having options to decide from. The question is first framed as 'how did you make the final decision' (line 3) and then followed up with 'how did you weigh these options up' (line 4). The first question remains open to different sorts of accounts of the decision-making process, but the second question, in using the phrase 'weigh up', again makes clear the interest in involving comparing options. This gives an example of the kind of specific action that might be involved in decision-making.

The questions are responded to initially by the designer (P2) with several pauses and hedging 'yeah erm so' (lines 5–6), indicating some difficulty. And then the participant provides a detailed description of three criteria used to weigh up the material options, describing 'cost' as the 'first filter' in lines 15–16, and then adding further detailed description of environmental and practical factors (lines omitted). This response shows what sort of things can be said about the specific activities involved in making a decision, by describing the criteria used.

The interviewer then probes further, which might imply that the initial answer was not sufficient. She asks in lines 21–22, 'and was it difficult to compare these different factors and make a decision'. Again, the emphasis is on comparing the options as the way to make a decision, which steers towards a particular kind of response, and the interviewer asks for an assessment of whether this was difficult. This question is met with a long pause (1.7), indicating some possible difficulty, and then an upcoming 'confession' is tentatively announced by saying 'er to be honest I guess erm' (line 23). The

'confession' is then given, saying 'I could have done a more er detailed comparison' (lines 24–25). Here the participant is indicating recognition that the initial answer was not treated as sufficient. Analysis of interview studies has found that participants often seek to give answers based on what they think a researcher is looking for (Rapley, 2012). Therefore, by responding more directly to the interviewer's suggestion of comparing factors in order to make a decision, the participant demonstrates conforming with the researcher's request. The 'confession' continues in lines 27–28 'in retrospect I'm citing these options but … ' which announces an alternative account is about to be provided. The participant says, 'when I was in the whole design process or the kind of process er ((product name)) like impressed me so much on the personal front and also the sustainability department people were like 'wow this is really nice'' (lines 28–32). This second account portrays a personal reaction of being 'impressed', rather than focusing on providing facts or criteria for analysis. Reported speech depicts assessments by colleagues of the product being 'really nice'. This response perhaps indicates difficulty in articulating the specific action of choosing, since this is portrayed as just based on liking one option.

The admission of having taken a decision based on a personal reaction continues, with the designer saying, 'so I feel we were a bit biased toward towards it so that I do realise now when I look at it so we were trying to make er that option work' (lines 34–36). The use of 'biased' indicates an ideal of making decisions objectively, in contrast the portrayal of actually of having made a decision based on being impressed. The phrase 'I do realise now' portrays reflecting back on the decision-making, and also again orients to the interviewer having indicated that the initial account given was not sufficient. However, the concluding statement in lines 38–39 'because it was so erm sort of er it ticked all the requirements which we had' shifts back to the idea of rational decision-making, in asserting that the material nevertheless met the criteria. This works to counteract any possible inference of cutting corners and making a less effective decision based on personal reactions. Overall, we see the designer orient to common normative assumptions, indicated in the interview questions, of design decision-making as being a rational, explicit process, but in the end shows difficulty and hesitancy in describing decision-making in this way.

In the next extract, we see how a different designer also first portrays seeking to make a decision in a methodical way, and then confesses that intuition is used. This participant had previously talked about an electric vehicle charge point he had designed while working for a small design company. An informal call had been held with this participant in preliminary research, where the use of decision support tools in design had been discussed.

Extract 4, from interview 12 – electric vehicle charge point

I okay >and< could you tell me a bit about the design process and how you went about making decisions

P12 (1.2) sure e:rm yeah i think I think we might have touched on our last call=

I =yeah

P12 i so i really (.7) er for this (.6) i really like a >I quite like a process< called morphological analy↑sis

I mhm

P12 (1.0) erm which is where i essentially break down e:r (.7) (.4) the functions of the product you're trying (.5) er to develop erm (.5) or you yeah >you break down er the product the product into its kind of< simplest basic functions

I mhm

P12 erm and then you can combine those to create e:rm (.6) concepts design concepts that you can then assess on how erm >on how good they are<

((55 lines omitted where the participant talks about some of the different options available when designing the electric vehicle charge point, and then gives detail about what the process looks like on paper, involving comparing lots of sketches in a huge table))

I yeah and so you've got you've got all these opti:ons in a big grid

P12 mhm

I >how do you actually< <then choose>

P12 (1.4) erm so it's (1.0) yeah this is one of th:e (1.2) this is maybe one of slightly controversial bits about the process is it's largely (.8) intuitive

I mhm

P12 erm or based on experience

In the initial question, the interviewer (I) asks an open question both about the design process for the project and specifically about decision-making, saying, 'can you tell me about the design process and how you went about making decisions' (lines 1–2). The initial response from the designer (P12) is hesitant, with pauses and hedging (sure erm year, line 3), before saying 'I think we might have touched on our last call' (lines 3–4), referring to the previous conversation held during pilot research. In lines 6–7, the designer begins his response by identifying a formal process for decision-making, saying, 'I really like a I quite like a process called morphological analysis' (lines 6–7). This process is described using an official name, indicating a technical, standardised approach. However, there is hesitancy around depicting using the process because the designer personally likes it, as we see a modification from 'I really like' to 'I quite like'.

Accounting for design decisions

The participant then gives an explanation of what morphological analysis is. This starts in lines 9–10 as a description of how the designer himself uses this process, saying 'I essentially break down er the functions'. But then 'I' switches to 'you' in the rest of the extract, to talk about how designers in general may use the process. The use of generalising language portrays the process as a commonly accepted method. For example, he says 'and then you can combine these to create erm concepts' (lines 15–16). This works to make the account appear universal and typical for designers. The participant then provides a lot of further detail (in lines omitted) about how the process is done, first in describing different options when designing the electric charge point, and then in describing what the method looks like on paper. This detailed account portrays knowledge and familiarity with the process which adds credibility.

The interviewer probes further by summarising what the participant has said about the process, saying, 'yeah and so you've got you've got all these options in a big grid' (lines 23–24), and then asking, 'how do you actually then choose' (line 26). Thus, we see that the detail given so far about looking at the options and considering criteria is not treated as giving enough information on the actual process of deciding. The quantity of options is highlighted through 'all these options' and 'big grid'. This question itself is very direct, and the use of 'actually' portrays seeking an honest response that reflects what really happened. The words 'then choose' are said more slowly, which works to emphasise the choosing as the important bit. In response to this direct question, we see some difficulty in continuing the narrative of the technical approach. There are first pauses and hedging ('erm so it's … yeah', line 27), and then the participant adds to his account in response to the signal from the interviewer that the initial account wasn't adequate. He gives a signal that a 'confession' of something different is coming, declaring, 'this is maybe one of the controversial bits' (line 28). The 'confession' that is then delivered is 'it's largely intuitive erm or based on experience' (lines 29–31). Decisions being made intuitively is portrayed as a possibly negative thing through the term 'controversial', or at least something that goes against expectations, through describing it as controversial. The modification of 'intuitive' to 'based on experience' works to counteract any possible negative connotations of intuition, by highlighting the role of the designer's expertise in the decision-making.

4 *Discussion*

In this analysis of interviews with sustainability-focused product designers, we have seen that the designers use different strategies to articulate how decision-making fits into the design process. When asked about design decision-making as an activity, decision-making is treated in different ways depending on the question, but there is difficulty in clearly describing what took place. From the analysis of the sixteen interviews, I identify the following findings. 1)

Asking what decisions were made in a project is treated as problematic. The idea of identifiable decisions as specific actions is brought into question, through giving accounts of many small decisions being embedded in a creative process, and through describing design as involving alternative actions to decision-making (seen in eleven interviews). 2) The assumption, implied in asking designers about the decisions they made in sustainable design projects, that designers make key design decisions related to product sustainability is also brought into question. Some designers instead claim that important design decisions were about seeking to inform and influence other stakeholders to select more sustainable options (seen in five interviews). 3) Asking how specific decisions were made in a project is responded to with accounts of rational decision processes, but which are then undermined with 'confessions' of decisions being made based on personal assessments and reactions (participants gave such contrasting accounts of the same decision process in seven interviews). Since the findings are based on analysis of lengthy sequences of talk, it is not possible to include the evidence for all of the instances in which the findings can be seen in all sixteen interviews in this paper. Instead, as is typical in DP (Wiggins, 2017), selected examples of sequential data representing the three findings have been presented and analysed.

From finding 1, we can infer that talking about design decisions as specific, identifiable actions is difficult and that it is instead easier to construct a retrospective account of the overall design process in which decisions must have somehow occurred. The design process is portrayed by the designers as messy and complex, rather than involving clear decision-making based on criteria. This is in response to being asked about design decisions in a general way, rather than about design decisions specifically related to sustainability. Thus, the participants produce accounts of the process of designing in general. These designers' accounts resonate with the idea put forward by several authors that their design practice involves intuition, embedded and embodied knowledge (Cross, 2001b; Liedtka, 2013). Empirical research based on observation of design has found collaborative decision-making to be often done intuitively rather than based on careful consideration (Lloyd & Busby, 2003) and at times to have involved what have been referred to as 'aha' moments (Cross, 1997). However, the conclusion from the present analysis is not that design decisions are made intuitively, as this would involve assuming what people say reflects an underlying reality. Instead, we can conclude from these findings that decision-making is a construct that is used to make sense of a variety of actions, rather than to describe one identifiable action. Rather than treating design decision-making as either rational, using clear criteria, or intuitive involving creativity and feelings, designers can portray and navigate creativity, skills, and agency in different ways when accounting for decision-making.

Accounting for design decisions

From finding 2, we can infer that the designers are taking opportunities to portray the limits of their individual agency over important design decisions related to sustainability. When asked what they had to make decisions about, we have seen an example of a participant claiming that he needed to make a decision about making the client aware of sustainable materials. The designer does not describe what the decision was, but describes what was important, which was to influence the clients to understand the importance of sustainable materials. Across the dataset designers who work in many different contexts similarly portray sustainability-relevant decision-making as often done by other stakeholders, rather than by the designers themselves. In theoretical literature, it has been argued that the designer's role should also include influencing other stakeholders towards sustainability, given the social nature of design (Fry, 2004; van de Poel & Verbeek, 2006). This study provides empirical evidence of sustainability-focused designers indeed portraying an aspect of their roles as involving trying to influence, in the absence of being able to decide. In these instances, the construct of decision-making is used to comment on agency, rather than specific actions or process.

From finding 3, we can infer that the participants are orienting to expectations that design decisions should be explicit and made in rational ways. These expectations are sometimes explicit in the way the interview question is asked ('how did you weigh these options up', extract 3). The participants knew in advance of the interview that the researcher's primary interest was decision-making, so some may have thought about how to describe their decisions beforehand. Since the participants speak as research participants, they were potentially seeking to fulfil the researcher's needs (Rapley, 2012) by providing a detailed account of considering different options in order to make a decision. We have seen two examples of participants first constructing accounts of organised, sequential processes involving comparing options, but then undermining these accounts by 'confessing' contrasting accounts of making decisions based on personal impressions or intuition. The latter are portrayed as controversial as they go against expectations of objective decision-making. This finding resonates with those of other authors who have concluded that expectations of fully rational decision-making in design are unrealistic. For example, Zannier et al. (2007) found software designers used a mixture of rational and intuitive approaches, and Guersenzvaig (2015) propose that rationality and intuition are two elements of human decision-making in design, rather than polar opposites. Calabretta et al. (2017) consider intuition and rationality to be contradictory in design decision-making but argue that they should nevertheless be somehow integrated. Looking more broadly than design, the fact that two different sorts of accounts of how the same decision was made are provided reflects research in social psychology that has found that people tend to retrospectively construct logical narratives of how and why they made decisions, after initially making them based on gut instincts (Haidt, 2001). While we cannot conclude from the analysis that the designers

in fact made the decisions intuitively and then constructed a rational account for the purpose of the interviews, we can conclude that there is something tricky about trying to articulate how a decision was made. Again, this indicates that the concept of decision-making is used to bring together many different actions and processes, and so is difficult to describe as a whole.

Based on these findings, design decision-making can be reconceptualised as a tricky yet flexible concept that is constructed in talk in different ways. Prior research has used different methods to seek to identify whether design decision-making in specific projects involves intuition, rationality, or both (Calabretta et al., 2017; Cross, 1997). The present study makes a new contribution to design studies by demonstrating that the ways in which design decision-making is articulated are situated in their specific interactional contexts. The same design process can be described in different ways, for example as involving intuition or rational comparing of options, depending in part on what the designer is specifically asked and on what expectations are inherent in the given context. The notion of decision-making can be used to focus on different actions and processes, or on agency. The DP analysis thus demonstrates the importance of considering reflective accounts of design work as constructed and constructive (Wiggins, 2017). For studies based on interviewing designers about their work, analysis should consider both questions and responses (Rapley, 2012), and should treat accounts as constructed within their interactional contexts, to achieve different things, and manage different assumptions, rather than reflecting a reality of what took place. Using DP to analyse interactional talk provides a new direction for understanding how being held to account for one's actions affects reflective practice in design.

The practical implications of these findings are as follows. Guidance on design decision-making could be adapted to reflect the trickiness of the concept. For example, guidance on making better design decisions could include some extracts of designers' reflective accounts of decision-making in specific projects. This would demonstrate the complexity of talking about decision-making in design and how designers make sense of their actions by sometimes portraying rational comparing of options, sometimes portraying a creative, intuitive process, and sometimes portraying negotiating a choice with others. This could encourage further reflection and discussion on how decision-making is done in design from designers' perspectives. Guidance on sustainable design decision-making in particular would benefit in particular from reflecting the complexity in how the designers' agency to make decisions is portrayed in their accounts. We have seen portrayal of some sort of agency over many small decisions embedded in the creative process, but not over making final decisions regarding how sustainable products are. Thus, when giving guidance on making explicit decisions that are key to sustainability, such as materials, it may resonate with designers to include advice on ways of negotiating with, persuading, and influencing other stakeholders to prioritise sustainability, as

well as on actions such as identifying and comparing options that may form part of deciding. Such advice may be informed by identifying successful negotiation and persuasion strategies employed in design meetings (Luck, 2009; McDonnell & Lloyd, 2009; Oak, 2011).

To improve this study, the interviews could have probed more to seek extended accounts of the actions involved in decision-making. The fact that the interviews covered multiple topics as part of a wider project meant that the number of follow-up questions was limited. Further research could involve interviews fully dedicated to the topic of decision-making in design, with specific and consistent questions on decision-making in general, and then on decisions related to sustainability.

5 *Conclusion*

We have seen in this article that a great deal of specific guidance exists on how to make better design decisions, including more sustainable decisions. Most of this reflects rational decision theory rather than a creative or intuitive perspective on how design decision-making is done. While observation of design work has provided a range of findings indicating creativity, subjectivity, and intuition in design decision-making, there is a lack of research involving detailed analysis of designers' own accounts of decision-making in their work. Through a DP analysis of sustainability-focused product designers giving accounts of design decision-making, this study provides the following findings. 1) Being asked to give an account of what decisions were made in a sustainable design project is treated as difficult in terms of being able to identify specific decision points in a creative process. 2) When talking about what important decisions related to sustainability were about, participants describe influencing other decision-makers, thus highlighting a challenge of agency to make key decisions. 3) Being asked to describe how decisions were made is met with constructions of rational decision-making processes, sometimes reflecting a rational framing in the question asked, involving comparing identified options. However, this is then this is undermined through 'confession' of the centrality of personal reactions and assessments in design decision-making. Overall, we see difficulty in articulating both what decisions were and how they were made in a design project, but nevertheless see attempts made to construct informative accounts of decision-making and design processes, and to portray the designer's effort in finding appropriate solutions to a sustainable design problem.

Based on the findings of this study, there is an opportunity to reconceptualise decision-making in design, as a flexible construct used to account for various actions and processes, that is used in different ways in different situations. Practical guidance on design decision-making could be modified to more closely reflect the variety of ways in which designers articulate their actions

and roles. For design researchers, this analysis has shown the usefulness of DP for analysing interviews as interactions and provides a different methodological direction for analysing reflective practice in design.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Dataset can be found at https://doi.org/10.5255/UKDA-SN-855100.

Acknowledgements

Thanks to Sue Widdicombe and Craig Martin for their guidance on this work, and to the anonymous reviewers for their useful feedback. Thanks to the University of Edinburgh for funding the project through a Principal's Career Development PhD Scholarship.

Appendix A. Interview participants

	Location	*Sex*	*Type of project talked about*	*Product type*
1	Germany	Male	Professional project - in-house	Furniture
2	India	Male	Professional project - in-house	Packaging
3	US/Netherlands	Female	Professional project - in-house	Luggage
4	Argentina/Italy	Male	Professional project - independent	Furniture
5	UK	Female	Internship - design agency	Child's bike
6	UK	Female	Postgraduate project plus previous work in industry	Cycling backpack
7	France	Female	Postgraduate project plus previous work in industry	Architectural outdoor space
8	Netherlands/Brazil	Female	Postgraduate project plus previous work in industry	Plant sensor
9	US	Male	Professional project - design agency	Packaging
10	Spain	Male	Design competition	Compost bin
11	Brazil	Male	Professional project - independent	Facemask
12	UK	Male	Professional project - in-house	Electric vehicle charger
13	US	Male	Professional project - in-house	Vehicles
14	Canada	Female	Professional project - in-house	Yoga mat

(*continued on next page*)

(*continued*)

	Location	*Sex*	*Type of project talked about*	*Product type*
15	Germany	Female	Professional project - independent	Lamp
16	UK	Male	Professional project - independent	Plastic cup

Appendix B. Transcription symbols (Wiggins, 2017, adapted from Jefferson, 2004)

(.)	A micro-pause around one tenth of a second
(1.2)	A pause or silence, measured in seconds and tenths of seconds
=	Latched talk, where there is no hearable gap between words (can occur within a turn at talk, or between speakers)
::	Stretched sounds in talk; the more colons, the longer the sound, as in rea::lly l::: ong sounds
CAPITALS	Talk that is noticeably louder in contrast to the surrounding talk (sometimes shouting)
Underlined	Emphasised words, or parts of words, are underlined
°	Degree symbols enclose noticeably °quieter° talk, with double degree signs indicating °°whispering°°
> <	'Greater than' and 'less than' symbols enclose talk that is at a faster pace (>speeded-up< talk) than the surrounding talk
< >	'Less than' and 'greater than' symbols enclose talk that is at a slower pace (<slowed down> talk)
↑ ↓	Upward arrows indicate a rising pitch in talk, downward arrows indicate falling pitch
£	British pound sign indicates smiley voice or suppressed laughter
#	Hashtag indicates 'creaky' voice such as when someone is upset.
[ ]	Square brackets indicate the start (and end) of overlapping talk
hh	hhs indicate audible breaths. A dot followed by hs (.h) indicate audible inbreaths; without the dot (as in hh) is an outbreath. Within a word (as in 'ye(h)s'), this indicates laughter while talking ('interpolated laughter'). The more hs, the longer the breath.
Huh/heh/hah	Laughter can be represented with outbreaths that have vowel sounds within them.
'yes'	Single quotation marks are used to indicate reported speech or thought
(( ))	Double brackets (sometimes without italics) contain details about other features that have not been transcribed, e.g., *((waves hand))*
(Unclear)	Words in single brackets are the transcriber's best guess at what was being said, or (unclear) or (inaudible) if it really can't be heard clearly

References

Akin, Ö., & Lin, C. (1995). Design protocol data and novel design decisions. *Design Studies, 16*(2), 211–236. https://doi.org/10.1016/0142-694X(94)00010-B.

Ball, L. J., Lambell, N. J., Reed, S. E., & Reid, F. J. (2001). The exploration of solution options in design: A 'Naturalistic Decision Making' perspective. In *Designing in Context* (pp. 79–93). Delft University Press.

Bouyssou, D., Dubois, D., Pirlot, M., & Prade, H. (2009). *Decision-making process: Concepts and methods*. London: ISTE. https://doi.org/10.1002/9780470611876.

Buchanan, R. (1992). Wicked problems in design thinking. *Design Issues, 8*(2), 5–21. https://doi.org/10.2307/1511637, JSTOR.

Buhl, A., Schmidt-Keilich, M., Muster, V., Blazejewski, S., Schrader, U., Harrach, C., Schäfer, M., & Süßbauer, E. (2019). Design thinking for sustainability: Why and how design thinking can foster sustainability-oriented innovation development. *Journal of Cleaner Production, 231*, 1248–1257. https://doi.org/10.1016/j.jclepro.2019.05.259.

Calabretta, G., Gemser, G., & Wijnberg, N. M. (2017). The interplay between intuition and rationality in strategic decision making: A paradox perspective. *Organization Studies, 38*(3–4), 365–401. https://doi.org/10.1177/0170840616655483.

Chick, A., & Micklethwaite, P. (2011). *Design for Sustainable Change: How Design and Designers Can Drive the Sustainability Agenda*. AVA Publishing.

Cooper, L. (2021 Aug 26). [Data Collection]. *Design Decision-making for Sustainability, 2020*. UK Data Service. https://doi.org/10.5255/UKDA-SN-855100.

Cooper, N., & Burnett, S. (2006). Using discursive reflexivity to enhance the qualitative research process: An example from accounts of teenage conception. *Qualitative Social Work, 5*(1), 111–129. https://doi.org/10.1177/1473325006061541.

Cross, N. (1997). Descriptive models of creative design: Application to an example. *Design Studies, 18*(4), 427–440. https://doi.org/10.1016/S0142-694X(97)00010-0.

Cross, N. (2001a). Chapter 5 - design cognition: Results from protocol and other empirical studies of design activity. In C. M. Eastman, W. M. McCracken, & W. C. Newstetter (Eds.), *Design knowing and learning: Cognition in design education* (pp. 79–103). Elsevier Science. https://doi.org/10.1016/B978-008043868-9/50005-X.

Cross, N. (2001b). Designerly ways of knowing: Design discipline versus design science. *Design Issues, 17*(3), 49–55. https://doi.org/10.1162/074793601750357196.

Devon, R., & van de Poel, I. (2004). Design ethics: The social ethics paradigm. *International Journal of Engineering Education, 20*(3), 451–469.

Dorst, K. (1995). Analysing design activity: New directions in protocol analysis. *Design Studies, 16*(2), 139–142. https://doi.org/10.1016/0142-694X(94)00005-X.

Dorst, K., & Dijkhuis, J. (1995). Comparing paradigms for describing design activity. *Design Studies, 16*(2), 261–274. https://doi.org/10.1016/0142-694X(94)00012-3.

Edwards, D. (1999). Emotion discourse. *Culture & Psychology, 5*(3), 271–291. https://doi.org/10.1177/1354067X9953001.

Edwards, D. (2000). Extreme case formulations: Softeners, investment, and doing nonliteral. *Research on Language and Social Interaction, 33*(4), 347–373. https://doi.org/10.1207/S15327973RLSI3304_01.

Ellen MacArthur Foundation and IDEO. (2018). *Material Selection*. Circular Design Guide. https://www.circulardesignguide.com/post/material-selection.

Faber, M. H., & Rackwitz, R. (2004). Sustainable decision making in civil engineering. *Structural Engineering International, 14*(3), 237–242. https://doi.org/10.2749/101686604777963919.

Fairs, M. (2019). Ellen MacArthur calls on designers to 'help regenerate the planet' by adopting circular design. *Dezeen*. https://www.dezeen.com/2019/06/11/ellen-macarthur-circular-economy-designers/.

Fry, T. (2004). The voice of sustainment: Design ethics as futuring. *Design Philosophy Papers, 2*(2), 145–156. https://doi.org/10.2752/144871304X13966215068038.

Fuente, A. de la, Armengou, J., Pons, O., & Aguado, A. (2017). Multi-criteria decision-making model for assessing the sustainability index of wind-turbine support systems: Application to a new precast concrete alternative. *Journal of Civil Engineering and Management, 23*(2), 194–203. https://doi.org/10.3846/13923730.2015.1023347.

Geissdoerfer, M., Bocken, N. M. P., & Hultink, E. J. (2016). Design thinking to enhance the sustainable business modelling process – A workshop based on a value mapping process. *Journal of Cleaner Production, 135*, 1218–1232. https://doi.org/10.1016/j.jclepro.2016.07.020.

Gould, R. K., Bratt, C., Lagun Mesquita, P., & Broman, G. I. (2019). Integrating sustainable development and design-thinking-based product design. In A. H. Hu, M. Matsumoto, T. C. Kuo, & S. Smith (Eds.), *Technologies and eco-innovation towards sustainability I: Eco design of products and services* (pp. 245–259). Springer. https://doi.org/10.1007/978-981-13-1181-9_19.

Gumienny, R., Meinel, C., Gericke, L., Quasthoff, M., LoBue, P., & Willems, C. (2011). Tele-Board: Enabling Efficient Collaboration. In C. Meinel, L. Leifer, & H. Plattner (Eds.), *Design Thinking* (pp. 47–164). Springer. https://doi.org/10.1007/978-3-642-13757-0_9.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological review, 108*(4), 814–834. https://doi.org/10.1037/0033-295x.108.4.814.

Hatamura, Y. (Ed.). (2006). What are Decisions? *Decision-Making in Engineering Design: Theory and Practice* (pp. 1–11). Springer. https://doi.org/10.1007/1-84628-261-6_1.

Haug, A. (2017). Educating ethical designers. *International Journal of Technology and Design Education, 27*(4), 655–665. https://doi.org/10.1007/s10798-016-9367-y.

Hazelrigg, G. A. (1998). A framework for decision-based engineering design. *Journal of Mechanical Design, 120*(4), 653–658. https://doi.org/10.1115/1.2829328.

Jacobsen, P. H., Harty, C., & Tryggestad, K. (2016). (Re)constructing the wicked problem through the visual and the verbal: The case of a dialogue based architectural competition. *Proceedings of ARCOM 32nd Annual Conference*. http://centaur.reading.ac.uk/69312/.

Jefferson, G. (2004). Glossary of Transcript Symbols with an Introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 13–31). John Benjamins.

Jin, Y., & Danesh, M. R. (2006). Value Aggregation for Collaborative Design Decision-Making. In K. E. Lewis, W. Chen, & L. C. Schmidt (Eds.), *Decision*

Making in Engineering Design (pp. 291–299). ASME Press. https://doi.org/10.1115/1.802469.ch24.
Kalantari, B. (2010). Herbert A. Simon on making decisions: Enduring insights and bounded rationality. *Journal of Management History, 16*(4), 509–520. https://doi.org/10.1108/17511341011073988.
Khatib, J. (2016). *Sustainability of Construction Materials*. Woodhead Publishing.
Kiker, G. A., Bridges, T. S., Varghese, A., Seager, T. P., & Linkov, I. (2005). Application of multicriteria decision analysis in environmental decision making. *Integrated Environmental Assessment and Management, 1*(2), 95–108. https://doi.org/10.1897/IEAM_2004a-015.1.
Le Dantec, C. A., & Do, E. Y.-L. (2009). The mechanisms of value transfer in design meetings. *Design Studies, 30*(2), 119–137. https://doi.org/10.1016/j.destud.2008.12.002.
Liedtka, J. (2013). *Design thinking: What it is and why it works*. ([Design at Darden Working Paper Series]).
Liedtka, J. (2015). Perspective: Linking design thinking with innovation outcomes through cognitive bias reduction. *Journal of Product Innovation Management, 32*(6), 925–938. https://doi.org/10.1111/jpim.12163.
Lloyd, P., Lawson, B., & Scott, P. (1995). Can concurrent verbalization reveal design cognition? *Design Studies, 16*(2), 237–259. https://doi.org/10.1016/0142-694X(94)00011-2.
Lloyd, P., & Busby, J. (2003). "Things that went well — No serious injuries or deaths": Ethical reasoning in a normal engineering design process. *Science and Engineering Ethics, 9*(4), 503–516. https://doi.org/10.1007/s11948-003-0047-4.
Luck, R. (2009). 'Does this compromise your design?' Interactionally producing a design concept in talk. *CoDesign, 5*(1), 21–34. https://doi.org/10.1080/15710880802492896.
Luck, R. (2012). 'Doing designing': On the practical analysis of design in practice. *Design Studies, 33*(6), 521–529. https://doi.org/10.1016/j.destud.2012.11.002.
MacAskill, K., & Guthrie, P. (2013). Risk-based approaches to sustainability in civil engineering. *Proceedings of the Institution of Civil Engineers - Engineering Sustainability, 166*(4), 181–190. https://doi.org/10.1680/ensu.12.00001.
McDonnell, J. A., & Lloyd, P. (2009). *About: Designing analysing design meetings*.
Oak, A. (2011). What can talk tell us about design?: Analyzing conversation to understand practice. *Design Studies, 32*(3), 211–234. https://doi.org/10.1016/j.destud.2010.11.003.
Oxman, R. (1995). Viewpoint Observing the observers: Research issues in analysing design activity. *Design Studies, 16*(2), 275–283. https://doi.org/10.1016/0142-694X(94)00006-Y.
Pedgley, O. (2009). Influence of stakeholders on industrial design materials and manufacturing selection. *International Journal of Design, 3*(1), 1–15.
Plessner, H., Betsch, C., & Betsch, T. (2011). *Intuition in Judgment and Decision Making*. Psychology Press.
Pomerantz, A. (1986). Extreme case formulations: A way of legitimizing claims. *Human Studies, 9*(2/3), 219–229.
Potter, J., & Wetherell, M. (1987). *Discourse and social psychology: Beyond attitudes and behaviour*. SAGE.
Pritchard, O. (2013). Designers have an obligation to make sure their work is truly sustainable. *The Guardian*. http://www.theguardian.com/sustainable-business/sustainable-design-sustainability.
Rapley, T. (2012). The (Extra)Ordinary practices of qualitative interviewing. In J. Gubrium, J. Holstein, A. Marvasti, & K. McKinney (Eds.), *The SAGE*

handbook of interview research: The complexity of the craft (pp. 541–554). SAGE Publications, Inc. https://doi.org/10.4135/9781452218403.n38.

Rapley, T. (2015). Questions of context: Qualitative interviews as a source of knowledge. In C. Tileaga, & E. Stokoe (Eds.), *Discursive psychology*. Routledge.

Ryd, N. (2004). The design brief as carrier of client information during the construction process. *Design Studies, 25*(3), 231–249. https://doi.org/10.1016/j.destud.2003.10.003.

Schöggl, J.-P., Baumgartner, R. J., & Hofer, D. (2017). Improving sustainability performance in early phases of product design: A checklist for sustainable product development tested in the automotive industry. *Journal of Cleaner Production, 140*, 1602–1617. https://doi.org/10.1016/j.jclepro.2016.09.195.

Schon, D. A. (2008). *The Reflective Practitioner: How Professionals Think In Action*. Basic Books.

Silva, N. D., Jawahir, I. S. O. D., Jr., & Russell, M. (2009). A new comprehensive methodology for the evaluation of product sustainability at the design and development stage of consumer electronic products. *International Journal of Sustainable Manufacturing, 1*(3), 251. https://doi.org/10.1504/IJSM.2009.023973.

Smith, R. L., & Ruiz-mercado, G. J. (2014). A method for decision making using sustainability indicators. *Clean Technologies and Environmental Policy; Berlin, 16*(4), 749–755. https://doi.org/10.1007/s10098-013-0684-5.

Sridhar, I. (2007). Materials Selection for Green Design. In *Environmentally Conscious Mechanical Design* (pp. 319–350). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470168202.ch9.

Steele, K., Carmel, Y., Cross, J., & Wilcox, C. (2009). Uses and Misuses of Multicriteria Decision Analysis (MCDA) in environmental decision making. *Risk Analysis, 29*(1), 26–33. https://doi.org/10.1111/j.1539-6924.2008.01130.x.

Surma-aho, A., Chen, C., Hölttä-Otto, K., & Yang, M. (2019). Antecedents and Outcomes of Designer Empathy: A Retrospective Interview Study. ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. https://doi.org/10.1115/DETC2019-97483.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458. https://doi.org/10.1126/science.7455683.

van de Poel, I. (2015). Conflicting Values: Value conflict in Design for Values. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), *Handbook of Ethics, Values, and Technological Design* (pp. 89–116). Netherlands: Springer. https://doi.org/10.1007/978-94-007-6970-0_5.

van de Poel, I., & Verbeek, P.-P. (2006). Editorial: Ethics and engineering design. *Science, Technology & Human Values, 31*(3), 223–236. https://doi.org/10.1177/0162243905285838.

Wiggins, S. (2017). *Discursive psychology: Theory, method and applications*. SAGE. https://doi.org/10.4135/9781473983335.

Woodhouse, E., & Patton, J. W. (2004). Introduction: Design by society: Science and technology studies and the social shaping of design. *Design Issues, 20*(3), 1–12.

Wooffitt, R., & Widdicombe, S. (2006). Interaction in Interviews. In P. Drew, G. Raymond, & D. Weinberg (Eds.), *Talk and Interaction in Social Research Methods* (1st ed., p. 22). SAGE.

Zannier, C., Chiasson, M., & Maurer, F. (2007). A model of design decision making based on empirical results of interviews with software designers. *Information and Software Technology, 49*(6), 637–653. https://doi.org/10.1016/j.infsof.2007.02.010.

Auditing design justice: The impact of social movements on design pedagogy at a technology institution

Madhurima Das[1], *Anastasia K. Ostrowski*[1], *Shelly Ben-David*[1] and *Gillian J. Roeder*[1], Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Kimberley Kimura, Wellesley College, 106 Central St, Wellesley, MA 02481, USA

Catherine D'Ignazio and *Cynthia Breazeal*, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Aditi Verma, Harvard Kennedy School, University of Michigan, 2355 Bonisteel Boulevard, Ann Arbor, MI 48109, USA

Engineering aims to solve consequential, real-world problems. To be effective problem-solvers in societal contexts, engineers and designers must be trained to conceptualize and operationalize ethics, equity, and justice in their practice. Our work develops a methodology — based on the Design Justice framework — for ethics, equity, and justice audits of design pedagogy. We apply this methodology to audit design and design critique courses at a technology institution — the Massachusetts Institute of Technology. Through this audit we identify exemplar courses as well as ways in which courses fail to engage with ethics, equity, and justice. This audit, the largest of its kind, is a proof of concept of how the audit methodology can be applied to design education and practice.

Keywords: design education, engineering design, research methods, inclusive design, design justice

Corresponding author:
Madhurima Das
rimadas@mit.edu

Engineering design pedagogy largely focuses on building scientific, engineering, and design expertise in future engineers (Cross & Cross, 1998; Magney & Bucciarelli, 1995). This focus on analytical and technical rigor, while essential, does not generally take into consideration the vital and complex role of social and environmental contexts in the design of technology (Rittel & Webber, 1973), and imparts an incomplete understanding to future designers and technology inventors of the role and purpose of the engineer, and engineered artifacts and systems, in society (Verma & Djokić,

www.elsevier.com/locate/destud
0142-694X *Design Studies* **86** (2023) 101183
https://doi.org/10.1016/j.destud.2023.101183

2021). Given that the shared purpose of engineers and of engineering is to solve consequential, real-world problems, our research project takes as its starting point the need to train engineers to conceptualize and operationalize ethics, equity, and justice in their work, especially their design practice. To that end, through this paper, we explore whether engineering design pedagogy equips engineers with the necessary intellectual frameworks to consider questions of ethics, equity, and justice. Engineering education and design researchers generally agree that pedagogical innovations are needed in order to ensure that current and future technologies are imagined, designed, built, managed, and disposed of in equitable and just ways (Costanza-Chock, 2020; Gallimore, 2021; Riley, 2008). In this paper, we audit design pedagogy and the inclusion of ethics, equity, and justice therein to provide a baseline from which to commence pedagogical innovation and renewal to better prepare engineers of the future. In our work, we draw and build on the pivotal work of Sasha Costanza-Chock (2020) which proposes design justice as a "framework for analysis of how design distributes benefits and burdens between various groups of people." Design justice researchers, such as Costanza-Chock (2020) and (D'Ignazio & Klein, 2020), and practitioners urge designers to consider the ways in which design can perpetuate or challenge systems of oppression and discrimination in contemporary societies (Collins, 1990; Costanza-Chock, 2020). Doing so requires that designers explicitly consider design justice principles such as: which values are encoded in design work, who is paid to do design (and recognized as a designer), where design is done, and who the design work is done with and for (Costanza-Chock, 2018). The focus of our work is how the normative logics underpinning design are rationalized and perpetuated through pedagogy (Costanza-Chock, 2018), particularly in light of the recent movements centering racial, social, and environmental justice and the resulting unprecedented amplification of justice in the social consciousness in the United States starting in 2020. We explore the impact of these rapidly mainstreaming social movements on design pedagogy at a technology institution. Our audit of design pedagogy at our shared home institution is carried out in the spirit of constructive critique and transparency, as a call to design instructors to reflect on the ways in which we can better train engineers of the future to serve society. The overarching aim of our research endeavor is to inform and stimulate the creation of design pedagogy that creates engineers who, through their design practice, aspire to advance equity and justice. We make the four following contributions through our paper: first, we extend the design justice framework for assessing design pedagogy; second, we explore which elements of the design justice framework are embedded in design pedagogy and how they engage with ethics, equity, and justice; third, we demonstrate a proof of concept design justice syllabus audit methodology that can be used at institutions of higher education; and fourth, we investigate what impact, if any, the social movements of 2020 have had on design pedagogy between

Fall 2019 and Fall 2020 at the institution. Such periodic audits can be used to understand, measure, and orchestrate systemic and institutional shifts towards educating and training future designers and engineers who are better able to engage with ethics and justice. The first ever design justice audit of design pedagogy at this technology institution carried out in our work creates a baseline for comparing future pedagogical improvements and innovations.

1 Related Works

Design justice is a framework of analysis as well as a community of practice that "ensure[s] a more equitable distribution of design's benefits and burdens; meaningful participation in design decisions; and recognition of community-based, Indigenous, and diasporic design traditions, knowledge, and practices" (Costanza-Chock, 2018). A design justice analysis of technological design examines whether a particular technology challenges or reinforces existing and systemic forms of oppression and suppression or the matrix of domination (Collins, 1990). Costanza-Chock's book on design justice proposes seven questions to critique existing technologies and institutional systems in which they operate, or to consider as part of the practice of designing new technologies. These questions include: (1) who gets to do design and whose work is recognized as design (*Equity*), (2) what users and communities do we design for and with (*Beneficiaries*), (3) what values are embedded implicitly or explicitly in technological artifacts and systems (*Values*), (4) how do we scope and frame design problems (*Scope*), (5) where is design work done and how does the location of the design work impact which sites are privileged whereas others are marginalized or ignored (*Sites*), (6) who receives the benefits of design work and how can the work be owned by communities instead of individuals (*Ownership, Accountability, & Political Economy*), (7) how do we rationalize and remember how and why technologies are designed as they are (*Discourse*) (Costanza-Chock, 2018). In order to add greater granularity to our analysis and pay particular attention to discussion of past harms of technology, we separated "Discourse" into two distinct categories of "Discourse" and "Histories."

Universalist, standardized, and one-size-fits-all approaches to design are too often prescribed as part of engineering pedagogy. These methods overlook how the potential benefits, burdens, and harms created by technologies are distributed on the basis of race, ethnicity, class, gender, disability, sexuality, and other markers of identity. While prior design epistemologies and paradigms—including (but not limited to) value sensitive design (Friedman, 1996), values in design (Knobel & Bowker, 2011), resource-constrained design (Anderson et al., 2012), human-centered design (Buchanan, 2001) and others—have sought to center users and communities, but none have the explicit attention towards intersectional equity as part of design work in the way that design justice does.

Impact of social movements on design pedagogy

Examples of well-intentioned but flawed engineering efforts aimed towards development abound in engineering practice and pedagogy. In a critique of such engineering-for-development initiatives, Nieusma and Riley (2010) observe that these engineering initiatives often make inaccurate and problematic assumptions about the role technology can and should play in development – be it in the West or in the Global South. Technology-for-development efforts also frequently ignore the power relations they create or perpetuate and forsake meaningful community engagement. This ultimately privileges technical performance and functionality as an end rather than a means towards achieving development. Similarly, Schneider et al. (2008) critique the colonial undertones of university-led engineering development projects through which students from Global North universities seek to launch development initiatives in the Global South. Here too, because of the limitations of their own training and the constraints of the development initiative as undertaken in a pedagogical context, student engagement with communities is fleeting and lacking in depth. Communities are designed *for*, and not *with*, as advocated by the design justice framework.

These efforts by engineering educators and students to engage in development work can be situated in a broader movement within engineering that Mitcham and Munoz (Mitcham & Munoz, 2010) describe as *humanitarian engineering*. They write that humanitarian engineering can be described "as working to escape what has been called the 'social captivity of engineering' by capitalism or nationalism or some other form of wealth and power" (Goldman, 1991; Mitcham & Munoz, 2010). Indeed, a pursuit of humanitarian engineering, or more broadly, humanitarian design, calls not only for an examination of our current design curricula but also reckoning with problematic practices taught to prior generations of engineers which continue to inform our design current practice and pedagogy (Lucena & Schneider, 2008). Understanding and addressing the limitations of engineering pedagogy specifically and design pedagogy broadly requires that we examine it through new conceptual lenses that bring social scientific and humanist ways of knowing to bear on the role of science and technology in society (Verma, 2021b). Pritchard and Baillie (Pritchard & Baillie, 2007) carry out such an analysis through a survey of Science and Technology Studies (STS) faculty and identify participation, politics, and citizenship as key analytical themes. Our paper builds on this work and adds additional analytical themes from the design justice framework—such as Values, Scope, Discourse, and Histories—for analyzing design pedagogy.

While Costanza-Chock's Design Justice scholarship (Costanza-Chock, 2018, 2020) was published relatively recently, it coalesces under one framework of ethics, equity, and justice considerations that have long been foregrounded by engineering ethics and engineering education researchers (Baillie et al., 2012; Riley, 2008). For this reason, we believe that a 2019 and 2020 audit of design pedagogy using this framework is apt. Our work marks the first

scholarly effort to extend the design justice framework for assessing design pedagogy. While no previous work has explored how design justice is embedded in design education, prior studies have developed methods of reviewing course syllabi to find patterns concerning how particular topics, such as ethics, are covered in curricula. Syllabus analysis is an established, useful method of identifying areas of emphasis in curricula (Chong, 2015). Fiesler et al. (2020) conducted an analysis of syllabi with a focus on ethics curricula in computing education. They investigated whether ethics courses were typically standalone or if the topics were being integrated into core computing curricula. Through their analysis of 115 syllabi, they found that there is a myriad of ways in which instructors are engaging with ethics topics in their curricula and with a variation in the depth of these engagements. Their recommendations for integrating ethics content into computing courses include emphasizing, starting in introductory courses, the key principle that even a small design artifact like code can have social consequences so that students understand the responsibility of working in computing early on in their education. We conducted an audit of design course syllabi using a similar overall methodology to investigate the ways in which instructors engage with ethics, equity, and justice in design pedagogy in both engineering as well as non-engineering departments of a technology institution. A key difference in our methodology from that of Fiesler et al. (2020) was that our focus was on syllabi available within a single institution. Additionally, our dataset was built by requesting syllabi from professors teaching design courses, whereas Fiesler et al. (2020) used syllabi shared by a group who self-identified their courses as relevant. The design justice audit was guided by the following research questions in our study:

RQ1: How, if at all, are design classes engaging with equity, justice, and ethics considerations?

We hypothesize that most design classes are not directly engaging with ethics, equity, and justice considerations. We expect to see wide variation across classes and departments in their levels of engagement with these topics: from no presence to integrating ethics, equity, and justice as main focuses of a course.

RQ2: Which design justice paradigms are implicitly or explicitly embedded in design pedagogy?

We hypothesize that each department will have different design justice paradigms embedded in their pedagogy based on the parts of design they have traditionally engaged with. For instance, Mechanical Engineering courses may include more Beneficiaries since engagement with users is a common part, especially of product design courses in the field. Similarly, Nuclear Science and Engineering courses may include more Histories since some design courses involve critique of historical failures of nuclear technologies.

Impact of social movements on design pedagogy

RQ3: What impact, if any, have the racial, social, and environmental justice movements of 2020 had on design pedagogy?

We hypothesize that the *racial, social, and environmental justice* movements of 2020 will have minimal impacts on design pedagogy. We expect that due to the COVID-19 pandemic, courses may incorporate more practices that indicate social mindedness (such as extension policies, explicit disability accommodations, etc.) but we do not expect this to permeate into pedagogy.

2 Methodology

To answer our research questions, we analyzed the design courses and the design principles (Fu et al., 2016) embedded therein at a technology institution through a novel design justice centered syllabus auditing methodology that we developed.

2.1 Syllabus auditing process

Our initial approach included performing a keyword search of "design" in the course catalog for Fall 2020 after the racial, social, and environmental justice movements of winter, spring, and summer 2020. However, initial results (1006 courses for Fall 2020) did not comprehensively encompass design pedagogy at the institution. Recognizing the importance of design and critique as part of a reflective design practice (Bardzell, 2010; Verma, 2021a), we were interested in identifying not only design courses but also courses which taught the students to think critically *of* design. In our analysis, we refer to these as "critique" courses. In order to capture design justice considerations as they appear in both design and critique courses, we extended the design justice framework and developed elaborations of the design justice questions as they apply in the context of a design and critique course respectively. These elaborations of the design justice questions in design and critique pedagogical contexts, which we refer to as the design justice and pedagogy framework, are shown in Table 1.

Though we initially started with a keyword search for "design" in course descriptions, in order to capture the broadest possible range of design and critique courses, we developed a more rigorous approach that consisted of reviewing course listings across six departments. The departments chosen for our analysis are representative of our (authors') expertise and are known for their emphasis on design at the Massachusetts Institute of Technology (MIT): Mechanical Engineering (MechE), Electrical Engineering & Computer Science (EECS), Nuclear Science & Engineering (NSE), Architecture, Urban Studies & Planning (DUSP), and Media Arts & Sciences (MAS). Across these departments, we logged courses for Fall 2019 and Fall 2020, collecting the course descriptions, syllabi, and any additional course materials available. Additionally, we identified courses taught in both Fall 2019 and Fall 2020 for comparison.

Table 1 Design justice questions: Examples in the context of design and critique

Question	*Design example*	*Critique example*
Equity: Who gets to do design?	The course considers how identity and background (race, class, ethnicity, gender, disability, and sexuality) shape who is able to do design and be recognized as a designer, or the course explicitly treats users and communities potentially impacted by a technology as its designers.	The course critiques how identity and background shape who is able to do design and be recognized as a designer. The syllabus critiques design practices that do or do not view users and impacted communities as co-designers.
Beneficiaries: Who do we design for or with?	The course considers how designers identify users and whether user identity and background impact who is considered as a potential user for a new design and whose preferences and needs are accounted for as part of the design work.	The course includes theoretical or methodological resources or itself critiques design processes that give differential access to individuals (who is designed for and with), or individual preferences, based on their identity and background.
Values: What values do we encode and reproduce in the objects and systems that we design?	The course teaches student designers to reflect on their own biases, assumptions, and values and how these might become embedded in the artifacts or systems they design.	The course includes theoretical or methodological resources that students can use to examine or critique the values that are implicitly or explicitly encoded in technologies, and/or the course itself includes this critique.
Scope: How do we scope and frame design problems?	The course teaches students to pay attention to the ways in which design problems are framed and the extent to which justice, equity, and ethics considerations shape those framings.	The course includes theoretical or methodological resources that students may use to examine and critique the framing of design problems—focusing on whether those framings are attentive to justice, equity, and ethics,—and/or the course itself includes this critique.
Sites: Where do we do design? What design sites are privileged? Which sides are ignored or marginalized? How do we make design sites accessible to those who will be most impacted?	The course asks students to consider as part of their design work how their own design site may be privileged whereas others are overlooked or marginalized.	The course includes theoretical and methodological resources that enable the students to critically examine and interpret how the site of the design work shapes the designed artifact, and/or the course itself includes this critique.
Ownership, Accountability, & Political Economy: Who owns and profits from design outcomes? What social relationships are reproduced by design? How do we move towards community control of design processes?	The course asks students to reflect on who may ultimately own the artifact or system being designed, what social and power relationships the designed artifact may either reproduced or create anew, or how the ownership of the artifact or system being designed can, over its full lifecycle, be held by communities instead of by individuals.	The course includes theoretical or methodological resources that enable students to critique the designed artifact or system and its maintenance, by examining its ownership structures and social and power relationships created or reproduced by the design, and/or the course itself includes this critique.

(*continued on next page*)

Impact of social movements on design pedagogy

Table 1 (*continued*)

Question	*Design example*	*Critique example*
Discourse: What stories do we tell about how things are designed?	The course asks students to consider how the technology or artifact being designed may be interpreted and/or rationalized, and it's raison d'être described by future users.	The course includes theoretical and methodological resources that enable students to critically examine how and why prior or current technologies are interpreted, rationalized and the raison d'être described by past, current or future users, and/or the course itself includes this critique.
Histories: Acknowledging unequal histories and/or historical harms arising from technology design, use, or diffusion.	The course asks students to consider the historical harms that may have arisen from prior versions of the technology being designed.	The course includes theoretical and methodological resources that students can use to examine the historical harms that have arisen current, prior, or future technologies, and/or the course itself includes this critique.

We iteratively identified and agreed (two coders per course description) upon which courses were design and/or critique courses that included aspects of ethics, equity, or justice to include for further analysis. We engaged in this form of "purposeful" sampling (Welch & Patton, 1992) in our qualitative research approach to ensure we captured the variation of design courses within the institution and to test developing ideas of the presence of Design Justice. This approach in turn helped create a valid dataset with the potential for generalizability and expansion to other pedagogical contexts (Maxwell, 1992). This detailed approach enabled us to gather all course syllabi that had a design or critique component across the six departments. We gathered 121 course syllabi from Fall 2019 and 119 syllabi from Fall 2020—a total of 240 unique course syllabi. Of these, 65 were courses offered in both semesters and thus could be directly compared. Table 2 depicts the course distribution across departments and the corresponding number of syllabi that were analyzed.

2.2 Analysis

We analyzed all course descriptions and syllabi to identify the inclusion of ethics, equity, and/or justice content within design and critique courses.

Table 2 Number of total classes in each department of interest in Fall 2019 and Fall 2020 along with number of syllabi requested and number of syllabi analyzed for each department of interest

	Fall 2020			*Fall 2019*		
	Total number of classes	*Number of syllabi requested*	*Number of syllabi analyzed*	*Total number of classes*	*Number of syllabi requested*	*Number of syllabi analyzed*
MechE	130	89	34	128	83	28
EECS	161	48	26	152	46	22
NSE	41	9	6	38	6	5
MAS	22	16	17	20	14	11
Architecture	88	52	17	84	61	26
DUSP	64	34	19	65	33	29
Total	**506**	**248**	**119**	**487**	**243**	**121**

Each syllabus was analyzed comparatively by two coders using an iteratively developed rubric, similar to established rubric development processes (Das & Yang, 2021). A draft of the rubric was created after the research team conducted a preliminary review of the dataset across the departments to gain an initial understanding of syllabi organization and content styles. The rubric was then iterated upon four times and piloted on a subset of the syllabus dataset including nine syllabi gathered from across four of the departments (NSE, MechE MAS, and EECS). The rubric is composed of four sections: (1) Introduction, (2) Course Description from Course Catalog, (3) Syllabus, and (4) Summary.

The Introduction section of the rubric gathers background information (i.e., year course was taught, name of the course, department, etc). This section also asks about whether the course syllabus features any statements related to justice, ethics, and equity, including land acknowledgments (statements recognizing Indigenous Peoples as traditional stewards of the land) and statements concerning mental health and disability accommodations. The Course Description section uses a Python script to read the course description and record the presence of 26 terms that relate to design justice themes (e.g. "stakeholder," "participatory," "inclusive," "intersectionality"), gathered from our rubric development process. The full list of these terms is provided in Appendix A. These are some of the themes that were considered to be relevant to the design justice questions without explicitly using the language in the design justice questions.

The Syllabus section's questions examine the implicit or explicit presence of the design justice questions and design paradigms included in the syllabus (e.g. human-centered design, complex system design, value centered design, etc.). The Summary section of the rubric records whether external community partners are involved in the course, the core learning methods employed, whether and how the design course is designing and/or critiquing a design topic and/or object, and an overview of whether and how design justice principles are explicitly/implicitly and meaningfully/superficially incorporated into the course. We define "implicitly" as a consideration of design justice questions but without labeling them as such and without any inclusion of theoretical and methodological resources. This may include holistic discussions of themes such as inclusive design or universal design that are related to design justice that are not captured in the design justice questions. "Explicitly" is defined as direct mentions and considerations of design justice questions with exposure to theoretical and methodological resources. "Meaningfully" and "superficially" were defined across a range from "no mention in the syllabus" to "a cursory one-liner in the syllabus" (superficially) to the course being "focused on design justice" (meaningfully). The coding using the rubric was completed by two team members for each course syllabus. Each team member coded each syllabus independently using the rubric in the format of a Qualtrics survey.

Impact of social movements on design pedagogy

Researchers discussed the results and resolved any disagreements, similar to Daly et al.'s (2012) coding approaches. After all syllabi were coded through the rubric, we analyzed the results in several stages. The first stage involved aggregating numerical data from each department to track how many syllabi in each department were addressing each design justice question. We also recorded the implicit/explicit engagement levels, meaningful/superficial attention to design justice principles, and where design justice principles and questions were an afterthought vs. where they were thoughtfully integrated in the course syllabus. A final question on the rubric asked the respondent to reflect on opportunities for the inclusion of design justice considerations in the course being analyzed. Responses to this question were used to identify exemplar courses described in Section 3.2. These results were aggregated in the same way as the data from the design justice questions. We also compared the courses between Fall 2019 and Fall 2020 (61 courses) documenting changes in presence of design justice.

The syllabi ranged from course schedules and lists of topics covered in the year to in-depth descriptions of course policies and grading to slides in the presentation from the first class of the semester to pages on websites. We obtained syllabi through consulting the archives of the institution's course management system and working with professors and departmental instructors and administrators. All syllabi that we gathered from the course management system were publicly available to the community. We also reached out directly to professors and departmental administrators—an approach that posed its own challenges. We sent up to two emails requesting syllabi, but only 44% of emails yielded a syllabus of interest. We experienced several challenges and limitations with regards to gathering syllabi including the variety of dispersed locations that syllabi were stored (i.e., course management system, instructors, websites), low availability of syllabi if they were not posted online, only syllabi from one semester being available (often the most recent semester), and incorrect listing of instructors on the course website. A central repository of course syllabi would make it significantly easier to replicate and extend the audit approach in the future as well as in other university contexts.

Another limitation of this work has to do with the descriptiveness and level of detail contained in course syllabi. Course syllabi are intended to provide an overview of the course; however, they may not capture all strategies that instructors use to incorporate design justice into their curriculum. For example, depending on the detail provided in the syllabi, it can be difficult to gauge if assignments engage with design justice material in the course in ways that may be student-directed and otherwise not captured in the course description or syllabus (e.g. projects could be creating a space for students to explore design justice principles that are not reflected in the syllabi). Our approaches for addressing these methodological limitations are described in the Future Work section.

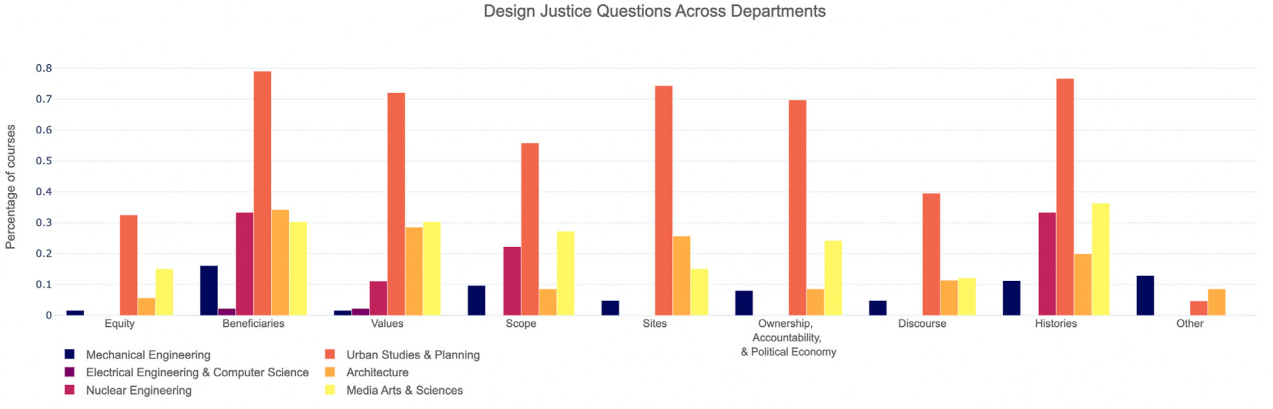

Figure 1 Design justice questions and their inclusion in course syllabi as compared across departments. The "Other" category in our analysis includes topics such as sustainability that were design-justice adjacent but did not fit one of the value categories

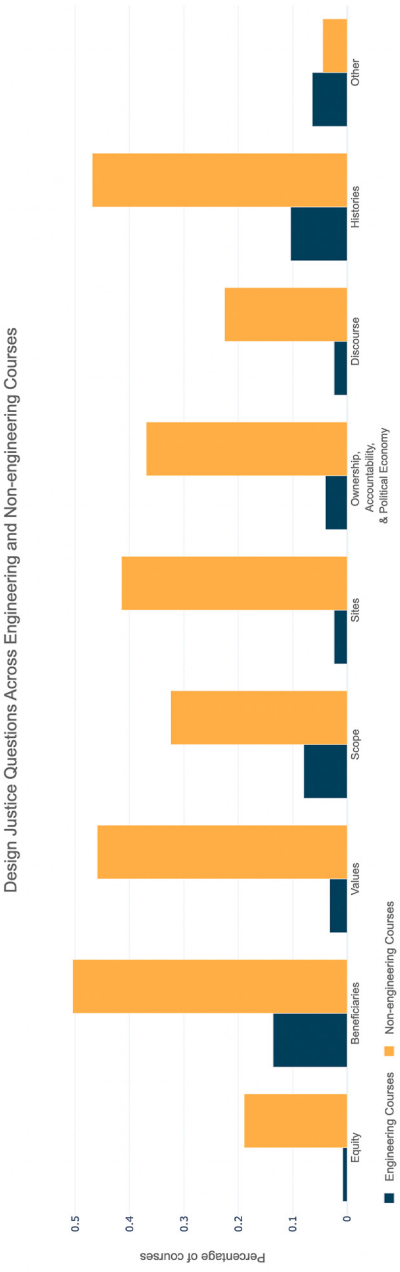

Figure 2 Design justice values compared across courses for engineering and non-engineering departments

Figure 3 Depiction of the implicit/explicit and meaningful/superficial focus given to design justice questions across departments. The marker is located at the average of both the x- and y-axes for each department. The size of the marker is proportional to the middle two quartiles of the level of explicitness of design justice content within a department to better depict the majority of the courses and exclude outliers

3 Results

A total of 505 classes were logged for Fall 2020 across these six departments. Among these, 368 classes were identified as classes of interest (design or other ethics/justice/equity-focused critique) for syllabi gathering. Our analysis of the syllabi from these departments and courses, which forms a baseline against which future curriculum development can be assessed, reveals that design justice is grossly underemphasized in design pedagogy. This is despite increased discussion around diversity, equity, and inclusion in academic spaces, including calls from engineering leadership to center equity in undergraduate curricula (ABET, 2021; Gallimore, 2021), particularly in light of the recent equity and justice centering movements. Coursework continues to be more focused on the design process and design outcomes without detailed consideration of societal and environmental implications of technology development and use. Principles of equity, justice, and engagement with user populations and impacted communities, when present, are more commonly treated as one-off engagements instead of sustained community partnerships, which have been shown to be more impactful (ABET, 2021; , Ostrowski, Breazeal, & Park, 2021; Gallimore, 2021). We observe some positive movement with classes discussing topics such as sustainability in engineering contexts. However, discussions around ethics, justice, and equity are more commonly found in classes that do not have design components or are outside of engineering disciplines, such as in DUSP or MAS. Though the COVID-19 pandemic and the social movements of 2020 served as catalysts for adapting courses for remote learning, there has not yet been a corresponding reform towards embedding design justice values in design courses, though it is possible that such changes are being planned. Though more time is necessary to determine a representative trend, we now have a critical first data point.

3.1 Design justice in department courses

We compared the courses from across departments to understand how various fields were integrating design justice into their courses. We compared the percentage of classes across each department that had any design justice component. As shown in Figure 1, DUSP courses had the highest prevalence of design justice (97.7%). NSE had the next highest (66.7%) followed by Architecture (54.3%), MAS (42.4%), and MechE (29.0%). EECS had by far the least prevalence of design justice themes, with only 4.3% of classes mentioning any design justice-related topic in the syllabus.

As shown in Figure 2, courses in non-engineering departments such as architecture, DUSP, and MAS consistently included design justice considerations at a much higher rate than courses in engineering departments. Courses in engineering departments performed especially poorly along the Equity, Sites, Values, and Discourse dimensions—categories that are no less applicable in an engineering context than a non-engineering one. Figure 3 provides an

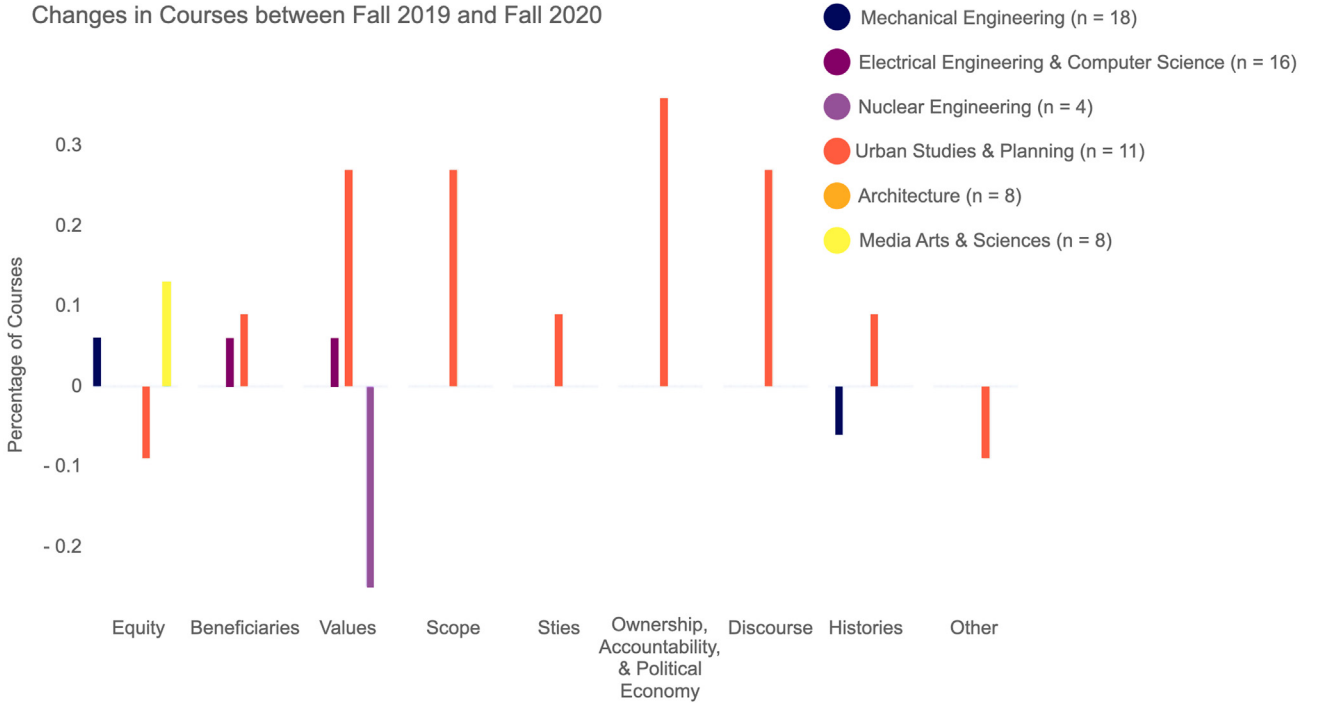

Figure 4 Changes in courses between Fall 2019 and Fall 2020 across the design justice areas. Bars above the zero axis indicate that courses in that department incorporated more design justice areas into their courses. Bars below the zero axis indicate that courses removed areas of design justice from the courses

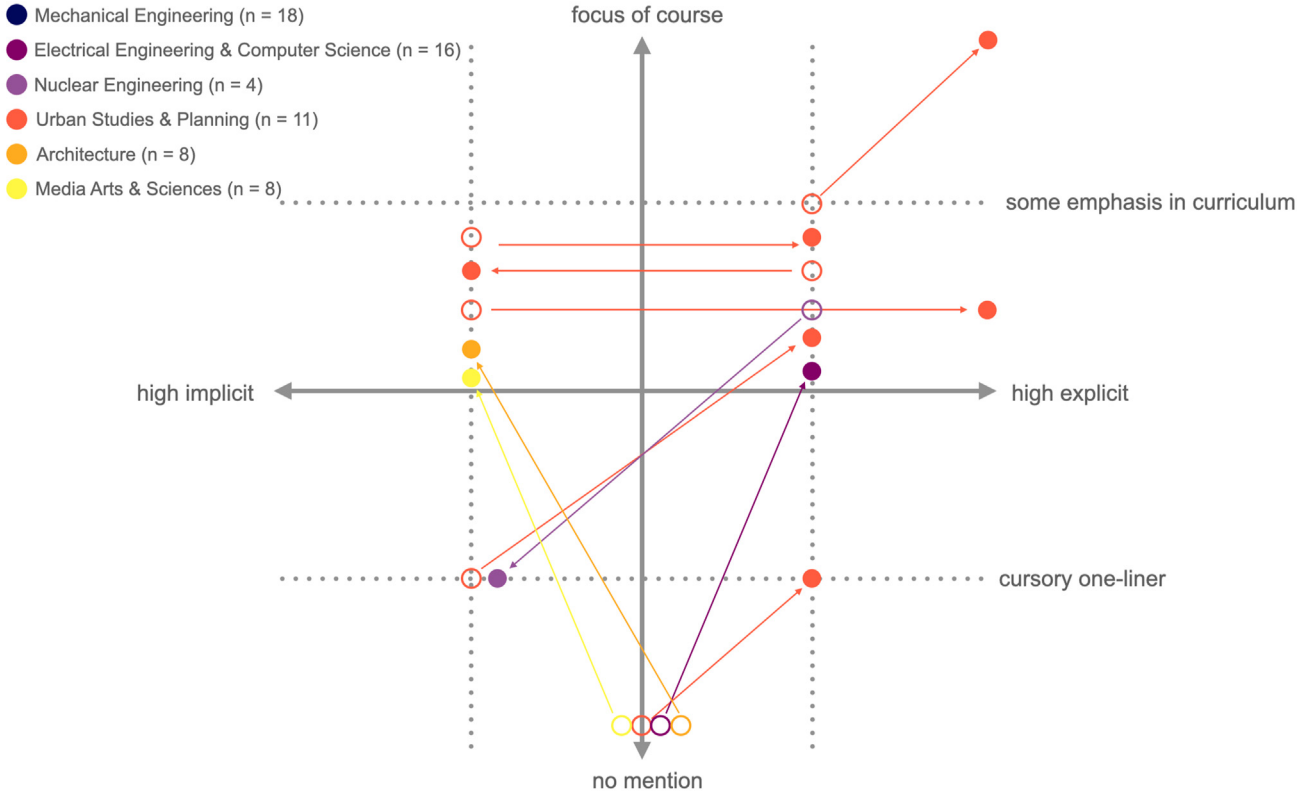

Figure 5 Plot demonstrating how classes shifted emphasis in design justice focus between Fall 2019 and Fall 2020. The open point represents Fall 2019 and the solid point following the arrow represents Fall 2020. The points are color coded by department. No changes were observed in Mechanical Engineering courses (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

additional way to visualize the data along the dimensions of explicit/implicit and meaningful/superficial engagement with design justice. DUSP had the highest prevalence of design justice with most of the courses exhibiting some emphasis on design justice in the curriculum. MechE and EECS had the least presence of design justice and, correspondingly, on average, the principles were not included or not mentioned. MAS, Architecture, and NSE clustered together with having some design justice presence but the overall average was a cursory one-liner in the syllabus without implicit or explicit mention of design justice, though the middle two quartiles demonstrate that there is a small range toward low implicit or explicit mentions in these departments.

3.2 Exemplar courses

Despite the relatively low emphasis on design justice considerations across departments, our syllabus analysis did reveal a number of exemplary pedagogical practices. In this section, we explicitly identify these courses as exemplars that other instructors could emulate in the future.

Design justice and related concepts as main learning objectives. Some courses had topics that were very aligned with design justice with an additional emphasis on technology. For instance, one way that design justice was incorporated into courses was having a focus and main objective of the course related to social justice and incorporating certain technologies into the course to facilitate discussions on how technologies can enable or prevent justice and perpetuate or help fix social inequalities. One course of this type in MAS began with an in-depth examination of racism and social injustice in the United States with readings and discussion focused on these and other justice-related topics. The course then introduced a technical foundation that merged technology applications with social justice. Students produced a proposal and/or research project that they worked on over the course of the semester that combined their interests and learning objectives of the course related to design justice.

Design justice concepts incorporated into technical pedagogy. One type of exemplar was a course that not only emphasized technical rigor, but also centered design justice considerations in the context of tool use and development. For example, a course on Geographic Information System (GIS) mapping offered in DUSP, in addition to emphasizing technical knowledge, also acknowledges the political significance of maps while critiquing associated historically oppressive practices. Specifically, it analyzes the use of maps and seeks to position map-making as a tool of empowerment for communities and activists that enables them to tell their own stories. The course achieves these learning objectives by interspersing technical readings on GIS with readings on power, colonialism, and the politics of maps such that students learn to critique GIS techniques and their potential uses even as they continue to acquire expertise

in the area. This practice of foregrounding social context and interspersing ethics, equity, and justice considerations with technical expertise is a practice that could usefully be emulated in the context of any course similarly built around teaching the use or development of an analytical tool.

Emphasis on community engagement. Some courses incorporated external groups relevant to the design topic into their pedagogy, like clients, experts, and community members. Select courses in sustainability in Mechanical Engineering and DUSP took students on field trips and met with local experts to better understand the leading factors and first-hand insights related to a specific environmental issue. The hands-on engagements that centered around different aspects of the problem—from health (of humans and ecosystems) to economic security—promoted a more well-rounded problem-solving approach. Similarly, a design-for-the-developing-world social impact course in Mechanical Engineering had community partners evaluate the proposed technologies as part of the course's evaluation process of whether or not the device would actually be useful to the community partners. Though these community members were not treated or identified as co-designers, they gave input throughout the design process in a formalized way. By collaborating with those who would be affected by design decisions, these classes in Mechanical Engineering and DUSP allowed students to better understand the social nuances of building technologies and/or designing urban infrastructures.

Just as the courses described above represent exemplary pedagogical practices, several others courses, as depicted in syllabi we analyzed, offer opportunities for learning lessons from prior pedagogical practices that have incorporated ethics, equity, and justice content in design pedagogy with limited success. In general these course fall into one of three categories.

(1) One set of courses tends to over-intellectualize ethics, equity, and justice as conceptual categories through the use of theories and frameworks that are far removed from the lived experiences of communities and individuals who actually experience inequities. This approach to the treatment of ethics, equity, and justice content was especially apparent in departments that deal with the designs of large sociotechnical systems. Because of the size and scope of such systems and their architecture of systems made up of several systems, questions of equity, ethics, and justice are easily overlooked. Instead, they ought to be integrated into every level of scale and technology design (Turner, Verma, & Wood, 2021).
(2) A second set of courses that are a study in how not to teach ethics, equity, and justice are those that seek to develop products or systems for underserved communities, without directly engaging with those communities during any stage of the design process. Such approaches to design, if normalized as part of the pedagogical practice at leading institutions of

higher learning, are likely to be amplified and perpetuated in design practice when students enter the workforce.

(3) A third set of courses are those that treat ethics, equity, and justice in cursory and inconsistent ways across course offerings. Several course syllabi we reviewed included ethics, equity, and justice content but did not show students how to operationalize it in design or critique work.

Overall, exemplar courses demonstrate a variety of ways to integrate design justice concepts into both engineering and non-engineering pedagogy, with several examples of exemplary pedagogical practices transcending specific departments. For example, both engineering and non-engineering departments could benefit from greater community engagement as part of teaching design or critique. Similarly, these departments could benefit equally from including reflections on, and critiques of, the analytical tools being developed in the course. Conversely, several courses whose syllabi we analyzed contain examples of how not to teach ethics, equity, and justice content. While our list of exemplary and non-exemplary pedagogical practices is not exhaustive, we recommend that instructors of design courses reference these examples, the exemplars in particular, to identify ways they might incorporate design justice into their own curricula. Instructors may also wish to consider and reflect on other novel pedagogical approaches that are not yet in practice.

3.3 Changes in courses between Fall 2019 and Fall 2020

The majority of courses (80%) that were taught in both Fall 2019 and Fall 2020 had no change in their inclusion of design justice questions. Figure 4 reveals the positive and negative changes that occurred between semesters with respect to design justice areas. DUSP exhibited the most change between Fall 2019 and Fall 2020; DUSP increased design justice in its syllabi in all areas except Equity. MAS, MechE, EECS, and NSE saw some change, though to a lesser degree than DUSP. Architecture showed no change between semesters.

In addition to comparing the design justice areas between semesters, we also explored how design justice emphasis changed between semesters by analyzing how implicitly or explicitly the design justice questions were addressed in the course and how meaningfully or superficially they were incorporated. Figure 5 depicts the results from 10 courses across the departments that exhibited change along these dimensions. There are three general trends that were identified from this subset of courses: (1) courses trend from being low implicit or no mention at all to explicitly incorporating design justice into the course with more meaningfulness; (2) courses move from having no mention at all to implicitly incorporating design injustice into the course with some curriculum emphasis; and (3) courses move from incorporating design justice explicitly to incorporating it implicitly. For example, courses that moved from low implicit or no mention to low or high explicit reshaped

the course to include design justice or added design justice case studies to complement technical material. Courses that moved from having no mention to implicitly incorporating design justice added limited readings around the topic area without specifically categorizing them as "design justice focused" whereas, courses that moved the opposite direction from explicit to implicit, though few, often removed readings or focus on design justice to prioritize other materials.

The changes we saw were incorporated into courses in a variety of ways. Some courses restructured lectures and curriculum materials to reframe the course more around social justice. Some courses completely redesigned the course to have a more even balance between technical material and case studies on design justice related areas. A few courses also removed course work, such as an inclusive design lecture, that decreased the emphasis of design justice in the course. Overall, the majority of courses held in both Fall 2019 and Fall 2020 did not change with regards to design justice areas or its emphasis. Those that did change generally changed positively, incorporating more design justice into the content and/or increasing emphasis on design justice through the methods described previously.

One of the courses exhibiting positive changes was transformed into what we have called an "exemplar." It is noteworthy that such a transformation occurred after a change of instructor. This suggests that courses lacking design justice content might not falter from lack of potential or irrelevance to course content; it simply takes someone with an eye and intention for justice—a skill that can be learned by engineering and non-engineering course instructors—to redesign curricula. Ultimately, however, even the majority of courses with positive improvement could be considerably further improved through more explicit, meaningful, and intentional inclusion of design justice considerations.

4 *Conclusions*

RQ1: How, if at all, are design classes engaging with equity, justice, and ethics considerations?

Most design courses are not engaging with equity, justice, and ethics considerations. In particular, there is a large gap in the level of engagement with these topics in engineering design courses compared to design courses outside of the engineering departments. This is a troubling finding and suggests that engineering design curricula could be reformed to meaningfully engage with equity, justice, and ethics considerations.

RQ2: Which design justice paradigms are implicitly or explicitly embedded in design pedagogy?

The design justice paradigms embedded in design pedagogy varied widely across departments. Urban Studies & Planning and Media Arts & Sciences tended to have a similarly high rate of all the design justice questions present in their courses and Electrical Engineering & Computer Science tended to have little to no design justice questions present in their courses. Mechanical Engineering and Nuclear Engineering courses both had a higher presence of Beneficiaries and Scope, but Nuclear Engineering also had a clearer focus on Histories, likely because of the presence of many critique courses that addressed past disasters. Architecture courses had a higher presence of Beneficiaries, Values, and Sites. We also found that some courses included topics adjacent to design justice that did not fit one of the explicit design justice paradigms. For instance, sustainability and climate change were focus areas for several engineering design courses.

RQ3: What impact, if any, have the racial, social, and environmental justice movements of 2020 had on design pedagogy?

We found that the majority of courses had no significant changes between 2019 and 2020. It is possible that this is too short of a time frame to see any significant changes, so we plan to continue this research in the future to determine how long it takes to see meaningful changes in the curriculum.

5 *Discussion*

Our paper highlights the notable absence of design justice principles in design and design critique courses, especially in engineering departments, at a technology institution, despite the racial, social, and environmental justice-centering movements of 2020. We call for a stronger emphasis of ethics, equity, and justice in design pedagogy. While these principles may be found in non-design courses and non-engineering departments, it is important for future developers and designers to responsibly engage with the social contexts and implications of their work, which is emphasized through incorporation of design justice principles in design education—both engineering and non-engineering. Several recommendations emerge from our work for how to include design justice principles. Our work develops a methodology for launching design justice audits as an ongoing process at technology institutions and identifying needs for institutional support for justice-, ethics-, and equity-minded work and education. Through these measures, we expand the epistemic scope of design education to holistically support a more socially minded design education centered on equity, justice, and ethics that empowers future designers to question and consider the societal impacts of technological designs from the earliest stages of ideation.

Overall, the results of this study indicate that design justice themes are, for the most part, not being prioritized in engineering courses. There is much room for improvement, especially in the engineering courses that are more focused on technology development for its own sake rather than viewing technology as a means to an end. Some courses are prioritizing this content through one-off lectures on ethics-related topics or short modules on relevant design justice-adjacent material, such as sustainability. However, most technology-design focused classes have little to no focus on ethics, equity, nor justice. This is a glaring gap in engineering design education that must be addressed. Non-engineering departments had a much more meaningful focus on design justice related topics. Although some non-engineering departments have a stronger focus on these topics, they are often not integrated throughout the entire curriculum and are not the main focus of the design process. Notably, the departments with the largest design justice focus tended to have the smallest number of students, so the overall reach of this content is very limited. At technology-focused institutions, these smaller departments often have fewer resources, less power, and prestige relative to engineering departments, which further limits their sphere of influence. Some possibilities for pedagogical reform may include engineering instructors learning from approaches adopted by colleagues, partnering with colleagues in other departments to develop novel cross-departmental course offerings, and urging their engineering students to take non-engineering courses as a way of improving their engineering practice through the development of intellectual breadth.

In addition to analyzing the department courses, we also examined changes between courses that were taught in both Fall 2019 and Fall 2020. In general, there were few changes in the inclusion of design justice questions addressed between Fall 2019 and Fall 2020. Though there were 17 examples of reform with respect to certain design justice paradigms, there were also five cases of courses reducing the presence of design justice related principles between the two years. While most courses did not change their overall explicit/implicit or meaningful/superficial engagement with design justice, those that did often incorporated more design justice aspects into their curricula, though this was done both implicitly and explicitly. Ideally, we would like to see courses consistently engage with design justice explicitly as demonstrated by the exemplars.

It is also important to note that the events of 2020, namely the COVID-19 pandemic and subsequent transition to virtual teaching, added stress and labor on educators that may have prevented them from embedding these values, as they focused more time on student wellbeing and acknowledging the emotional turmoil that students were facing in the context of the pandemic and the increasing visibility of police brutality and racism in the United States (Park et al., 2020). Virtual teaching and the added toll on educators may have limited instructors from experimenting with content: as such, we may expect to see more of the changes described added into 2021 syllabi when many courses

reverted to the in-person format. Regardless, it is crucial that design justice paradigms be incorporated into design courses directly, so that students see them as fundamental components of the design process. Students who do not learn and practice these concepts will become design practitioners who do not have experience with considering the implications of their work and may become more disengaged with these principles (Cech, 2014). Justice, ethics, and equity cannot be afterthoughts and must be foregrounded in design pedagogy.

We firmly believe that ethics, equity, and justice are fundamental components of engineering and must be incorporated into engineering education. Our results demonstrate that traditional engineering departments engaged much less with design justice than non-engineering departments. We are at a critical juncture to address the lack of design justice in engineering design pedagogy and we are beginning to see commitments in the field to address this. Recently, the deans of multiple engineering departments nationwide jointly authored a letter to The Accreditation Board for Engineering and Technology, Inc. (ABET, 2021) emphasizing the need for implementing a Diversity, Equity, and Inclusion requirement in engineering curricula (ABET, 2021). Some universities have already embarked on significant programs of curricular reform. For example, the University of Michigan has begun incorporating these topics into their engineering education already to work towards "equity-centered engineering" (Gallimore, 2021). At this institution, the Social and Ethical Responsibilities of Computing (SERC), has been working towards centering on social, ethical, and policy considerations in computing pedagogy ((SERC), 2021). In addition to these advancements, there are opportunities and methods for including design justice in future course offerings—as we've seen in our syllabus analysis—such as the inclusion of ethics and equity considerations in courses built around analysis tools and techniques. Courses can also include curriculum components that investigate and discuss the social implications of technology by questioning who has access to the technology, who benefits from it, and who designs it.

While our focus in this paper is on undergraduate and graduate courses, we believe that our work provides contributions and extensions to design practice as well. The training that designers and engineers receive in their coursework in undergraduate and graduate schooling is essential to shaping how these students will develop future technologies. Our work here also provides emphasis on Design Justice that can be translated to corporate and industrial design projects. The audit we've done in this work can be expanded to examining corporate and industrial design projects, encouraging professional designers and engineers to embed Design Justice principles in practice. It could also be used as a benchmark or assessment of corporate attention to Design Justice principles. This extension also provides future work directions for the interface and impact of Design Justice grounded coursework for design in practice.

Impact of social movements on design pedagogy

6.1 Limitations

The primary limitation of this work is the exclusive use of syllabi and course descriptions to assess whether or not courses are including design justice related content. Additionally, we are limited by our access to materials only at a single institution and during a short time frame. It is our hope that the syllabus audit framework allows others who are embedded in their institutions to carry out similar audits of the curricula at their universities.

6.2 Future work

As previously acknowledged, syllabi are not a complete representation of the full pedagogical and learning experience imparted and experienced in a course. Future work will explore additional ways to understand how instructors include design justice considerations in their courses apart from the material listed on the syllabus. To this end, we have developed and disseminated a survey that allows course instructors to articulate how they incorporate design justice principles into their courses to build a more complete picture of how design justice is integrated into pedagogy. We also plan to interview instructors of the courses to understand why and when they incorporate design justice in pedagogy, in addition to any institutional incentives or barriers they encounter while doing so. We expect that this future research will be able to offer recommendations not only for curricular reform but also educational policy reform at institutions of higher learning as well as through accreditation organizations such as ABET.

In our ongoing work, we have already begun collecting syllabi from other departments and additional years to extend this investigation further. We also plan to conduct a comprehensive analysis of our recommendations that we noted when we were completing the syllabus coding. These recommendations will be used to inform how design justice could be incorporated into courses more specifically (i.e., case studies, section of project, etc.). We have also developed a Python script that reads each syllabus and identifies the design justice related keywords within it. We plan to elaborate on this analysis in a future publication to show which terms appear in syllabi across various departments.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

We wish to thank the course administrators at MIT whose assistance was invaluable in gathering the syllabi that constitute our primary source of data for this analysis. We are also grateful to Sasha Costanza-Chock, Maria Yang, Julie Shah, and David Kaiser for their generous and insightful input on our methodology and for their ongoing support of our design justice and design pedagogy project.

Appendix A. List of Course Description Search Terms

1. Justice/Injustice/Just/Unjust
2. Ethics/Ethical
3. Equity/Inequity/Equitable/Inequitably
4. Sustainability
5. Climate/climate change
6. Community
7. Participatory
8. Inclusive
9. Stakeholder
10. Society
11. Impact
12. Self-determination
13. Social Hierarchy
14. Oppression
15. Colonialism
16. Racism
17. Capitalism
18. Imperialism
19. Patriarchy
20. Matrix of Domination
21. Intersectionality
22. Liberatory/Liberation
23. Identity (Group)
24. Emancipation
25. Other
26. None

Notes

1. Authors contributed equally to this work.

References

ABET. (2021). *Big 10+ Universities Deans of Engineering Letter of Support*. ABET Diversity, Equity, & Inclusion. https://www.abet.org/about-abet/diversity-equity-and-inclusion/.

Anderson, R. E., Anderson, R. J., Borriello, G., & Kolko, B. (2012). Designing technology for resource-constrained environments: Three approaches to a multidisciplinary capstone sequence. In *2012 Frontiers in Education Conference Proceedings* (pp. 1–6). https://doi.org/10.1109/fie.2012.6462501.

Baillie, C., Pawley, A. L., & Riley, D. (2012). *Engineering and social justice: In the university and beyond*. JSTOR: Purdue University Press. https://doi.org/10.2307/j.ctt6wq5pf.

Bardzell, S. (2010). Feminist HCI. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10* (pp. 1301–1310). https://doi.org/10.1145/1753326.1753521.

Buchanan, R. (2001). Human dignity and human rights: Thoughts on the principles of human-centered design. *Design Issues, 17*(3), 35–39. https://doi.org/10.1162/074793601750357178.

Cech, E. A. (2014). Culture of disengagement in engineering education? *Science, Technology, & Human Values, 39*(1), 42–72. https://doi.org/10.1177/0162243913504305.

Chong, F. (2015). The pedagogy of usability: An analysis of technical communication textbooks, anthologies, and course syllabi and descriptions. *Technical Communication Quarterly, 25*(1), 12–28. https://doi.org/10.1080/10572252.2016.1113073.

Collins, P. H. (1990). Black feminist thought in the matrix of domination. In *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment* (pp. 221–238). http://www.oregoncampuscompact.org/uploads/1/3/0/4/13042698/patricia_hill_collins_black_feminist_thought_in_the_matrix_of_domination.pdf.

Costanza-Chock, S. (2018). Design justice: Towards an intersectional feminist framework for design theory and practice. *Design Research Society, 2018*. https://doi.org/10.21606/dma.2017.679.

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. The MIT Press. https://doi.org/10.7551/mitpress/12255.001.0001.

Cross, N., & Cross, A. C. (1998). Expertise in engineering design. *Research in Engineering Design, 10*(3), 141–149. https://doi.org/10.1007/bf01607156.

Daly, S. R., Yilmaz, S., Christian, J. L., Seifert, C. M., & Gonzalez, R. (2012). Design heuristics in engineering concept generation. *Journal of Engineering Education, 101*(4), 601–629.

Das, M., & Yang, M. C. (2021). Assessing Early Stage Design Sketches and Reflections on Prototyping. *International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, 85420*. V006T06A004.

D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. https://doi.org/10.7551/mitpress/11805.003.0003.

Fiesler, C., Garrett, N., & Beard, N. (2020). *What do we teach when we teach tech ethics?* 289–295. https://doi.org/10.1145/3328778.3366825.

Friedman, B. (1996). Value-sensitive design. *Interactions, 3*(6), 16–23. https://doi.org/10.1145/242485.242493.

Fu, K. K., Yang, M. C., & Wood, K. L. (2016). Design principles: Literature review, analysis, and future directions. *Journal of Mechanical Design, 138*(10), 101103. https://doi.org/10.1115/1.4034105.

Gallimore, A. D. (2021, August 30). *It's time for engineering to be equity-centered*. Inside Higher Ed. https://www.insidehighered.com/views/2021/08/30/diversity-equity-and-inclusion-should-be-required-engineering-schools-curricula.

Goldman, S. L. (1991). The social captivity of engineering. *Critical Perspectives on Nonacademic Science and Engineering, 4*, 121–143. https://www.structuremag.org/wp-content/uploads/2014/08/C-Infocus-Schmidt-May101.pdf.

Knobel, C., & Bowker, G. C. (2011). Values in design. *Communications of the ACM, 54*(7), 26–28. https://doi.org/10.1145/1965724.1965735.

Lucena, J., & Schneider, J. (2008). Engineers, development, and engineering education: From national to sustainable community development. *European Journal of Engineering Education, 33*(3), 247–257. https://doi.org/10.1080/03043790802088368.

Magney, J., & Bucciarelli, L. L. (1995). Designing engineers. *Technology and Culture, 36*(4), 1030. https://doi.org/10.2307/3106930.

Maxwell, J. (1992). Understanding and validity in qualitative research. *Harvard Educational Review, 62*(3), 279–301. https://doi.org/10.17763/haer.62.3.8323320856251826.

Mitcham, C., & Munoz, D. (2010). Humanitarian engineering. *Synthesis Lectures on Engineers, Technology and Society, 5*(1), 1–87. https://doi.org/10.2200/s00248ed1v01y201006ets012.

Nieusma, D., & Riley, D. (2010). Designs on development: Engineering, globalization, and social justice. *Engineering Studies, 2*(1), 29–59. https://doi.org/10.1080/19378621003604748.

Ostrowski, A. K., Breazeal, C., & Park, H. W. (2021). *Long-Term Co-Design Guidelines: Empowering Older Adults as Co-Designers of Social Robots* 1165–1172. https://doi.org/10.1109/ro-man50785.2021.9515559.

Park, J. J., Park, M., Jackson, K., & Vanhoy, G. (2020). Remote engineering education under COVID-19 pandemic environment. *International Journal of Multidisciplinary Perspectives in Higher Education, 5*(1), 160–166. https://eric.ed.gov/?id=EJ1273006.

Pritchard, J., & Baillie, C. (2007). How can engineering education contribute to a sustainable future? *European Journal of Engineering Education, 31*(5), 555–565. https://doi.org/10.1080/03043790600797350.

Riley, D. (2008). Engineering and social justice. *Synthesis Lectures on Engineers, Technology and Society, 3*(1), 1–152. https://doi.org/10.2200/s00117ed1v01y200805ets007.

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences, 4*(2), 155–169. https://doi.org/10.1007/bf01405730.

Schneider, J., Leydens, J. A., & Lucena, J. (2008). Where is 'Community'?: Engineering education and sustainable community development. *European Journal of Engineering Education, 33*(3), 307–319. https://doi.org/10.1080/03043790802088640.

(SERC), M. S. and E. R. of C. (2021). *SERC framework*. Schwarzman College of Computing. https://computing.mit.edu/cross-cutting/social-and-ethical-responsibilities-of-computing/#SERC-Framework.

Turner, K., Verma, A., & Wood, D. (2021). *Intersectional Antiracism & Technology Design: Building Frameworks to Advance Justice and Equity in Complex Sociotechnical Systems*. Dubai: UAE.

Verma, A. (2021a). The nuclear, humanities, and social science nexus: challenges and opportunities for speaking across the disciplinary divides. *Nuclear Technology, 207*(9), iii–xv. https://doi.org/10.1080/00295450.2021.1941663.

Verma, A. (2021a). What can nuclear engineers learn from design research? A review of the theory and evidence from contemporary American and French nuclear reactor design projects. *Nuclear Engineering and Design, 379*, 111114. https://doi.org/10.1016/j.nucengdes.2021.111114.

Verma, A., & Djokić, D. (2021). Reimagining Nuclear Engineering. *Issues in Science and Technology, 37*(3), 64–69.

Welch, J. K., & Patton, M. Q. (1992). Qualitative evaluation and research methods. *The Modern Language Journal, 76*(4), 543. https://doi.org/10.2307/330063.

Understanding inspiration: Insights into how designers discover inspirational stimuli using an AI-enabled platform

Elisa Kwon, Vivek Rao and *Kosa Goucher-Lambert*, Department of Mechanical Engineering, University of California, Berkeley, USA

Throughout the design process, designers encounter diverse stimuli that influence their work. This influence is particularly notable during idea generation processes that are augmented by novel design support tools that assist in inspiration discovery. However, fundamental questions remain regarding why and how interactions afforded by these tools impact design behaviors. This work explores how designers search for inspirational stimuli using an AI-enabled multi-modal search platform, which supports queries by text and non-text-based inputs. Student and professional designers completed a think-aloud design exploration task using this platform to search for stimuli to inspire idea generation. We identify expertise and search modality as factors influencing design exploration, including the frequency and framing of searches, and the evaluation and utility of search results.

Keywords: design tools, design behavior, creativity, protocol analysis, design inspiration

In design and other creative domains, becoming inspired may be associated with experiencing a serendipitous encounter. For designers, inspiration is important for assisting with the generation of creative solutions. One definition of inspiration proposed by Gonçalves et al. (2016) references the role of an external stimulus in altering the creative process by influencing problem framing or solution generation. Significant effort has been made to describe and understand inspiration more formally, such as through an exploration of the influence of features of inspirational stimuli on ideation and design outcomes (e.g., by Chan et al., 2011; Fu et al., 2013b; Goucher-Lambert, Gyory, Kotovsky, & Cagan, 2020), the cognitive processes underlying designers' search processes (Gonçalves et al., 2013, 2016), and the methods and systems used to derive and retrieve inspirational stimuli using, e.g., data-driven techniques (Jiang et al., 2022).

In the present work, designers engage with a new AI (Artificial Intelligence)-enabled search platform to discover inspirational stimuli to aid solution

Corresponding author:
Kosa Goucher-Lambert
kosa@berkeley.edu

www.elsevier.com/locate/destud
0142-694X *Design Studies* **88** (2023) 101202
https://doi.org/10.1016/j.destud.2023.101202

generation to a design task. Search by novel interaction modalities, including by non-text-based search inputs, are made available in the developed platform, which is described in prior work (Kwon et al., 2022). Due to the possibility that the modes of search presented in our search platform are less familiar, the role of expertise when engaging with these search inputs is also studied. The aim of this work is to extend upon knowledge regarding processes employed by designers to search for inspirational stimuli, especially when facilitated by design support tools using new interaction mediums. Specific research questions guiding this work include the following.

- RQ1: How does input modality in an AI-enabled platform impact search for inspiration?
- RQ2: How do students and professional designers compare in their search for inspiration using an AI-enabled platform?
- RQ3: What rationale do designers provide for their evaluation and selection of inspirational stimuli?

In Figure 1, the approach taken to answer these research questions, **RQ1**-**RQ3**, is presented. First, participants completed a design task using our search platform from which we collected their platform interactions and think-aloud descriptions of their search processes. This experimental data is used to describe how designers search for inspirational stimuli in terms of search activities, behaviors, and pathways, as defined and outlined in Sec. 2.3. The developed framework is used to code experimental data into search behaviors, including how searches were defined and how the retrieved results were evaluated and selected. Search pathways then explore the relationships between search behaviors, such as how designers' selections of platform-retrieved stimuli are related to their evaluations of the same stimuli. Quantitative comparisons between search activities, behaviors, and pathways made using the available search modalities (keyword, part, and workspace, as defined in Sec. 2.1) by student vs. professional designers are detailed throughout Sec. 3.1, addressing **RQ1** and **RQ2**. As a final contribution of this work, answering **RQ3**, rationale and motivations for following specific pathways are discussed in Sec. 3.2 through select examples. The presented examples demonstrate how the search platform both accomplishes and influences designers' search goals. These results can be helpful for the further development and use of design tools, including search interfaces, by leveraging insight gained into the cognitive processes underlying the search for, evaluation, and selection of inspirational stimuli.

1 *Background*

To deepen our understanding of how designers search for inspiration, we consider three main components influencing this process. First, insights from past work are reviewed to motivate why designers should be exposed

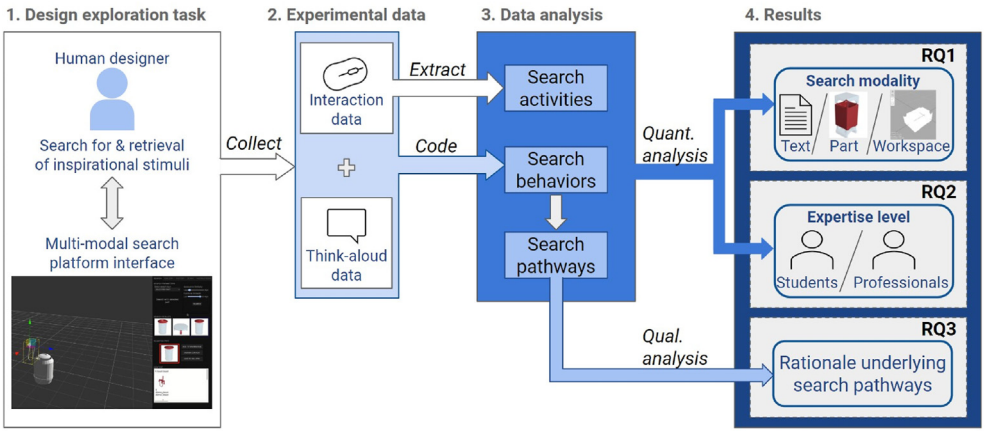

Figure 1 ***Paper overview:*** *Alignment of research questions to (**1**) design exploration task conducted (**2**) data collected (**3**) analysis approach, and (**4**) results presented*

to external stimuli during the design process. Second, cognitive processes and preferences underlying designers' search for inspiration are explored. Third, methods to support designers' search for and retrieval of inspirational stimuli, including various AI-enabled methods, are reviewed. This background is relevant to the work presented in the current paper, which investigates how designers search for inspirational stimuli when using an AI-enabled multimodal search platform.

1.1 Impact of inspirational stimuli on design

Inspiration is discussed in this work as a process where a stimulus influences the thought process used towards problem framing or solution generation (Gonçalves et al., 2016). Accordingly, *inspirational stimuli* is used to describe external stimuli providing inspiration. Inspirational stimuli play a key role at many points across the design process: Lucero (2012) found that inspirational stimuli manifested in moodboards helped designers frame, align, abstract, and direct their work across various design activities. Inspirational stimuli can importantly aid designers by triggering idea generation and providing an anchor for mental representations of designs (Eckert & Stacey, 2000), but can also negatively lead to design fixation, where designers unconsciously focus on particular aspects of an object or task, resulting in limited idea generation (Jansson & Smith, 1991). Across many controlled experiments, the influence of external stimuli on design ideation has been studied to identify characteristics that make them useful or beneficial to designers, while aiming to avoid such fixation effects.

One significant area of prior work on the role of inspirational stimuli on design has focused on stimuli promoting analogical reasoning, defined as the process

where a mapping association is made based on relations between a source and target (Gentner, 1983). Analogies are one form of external design stimuli suggested to be beneficial for creativity by encouraging new inference formation and problem construal (Gentner & Markman, 1997). Various features of analogical stimuli have been investigated, notably analogical distance, referring to the proximity of domain of the given stimulus to the designer's current problem. Far-field stimuli, for example, have been shown to lead to idea novelty (Chan et al., 2011; Goucher-Lambert & Cagan, 2019), compared to near-field stimuli, which can improve feasibility, relevance, and idea quantity (Chan et al., 2015; Goucher-Lambert et al., 2019, Goucher-Lambert, Gyory, Kotovsky, & Cagan, 2020). Fu et al. propose a "sweet spot" of analogical distance, discounting examples that are "too near" or "too far" to be beneficial to designers to apply analogical reasoning (Fu et al., 2013b).

Several factors of inspirational stimuli other than analogical distance can also impact design outcomes. The timing of when the stimulus is presented to the designer is important: it is more effective to provide once a designer has started to generate ideas for a design task than before idea generation has begun (Tseng, Moss, Cagan, & Kotovsky, 2008). The current ideation state of the designer is also relevant, where stimuli received when the designer is stuck can help produce more ideas, than when provided at predefined intervals (Siangliulue et al., 2015). In prior work, the level of detail or concreteness of the stimulus is another explored feature. Descriptions of design stimuli can be more general vs. domain-specific (Linsey et al., 2008) or constitute concrete design examples vs. abstract system properties (Vasconcelos, Cardoso, Sääksjärvi, Chen, & Crilly, 2017). While concept-level design stimuli (e.g., keywords extracted from patents) can provide more rapid inspiration, more comprehensive stimuli (e.g., patent documents) can provide rich engineering design details (Luo et al., 2021).

The modality in which the stimulus is represented to the designer is also considered. The impact of visual stimuli compared to physical stimuli (Toh & Miller, 2014), or in combination with textual stimuli (Borgianni, Rotini, & Tomassini, 2017; Han et al., 2018; Malaga, 2000), or other images (Hua et al., 2019), are examples of how representation modalities have been explored in prior work. Designers are found to tend to prefer visual information (Gonçalves et al., 2014; Linsey et al., 2011), which can lead to increased idea novelty (Linsey et al., 2008). When combined with unrelated semantic elements, images can promote creative idea generation (Han et al., 2018), especially when compared to words presented alone (Malaga, 2000). Sketches represent one specific form of visual 2D stimuli. Students have been found to seek and be most influenced by highly resolved sketch stimuli rather than rough sketches (Wallace et al., 2020). Experts may value sketch stimuli for

their contextual content, while students value sketch stimuli for their real-life resemblance and direct connection to the task in question (Cai, Do, & Zimring, 2010).

Visual stimuli can also be represented in 3D, such as in 3D modelling. When comparing the use of 2D and 3D stimuli, further differences in designers' expertise level are found: Gonçalves et al. (2014) demonstrated that professional designers valued 3D object- and 2D image-based stimuli equally for inspiration, while student designers valued image-based stimuli more than other modalities. One key factor in motivating this difference is professionals' valuation of the amount of information object stimuli present to them. Their valuation of this information is reflected by their work on 'real' design solutions as opposed students' work on conceptual design solutions. Our previous work presented 3D-model parts to designers as stimuli based on chosen input modalities and analogical distance parameters (Kwon et al., 2022). In this work, we found that the modality used to *search for* inspirational stimuli affects what is discovered and how it is used. The present work extends upon these results by further examining the role of expertise when using various search modalities.

The impact of various features of inspirational stimuli on design outcomes are reviewed to motivate the present study of designers' search for stimuli to inspire idea generation. While much is known regarding how inspirational stimuli can impact the design process, the search behaviors employed by designers, as well as the methods enabling these processes, are less understood. In the present work, designers' use of an AI-enabled search platform is investigated, providing insight into designers' search for inspiration. The cognitive processes underlying these behaviors, and the design tools used to support them, are next reviewed.

1.2 Cognitive perspectives of search for inspiration

Sio et al. (2015) describe designing as a process of searching for task-relevant concepts and integrating these concepts into a design solution. Gonçalves et al. (2016) further define the search for inspiration process as initiated by a specific intention and goal, often expressed by keyword or other search input. To select keywords to initiate the search process, they discovered that designers search for closely related terms to the design problem earlier in the task and more distantly related terms later in the task. These search strategies are supported by related research on analogical stimuli that suggests the importance of both analogically near and far stimuli on promoting beneficial design outcomes (Fu et al., 2013b).

However, when the goal of a designer is not well defined, how is the search process initiated? Two search processes are proposed by Gonçalves et al. (2013, 2016): active search, which is an intentional process driven by specific goals, and passive search involving an accidental, non-deliberate discovery of relevant inspiration sources. Passive search is attributed to the random discovery of unexpected results, which can provide beneficial sources of inspiration (Gonçalves et al., 2016; Herring et al., 2009). Similar to the dichotomy between active and passive search, information retrieval theory differently defines lookup vs. exploratory search behaviors (Sutcliffe & Ennis, 1998). Exploratory search promotes knowledge acquisition and supports evolving needs, compared to lookup search activities which are used to meet precise search goals (Marchionini, 2006). Exploratory search is related to the examination of more results than lookup search (Athukorala et al., 2016). Passive and exploratory search strategies may be used when task constraints are low. Biskjaer et al. (2020) investigate the effect of task constraint on inspiration search strategies, finding that low task constrainedness was associated with more frequent and divergent search. When searching for inspiration, both active and passive search strategies are relevant. Designers are expected to find relevant inspirational stimuli through expressing specific search intent as well as through passive encounters with inspirational stimuli when search goals are not as clearly defined or unexpected search results are encountered. This intentional search for and passive discovery of inspirational stimuli can be facilitated by design-support tools, such as the search platform presented in this work and others reviewed in the next section.

1.3 Design support tools for inspirational stimuli retrieval

The discovery of inspirational stimuli is a process that can be supported by design support tools, including those that rely on AI. The interactions enabled by these systems and used by designers are important to consider towards understanding design behaviors, such as search for information and inspiration. Different computational and AI-enabled methods and tools have been proposed to provide inspiration to designers through external stimuli, applied in contexts like biologically inspired design (Vattam, Wiltgen, Helms, Goel, & Yen, 2011; Goel, Vattam, Wiltgen, & Helms, 2012; Nagel & Stone, 2012; Sartori et al., 2010), and using sources of designs such as patent databases (Murphy et al., 2014; Fu et al., 2013a, 2013b) or crowd-sourced solutions (Goucher-Lambert & Cagan, 2019; Kittur et al., 2019). Different from these studies, the present work focuses on the search for and retrieval of inspirational design stimuli, rather than on the stimuli provided by these systems. The use of multi-modal inputs is specifically studied to understand how they can support inspirational search. Various methods have also been developed that utilize non-text inputs, such as through image or sketch-based inputs.

Sketch-based retrieval of visually similar examples can importantly support visual analogy (Zhang & Jin, 2020, 2021). Image-based search using visual similarity can also extract relevant examples from sources such as patent documents (Jiang, Luo, Ruiz-Pava, Hu, & Magee, 2020, 2021). DreamSketch is an example of a sketch-based user interface that provides designers with 3D-modeled design solutions based on early stage 2D-sketch-based designs (Kazi, Grossman, Cheong, Hashemi, & Fitzmaurice, 2017). SketchSoup inputs rough sketches and generates new sets of sketches, which can inspire further concept generation (Arora et al., 2017). 3D-represented design ideas can be recognized by tools such as the InspireMe interface, which provides suggestions for new components to add to a designer's initial 3D model (Chaudhuri & Koltun, 2010). Design support tools that recognize these inputs can be beneficial since sketching itself is a process that can assist idea formation (Botella, Zenasni, & Lubart, 2018). In general, interactions with visual stimuli can help trigger new mental images and thus new ideas for design (Menezes & Lawson, 2006). By recognizing a designer's sketch as it is developing, the system can also provide relevant computational aid when it is advantageous to the designer during the design process (Do, 2005).

These examples suggest that multi-modal inputs may be used to more effectively recognize the idea or query expressed by a designer, and support the further search and exploration of the design space. The present work extends on these examples by directly assessing how these modalities are used to search for inspirational stimuli. We aim to describe the behaviors that interactions within these systems represent and to understand the cognitive processes involved in how designers search.

2 Experimental approach

To support the main aims of this research, we conducted a study facilitated by the use of a multi-modal search platform to investigate how designers search for inspirational stimuli. The study was conducted using Zoom, where participants' progress was screen and audio recorded. Screen recordings were used to capture how participants engaged with the AI-enabled design tool provided. In this section, the details of the search platform used, participants, and the design exploration task they completed are described. The methods and approach taken to analyze the results presented in this paper are also introduced.

2.1 AI-enabled multi-modal search platform

The design tool, a multi-modal search platform, relies on a deep-learning approach to efficiently retrieve relevant 3D-model parts based on the user's input query. Deep-neural networks are used to model similarities between various 3D-model parts from the PartNet dataset, consisting of 24 object categories and 26 671 3D-model assemblies. The platform is extensively described

in our prior work (Kwon et al., 2022). The search platform supports search for parts in the dataset using three types of input. The first search type is keyword-based, where parts with related text labels are returned. The second and third search types are part-based and workspace-based, where new parts are retrieved using visual snapshots taken of a selected 3D-model part or the participant's current workspace (composed of 3D-model parts), respectively. In part and workspace searches, sliders in the user interface also specify how similar the desired results are from the inputs by visual and functional similarity. For each search made, three parts are retrieved and shown in the user interface. Examples of keyword and part searches and results in the interface are shown in Figure 2a, b.

The interface also allows three additional actions to further interact with the retrieved results. Parts can be added to and modified in the user's 3D workspace using an 'Add to Workspace' button. Workspace-based searches are made with snapshots of the entire workspace with parts added to the workspace using this action. Since all results are retrieved from the PartNet dataset, which contains information on neighboring parts in the same assembly of a given result, this information may also be viewed using a 'View in Context' button. For a selected part, this action allows further understanding of the retrieved part's placement in its original context. Uses of these features for a keyword search result for "container" are also shown in Figure 2c, d. Finally, parts can be added to a gallery of collected 3D parts using an 'Add to gallery' button. During the design task, the gallery was available for participants to access and select parts from at any point. For any given search made, none to all actions can be performed, in any order.

Interactions afforded by this platform were investigated in our prior work. Using this search platform and the same design prompt provided in the present study (described in Sec. 2.2.2), a controlled experiment was conducted ($n = 21$) in which keyword, part, and workspace searches were engaged separately in three subtasks (Kwon et al., 2022). Participants were instructed to conduct a minimum of five searches using each input, and to save a minimum of three parts to their gallery of parts. The goal of this prior study was to analyze participants' interactions in the platform and relate these actions to strategies involved in searching for inspirational examples. Understanding how each modality was used and interacted with was the main aim of this study, instead of how designers may have naturally used them to achieve specific design outcomes. Distinct outcomes using each search modality were found, including the most frequent use of the part-based search, but low engagement with the returned parts (e.g., by viewing related parts in the same object assemblies or adding them to the 3D workspace). We speculated that increased part-based search but decreased engagement may have been due to the task requirement to continue to search until desired results were obtained. Based on these findings, we aim to further understand in the present

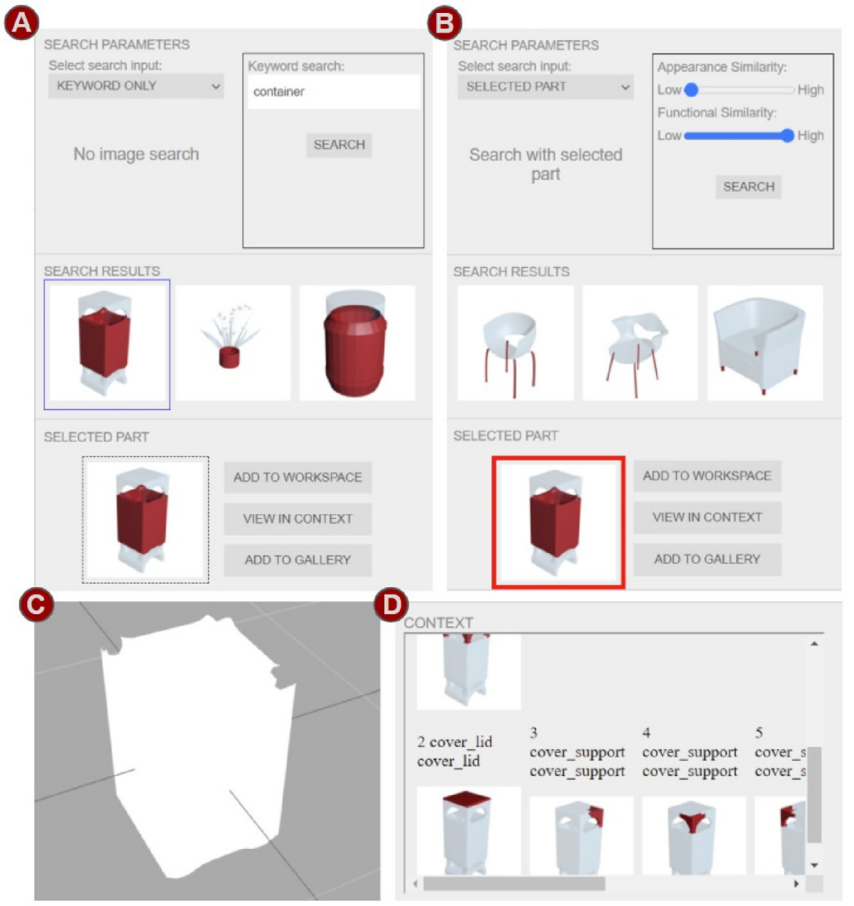

Figure 2 ***Features of multi-modal search platform:*** ***(a)*** *Keyword search and results for "container";* ***(b)*** *Part search with selected container result;* ***(c)*** *Container added to 3D workspace;* ***(d)*** *Container viewed in context*

study how each search modality supports designers' search goals when used freely in the same task, and to elucidate their intentions and discoveries by introducing a think-aloud protocol.

2.2 Design exploration task and think-aloud protocol

2.2.1 Participants

Participants were recruited for the study via email solicitation among graduate students at the University of California, Berkeley, and industry professionals. All participants were required to meet the minimum eligibility of having at least one year of Computer-Aided Design (CAD) experience. Fifteen participants volunteered for the study, including eight professionals recruited from industry and seven students recruited from the university. Self-reported experience with CAD tools of students (three males, four females) and

professionals (seven males, one female) is summarized in Table 1. Students consisted of six Ph.D. students in Mechanical Engineering and one Master of Design student. Professionals included five designers and three engineers by job title, across organizations ranging from <10 to >10 000 employees. Participants were offered $20 compensation for their participation in the 1-h study, detailed below. This study was approved by the Institutional Review Board (IRB) at the University of California, Berkeley.

2.2.2 Study objective and instructions

The study objective presented to designers was to use the multi-modal search platform to search for parts that inspire solutions to the design for "a multi-compartment disposal unit for household waste". No additional design requirements or specifications on the relationship between the searched for parts to the design problem were provided. Participants completed the task in <30 min., including approx. 15 min. learning how to use the interface through a guided tutorial embedded in a Qualtrics link accessed at the start of the study. Participants read descriptions and viewed videos of the interface in use and followed instructions for completing example searches in the interface. Instructions for following a think-aloud protocol directed participants to explain their interactions aloud, with particular attention to: (1) why the specified search type and input were used before executing a search and (2) whether the returned result was what was expected, or not, after executing a search. Based on prior work in which the same task was completed without think-aloud instructions, these prompts were specified to elucidate motivations behind previously observed search behavior during the task. Designers were provided with the suggestion to conduct five of each search type (keyword, part, and workspace). These guidelines were not strictly enforced during the task to allow designers to freely use the search types in any order.

2.3 Analysis of design exploration task and think-aloud data

The main approach taken to analyze results from the design exploration task is to examine three levels of search: activities, behaviors, and pathways. Further elaborating on Figure 1, the relationships between the task data and these search levels are summarized in Figure 3. These search levels are defined to understand designers' search processes through interactions with the search platform and transcriptions of think-aloud data. Search activities describe how designers conducted multi-modal searches. Search behaviors are extracted from both platform interactions and accompanying think-aloud data before and after executing searches. Search pathways are then used to discuss how search behaviors are related.

Firstly, **search activities** are studied, related to the frequency of use of the multi-modal inputs in the search platform. Task data captured by the search platform was extracted, including individual button presses to conduct

Table 1 CAD experience of student and professional designers

CAD experience (years)				
Participant type	*1–2*	*3–5*	*6–9*	*>10*
Students	5	0	2	0
Professionals	0	2	3	3

searches, view parts in context, save parts to the gallery, add parts to the workspace, and all individual part data. The frequencies of searches made using each input type are specifically explored in this work.

Secondly, to abstract and classify **search behaviors** from platform interactions and think-aloud data, a framework was developed. This framework is an extension from Gonçalves et al.'s description of the inspirational search process, which outlines the formulation of search inputs, the (successful or unsuccessful) search for and selection of a stimulus, assessment of its correspondence to the designer's expectations, and finally the designer's choice to incorporate and adapt the stimulus to the problem at hand (2013). In the present work, the behaviors identified include: how designers defined searches (whether new or continued searches for results were made), evaluated search results (whether results were expected or unexpected), and selected search results (whether results were accepted or rejected from their design). This framework is further detailed in Table 2. For each behavior (search definition, evaluation, and selection), two possible levels were assigned by following the listed criteria, shown in Table 2. Representative examples of quotes from the think-aloud data associated with each search behavior are also provided.

Two coders, each with at least three years of postgraduate design research experience, assessed the data using the framework. Coder 1 manually transcribed think-aloud data from screen and audio recordings of the design task sessions. Coder 1 identified user interaction behavior and think-aloud quotations pertaining to the three defined behaviors (definitions, evaluations, selections). A total of 235 search actions were identified, an average of 15.7 searches per participant. To validate the framework, Coders 1 and 2 independently applied framework codes to 15% of the dataset. A minimum of 84% interrater reliability for search definition codes was determined using percentage agreement, and 0.69 using Cohen's Kappa, indicating substantial agreement (Stemler, 2004). This suggests that the developed coding framework was relatively consistent across coders.

However, exceptions emerged to the defined criteria when codes for search definitions and evaluations were assigned. An example of an exception to the defined criteria is when a 'new search' followed a 'rejected' outcome, e.g., when a participant made a new search for a "lid" without accepting results for their previous search for a "handle". Based on the criteria defined, this

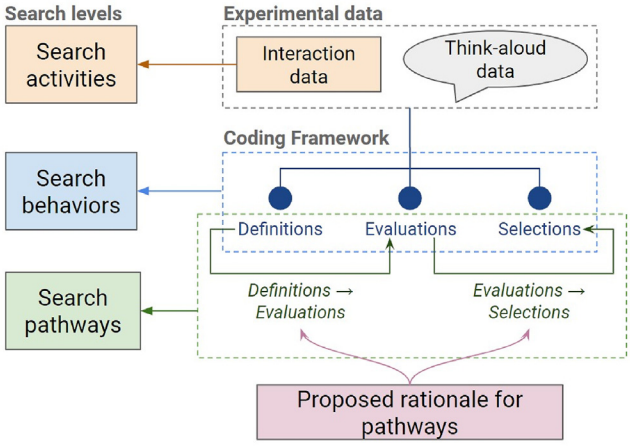

Figure 3 ***Overview of relationships between three levels of search examined in results:*** *activities describing interactions, behaviors, coded from experimental data with the developed framework, and pathways discussing relationships between behaviors. Rationale for select pathways are discussed*

Table 2 Search behavior framework: Classification scheme for search behaviors from task and think-aloud data

Behavior: *Description*	*Classification criteria*	*Representative example of associated quote*
Search Definition		
New: Beginning of a new search for a result	Follows an 'accept' outcome of a previous search (see below)	*"I want to see a disposal unit" (**P8**)*
Continuing: Continuation of a search for a result	Follows a 'reject' outcome of a previous search (see below)	*"Maybe instead of cylinder, some kind of rectangular cube" (**P7**)*
Search Evaluation		
Expected: Results match designer's expectation	Explicit acknowledgement that the result is what was searched for or preceding an 'accept' outcome, if no accompanying verbal statement	*"Yes, I like these features. This is providing what I'm looking for" (**P10**)*
Unexpected: Results do not match designer's expectation	Explicit acknowledgement that the result is not what was searched for/is unexpected or preceding a 'reject' outcome, if no accompanying verbal statement	*"This is not what I was expecting - I was expecting to see more lids, whereas these are table tops" (**P4**)*
Search Selection		
Accept: Results are accepted by designer	Result is added to the designer's developing design in the 3D workspace or saved to their gallery of parts	*"This is a shape that could possibly be used in my design. So I'm going to add it to my gallery." (**P12**)*
Reject: Results are rejected by designer	Result is not added to the designer's developing design in the 3D workspace or saved to their gallery of parts. Designer continues to search again.	*"This is not what I was thinking, but this is a trashcan, for sure."* [makes continued search] *"I'm maybe more looking for a cabinet" (**P5**)*

search should be labelled as a continuation of a prior search, but is clearly indicated by the designer to be a new search for a different part. By identifying these characteristics of designers' search behavior, the relationships between what designers search for and what they actually find useful can be explored. Coder 1 coded the entire dataset accounting for these exceptions.

Linking related search behaviors, **search pathways** are the third level of search explored in the present analysis. For a given search, designers follow pathways between defining and evaluating searches and evaluating and selecting parts to incorporate into designs. Illustrative examples from the study of various search pathways can be found in Table 3. Investigating the link between search definitions and evaluations can help uncover if designers have different expectations regarding search results they have repeatedly searched for, or are searching for, for the first time. By studying search evaluation–selection pathways, the influence of encountering unexpected search results on stimuli selection can be examined. Designers may be inspired positively or become negatively fixated on parts they are originally intending to find. These pathways are studied since stimuli selection is known to depend upon how a search is defined and the goal associated with the search (Gonçalves et al., 2016).

3 Results

Following the analysis approach introduced, results detailing participants' search activities, behaviors, and pathways are presented and discussed in this section. In Sec. 3.1, quantitative analyses of each level of search are conducted to examine differences between searches made using keyword, part, and workspace inputs and made by students and professionals, addressing **RQ1** and **RQ2**, respectively. **Search activities** describe how designers used the different search modalities in the platform in terms of frequencies of use. Using the classification scheme established in Table 2, **search behaviors** are investigated. **Search pathways** provide further insight into the relationship between search behaviors, linking search definitions with evaluations, and evaluations with selections. Finally, in Sec. 3.2 an exploration of various search pathways is also made to address **RQ3**, revealing insights into the rationale designers express for defining, evaluating, and selecting inspirational stimuli.

3.1 Quantitative analyses of search activities, behaviors, and pathways

3.1.1 Search activities: designers' use of keyword, part, and workspace searches

The frequency of use of each search modality (keyword, part, workspace) by designers of each level of expertise (student, professional) are first compared. A Poisson regression model, which is used to model count variables, was selected to analyze these differences. A mixed effects model was constructed

Table 3 Illustrative examples of search pathways linking search behaviors

#	*Search pathway*	*Group*	*Type*	*Associated quote/action*
1	New → Unexpected	Student	Workspace	*"I can search for something like. I can use the current workspace … maybe 50% appearance and full functionality to find some other stuff.* ***These are all irrelevant****"*
2	Expected → Rejected	Professional	Part	*"Ahh, yes that's good, I'm seeing kind of like very close matches … I'm going to keep playing around with sliders till I* ***get something closer****"*
3	Expected → Rejected	Student	Keyword	*"I'm going to look for a 'lid' … Ok, yes,* ***I'm looking for something like this****, something square and flat …* ***I want it to be flat*** *and cover [the bin] completely."* [Searches again]
4	Unexpected → Accepted	Student	Workspace	*"I'm looking for something similar to this waste bin so that it can look for the top of the waste bin … Well* ***that's kind of funny****"* [referring to wheel results]. *"Now we can* ***add wheels*** *to this and make it mobile, which is good!"*

using R in RStudio, leveraging the lme4 package to incorporate both fixed (modality, expertise) and random (participant) effects using Laplace Approximation. The model predicts the effects of modality and expertise on the log of frequency of searches made by participants using each search type (N = 45, 15 participants x 3 modalities). Results of the Poisson regression are summarized in Table 4.

Model estimates (β) define the change in the log of frequency associated with each predictor compared to the specified reference (i.e., part or workspace search compared to keyword search and student compared to professional designer). To analogously describe the change in expected search frequency (rather than the change in log of frequency) given the predictor compared to the reference, incidence rate ratios with 95% confidence intervals are also reported. In the context of Poisson regression models, incidence rate ratios are equivalent to e^{β}. The average number of searches made by student (blue) and professional (red) designers using keyword, part, and workspace searches are visually presented in Figure 4.

R1.1. Search activities: most searches are made by keyword. The first comparison made between designers' use of keyword, part, and workspace inputs when searching is in the frequencies of searches conducted using each search type. Significant differences were found in the expected frequency of searches made using part and workspace, compared to keyword searches. Search frequencies for part and workspace searches are 0.39 ($p < 0.001$) and 0.19 ($p < 0.001$) times the search frequency of keyword searches, respectively. Workspace searches represent the most comparatively novel feature offered by the tool, while keyword searches are likely the most familiar input to designers. These results present an important consideration in the design of

Table 4 Poisson regression model predicting search frequency using each modality (n = 45)

Outcome variable	*Predictor*	*Level*	β	p	*Incidence rate ratio* (e^{β})	*95% C.I.*
Search frequency	**Modality**	*Keyword*	(Ref)	(Ref)	(Ref)	(Ref)
		Part	−0.94	<**0.001**	0.39	(0.29, 0.52)
		Workspace	−1.7	<**0.001**	0.19	(0.12, 0.28)
	Expertise	*Professional*	(Ref)	(Ref)	(Ref)	(Ref)
		Student	0.15	0.40	1.16	(0.80, 1.7)

multi-modal inspirational search tools for engineering design: designers, regardless of experience level, more readily use familiar search modalities in their search process.

R1.2. Search activities: Professionals and students do not differ by frequency of search type use. No significant difference was found between participant groups in the frequencies of searches made ($\beta = -0.15$, $p = 0.40$). Student and professional designers therefore do not appear to differ in the modality of search for inspiration they engage when using the multi-modal platform. Adding to Result **R1.1**, both students and professionals used keyword searches the most and workspace searches the least.

3.1.2 Search behaviors: designers' definition, evaluation, and selection of search results

The second level of search examined are behaviors, including how designers define searches and evaluate and select search results. The average proportions across participants of search behavior outcomes made using keyword, part, or workspace searches and by professional or student designers are summarized in Table 5.

To determine the impacts of search modality and designer expertise on search behavior outcomes, mixed effects binary logistic regression models are used. Three models were constructed to demonstrate whether modality and expertise are significant predictors for whether a search was new (vs. continuing), and its result was evaluated as expected (vs. unexpected) and accepted (vs. rejected). Mixed effects logistic regression models were also constructed in R using the lme4 package in RStudio, and incorporated both fixed (modality, expertise) and random (participant) effects using Laplace Approximation.

The results from each regression model are summarized in Table 6, where search definitions (as new), evaluations (as expected), and selections (as accepted) are analyzed as separate outcome variables. Model estimates (β), significance values (p), odds ratios (e^{β}) and their corresponding 95% confidence intervals are reported in Table 6. Estimates for modality are in reference to keyword searches, and for expertise in reference to professionals. Findings

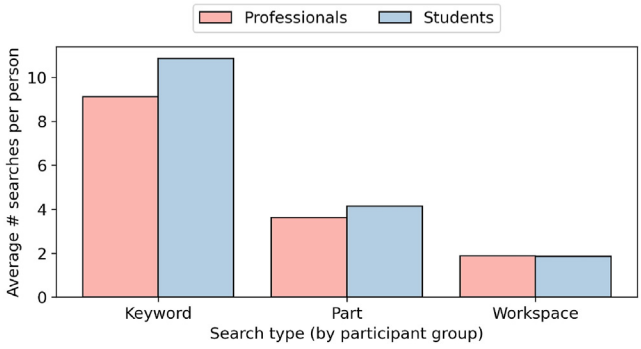

Figure 4 ***Average (per person) frequency of search type use:*** *Comparison between search types (keyword, part, workspace) and participant groups (professionals ($n = 8$), students ($n = 7$))*

across these models are discussed further in this subsection in terms of search modality and designer expertise, separately. To aid with the interpretation of these results, Figure 5 visualizes the odds ratios of each estimate compared to the indicated references for selection and evaluation outcomes. Odds ratios < 1 with confidence intervals that do not cross odds $= 1$ represent that the predictor is significantly less likely than the reference to result in the behavior. Odds ratios > 1 would indicate that the predictor is more likely to result in the behavior than the reference.

R2.1. More keyword search results are expected and accepted. Considering the impact of search modality on the generation of new vs. continued searches, no significant differences between keyword and part or workspace searches were found. Designers are known to rely on "random active search processes" to discover inspiring stimuli when they have a search intention, but do not have a keyword in mind to conduct the search (Gonçalves et al., 2016). Designers' use of part and workspace inputs to formulate new searches demonstrates that these modalities may help achieve the gap between intentional search and uncertainty of what to search for.

However, workspace searches are significantly less likely by 0.25 times than keyword searches to result in an expected evaluation ($p = 0.015$). In other words, workspace search results are 4 times more likely be unexpected than keyword search results. In total, 156/235 (66.4%) searches retrieved results that were identified as unexpected. As shown in Table 5, this high proportion of unexpected search results is disproportionately true for searches made with workspace inputs (24/28, 85.7%) in comparison to keyword searches (91/149, 61.1%). This finding may reflect that designers did not know what to expect

Table 5 Average proportions (%) of search behaviors across search types and participant groups

Search behavior		*Search types*			*Participant group*	
		Keyword	*Part*	*Workspace*	*Professional*	*Student*
Definition	New	40.94%	32.76%	53.57%	45.30%	35.59%
	Continuing	59.06%	67.24%	46.43%	54.70%	64.41%
Evaluation	Expected	38.93%	29.31%	14.29%	40.17%	27.12%
	Unexpected	61.07%	70.69%	85.71%	59.83%	72.88%
Selection	Accept	40.94%	25.86%	35.71%	43.59%	29.6%
	Reject	59.06%	74.14%	64.29%	56.41%	70.34%

when engaging workspace searches. One student designer noted: *"If I want the same functionality in the entire workspace in one part, I don't quite know what that means in this context"*. This example can help to explain results in Figure 4, and why workspace searches were less frequently used: designers often had different expectations of what such searches would yield, than what was actually returned. Beyond the designer's ability to interpret these results, also reflected is the computational difficulty of retrieving relevant and expected parts using visual and functional features. This suggests the need for further work to improve the effectiveness of this search modality to better meet designers' expectations.

A significant difference in the acceptance of part and keyword searches was found, where part searches were 0.49 times less likely to be accepted ($p = 0.041$). On average, designers accepted results from only 25.7% of part searches, while 40.9% of keywords search results and 35.71% of workspace search results were accepted (Table 5). This low likelihood of acceptance corresponds to insights from our prior study, as described in Sec. 2.1, where part

Table 6 Binary logistic regression models predicting search behavior outcomes (n = 235)

Outcome variable	*Predictor*	*Level*	β	p	*Odds ratio* (e^{β})	*95% C.I.*
Definition:	**Modality**	*Keyword*	(Ref)	(Ref)	(Ref)	(Ref)
New = 1, Continued = 0		Part	−0.36	0.27	0.70	(0.36, 1.3)
		Workspace	0.50	0.23	1.64	(0.73, 3.8)
	Expertise	*Professional*	(Ref)	(Ref)	(Ref)	(Ref)
		Student	−0.40	0.14	0.67	(0.39, 1.1)
Evaluation:	**Modality**	*Keyword*	(Ref)	(Ref)	(Ref)	(Ref)
Expected = 1, Unexpected = 0		Part	−0.49	0.17	0.61	(0.30, 1.2)
		Workspace	−1.4	**0.015**	0.25	(0.068, 0.69)
	Expertise	*Professional*	(Ref)	(Ref)	(Ref)	(Ref)
		Student	−0.63	**0.044**	0.53	(0.27, 1.0)
Selection:	**Modality**	*Keyword*	(Ref)	(Ref)	(Ref)	(Ref)
Accepted = 1, Rejected = 0		Part	−0.71	**0.041**	0.49	(0.24, 0.96)
		Workspace	−0.26	0.56	0.77	(0.32, 1.8)
	Expertise	*Professional*	(Ref)	(Ref)	(Ref)	(Ref)
		Student	−0.62	**0.025**	0.54	(0.31, 0.92)

Understanding inspiration

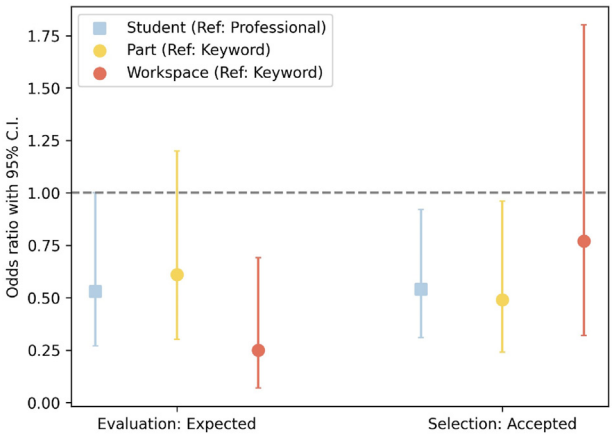

Figure 5 Odds ratios with 95% confidence intervals for predictors of evaluation and selection outcomes. Odds are computed with respect to the indicated reference

searches were most frequently used. Participants may have conducted many part searches because they did not immediately find desirable results, prompting further search.

R2.2. Students exhibit narrower search behaviors than professionals. The impact of expertise on the definition of searches was not found to be significant, but professionals and students did differ by how searches were evaluated and selected. Students, when compared to professionals, were 0.53 ($p = 0.044$) times less likely to evaluate results as expected, and were 0.54 ($p = 0.025$) times less likely to accept parts into their final designs. These behaviors can be linked broadly to narrower search processes and design fixation, where instead of fixating on aspects of an external solution, an adherence to their initial ideas and internally imagined parts may occur.

3.1.3 Search pathways: linking prior behaviors with subsequent outcomes

The relationship between search behaviors is further analyzed through **search pathways**. A similar approach as used in Sec. 3.1.1 and 3.1.2 is used to determine how modality and expertise influences pathway outcomes, such as how new vs. continued searches were evaluated and how expected vs. unexpected were selected. Additional mixed effects binary logistic regression models were constructed to model whether modality and expertise differently predict how new (N = 95) and continued (N = 140) searches were evaluated and expected (N = 79) and unexpected (N = 156) search results were selected. Across these four models, modality and expertise are only found to significantly impact the evaluation of new searches. Results of the model predicting the evaluation of new searches are summarized in Table 7. Observations regarding

all definition-evaluation and evaluation–selection pathways are discussed further in this subsection.

R3.1 Search modality impacts the evaluation of new searches. The relationship between search behaviors is analyzed through search pathways related to searches made with keyword, part, and workspace inputs. As represented in Sankey diagrams shown in Figure 6, (A) definition-evaluation and (B) evaluation–selection pathways are displayed. These diagrams visually depict the average number of searches made in each pathway per designer. Associated pathway frequencies combined across all participants are shown in Table 8.

Differences between the evaluation of new vs. continued keyword, part, and workspace searches are shown in Figure 6a. New workspace compared to keyword searches were 0.91 times less likely to be evaluated as expected ($p = 0.028$). This finding is driven by the observation that only one new workspace search was evaluated as expected (Table 8). By contrast, a higher proportion of new keyword ($26/61 = 42.6\%$) and part ($7/26 = 26.9\%$) searches were evaluated as expected. As stated previously (**R2.1**), more workspace than keyword searches were evaluated as unexpected, across designers, possibly attributable to the limitations in the system's ability to retrieve expected results and the designer's ability to anticipate and understand how the system is conducting non-text-based searches.

Using workspace searches without having a clearly defined search goal may influence why the results are then evaluated as irrelevant. For instance, Example 1 in Table 3 presents an example of a new workspace search made with a vaguely expressed intent. In addition to highlighting limitations of the system discussed previously, these findings suggest that for non-text searches to be more aligned with designer expectations, further support, curation, or instruction may be necessary. This is an important result for the design of future inspirational search systems, which may leverage diverse media beyond text for queries. To understand how to encourage designers to evaluate more AI-provided results as expected and acceptable, designer rationale for following these pathways are explored in Sec. 3.2.

Once a search is made and the returned parts are evaluated as expected or unexpected, results may then either be accepted (incorporated into the participant's current design) or rejected. No significant differences were found between workspace and part searches compared to keyword searches in the evaluation of expected or unexpected results. While modality was found to affect how designers evaluate search results, it does not appear to affect how the expected results are then selected. In other words, if a search result was expected or unexpected, whether the search was made using a keyword, part, or workspace search did not significantly influence designers' acceptance or

Table 7 Binary logistic regression model predicting evaluation of new searches (n = 95)

Outcome variable	*Predictor*	*Level*	β	p	*Odds ratio* (e^{β})	*95% C.I.*
Evaluation: Expected = 1, Unexpected = 0	**Modality**	*Keyword*	(Ref)	(Ref)	(Ref)	(Ref)
		Part	−0.14	0.81	0.87	(0.28, 2.6)
		Workspace	−2.4	**0.028**	0.094	(0.005, 0.53)
	Expertise	*Professional*	(Ref)	(Ref)	(Ref)	(Ref)
		Student	−1.0	**0.033**	0.37	(0.14, 0.90)

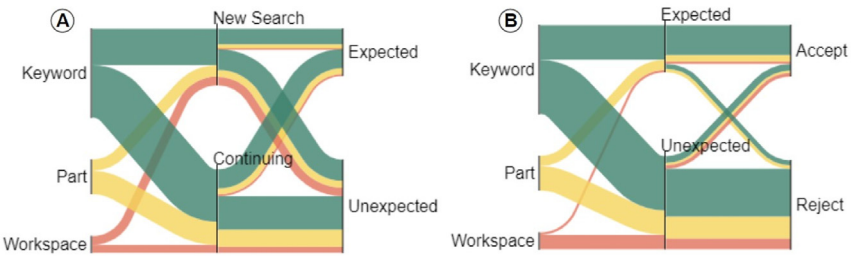

Figure 6 ***Search pathways compared across*** *keyword* ***(green), part (yellow), and workspace (orange) searches*** *linking* ***(a)*** *definition and evaluation behaviors and* ***(b)*** *evaluation and selection behaviors.*

Table 8 Summary of search pathways made using keyword, part, and workspace search inputs

Search behavior		*Search type*			*Total*
Definition	*Evaluation*	*Keyword*	*Part*	*Workspace*	*# of searches*
New	Expected	26	7	1	34
	Unexpected	35	12	14	61
Continuing	Expected	32	10	3	45
	Unexpected	56	29	10	95
Evaluation	*Selection*	*Keyword*	*Part*	*Workspace*	*# of searches*
Expected	Accept	50	11	4	65
	Reject	8	6	0	14
Unexpected	Accept	11	4	6	21
	Reject	80	37	18	135

rejection of results. The relative proportions of expected and unexpected keyword, part, and workspace searches that are accepted and rejected are shown in Figure 6b.

More surprisingly, two additional evaluation−selection pathways are notable. A small proportion of searches made with each search input that are expected are rejected, and that are unexpected are accepted. Table 8 shows that, across all participants, 8/58 (13.8%) keyword and 6/17 (35.3%) part search results evaluated as expected were rejected. Examples 2 and 3 in Table 3 illustrate these behaviors, where designers reference looking for a closer match than

what has already been found. Expected search results may encourage designers to search further, as they may consider themselves 'on the right track'. The use of slider repositioning when defining part and workspace searches can further aid this process. Another less explored and less intuitive pathway is the acceptance of unexpected stimuli, including 11/91 (12.1%) keyword, 4/37 (10.8%) part, and 6/24 (25.0%) workspace search results. Example 4 in Table 3 shows how a result from a workspace search that does not match the designer's original intention can be nonetheless useful for, e.g., introducing a design feature such as wheels to add mobility to a waste bin. These findings suggest that cognitive behaviors exist when searching that challenge designers' fixation on a given objective, and are explored further in Sec. 3.2.

R3.2 Expertise impacts the evaluation of new searches. Next, comparing definition evaluation pathways followed by students and professionals, the Sankey diagram in Figure 7a represents the average number of searches made in each pathway per designer in each group. Corresponding pathway frequencies are summarized in Table 9. The binary logistic regression model for new searches demonstrated that new searches made by students compared to professionals were 0.63 times less likely to be evaluated as expected ($p = 0.033$). Figure 7a emphasizes that professionals find more new searches provide expected results than students. On average per participant, professionals evaluated 3.0 new searches as expected, compared to 1.4 by students (see Table 9). Expressed differently, professionals evaluate, on average, 45.3% of new searches as expected, compared to 23.8% by students. No significant results are found regarding the evaluation of continuing searches.

While professionals and students do differ by the proportion of searches that are evaluated as expected and accepted (Result **R2.2**), their selection of expected and unexpected search results do not differ significantly. These relative frequencies of pathways can be compared visually in Figure 7b. Intuitively, across participants, a high proportion of results that are evaluated as expected are accepted, and unexpected results are rejected. For professionals, 41.0% of searches are evaluated as expected, 80.9% of which are accepted. Students evaluate fewer searches as expected (27.1%), but accept a relatively high proportion of these results (84.4%). Both professionals and students reject a similar percentage of searches evaluated as unexpected (81.4% and 90.7%, respectively). Therefore, although students and professionals exhibit different search evaluation and selection behaviors, they similarly evaluate expected and unexpected search results.

As noted when comparing evaluation–selection pathways across search modalities, both students and professionals also reject expected results and accept unexpected results. Only a small proportion of searches made by both participant groups are represented in these pathways. To understand why unexpected results may be accepted, examples are presented in Sec. 3.2 to

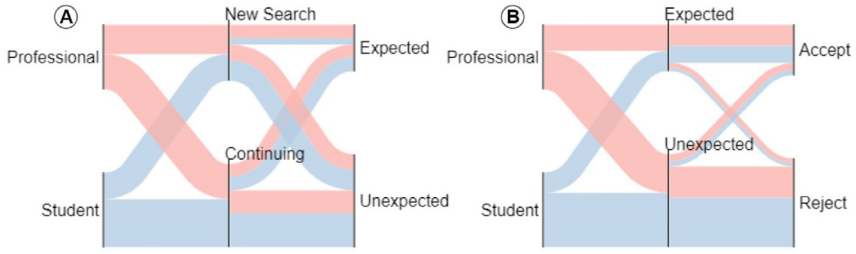

Figure 7 ***Search pathways compared across professionals (red) and students (blue)*** *linking* ***(a)*** *definition and evaluation behaviors and* ***(b)*** *evaluation and selection behaviors. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.).*

Table 9 Summary of search pathways by professional and student designers

Search behavior		*Participant group*		*Total*
Definition	*Evaluation*	*Professional (n = 8)*	*Student (n = 7)*	*# of searches*
New	Expected	24	10	34
	Unexpected	29	32	61
Continuing	Expected	23	22	45
	Unexpected	41	54	95
Evaluation	*Selection*	*Professional (n = 8)*	*Student (n = 7)*	*# of searches*
Expected	Accept	38	27	65
	Reject	9	5	14
Unexpected	Accept	13	8	21
	Reject	57	78	135

uncover rationale for following this particular pathway. This pathway, in addition to the evaluation of new search results as expected, represent desirable behaviors to better understand and encourage regarding the use of design support tools.

3.2 Designer rationale motivating search pathway outcomes

Finally, to gain further insight into specific search pathways followed by designers, the rationale provided for their evaluation and selection of search results are explored. To identify rationale, a mixed-methods approach is used where quantitative analyses of interaction and think-aloud data first enabled the isolation of individual search pathways, as fully described in Sec. 3.1.3. Qualitative insights from think-aloud data are now used to describe rationale underlying three search pathways. Two pathways with desirable outcomes are considered: when search results from a new search are evaluated as expected and when unexpected search results are accepted. Both pathways represent less explored, but desirable outcomes from interacting with the search platform. A third pathway is discussed, constituting a more frequent, but potentially less desirable outcome: the rejection of unexpected results.

3.2.1 New search results that meet designers' expectations

The first pathway of interest involves a new search for a part, for which the system retrieves results that the designer evaluates as expected. This pathway constitutes 34/235 of all searches, across participant groups and search modalities (see Tables 8 and 9). Examples to characterize this pathway are presented to understand why some searches lead to parts that do or do not match expectations to an initial search goal. We propose that both the platform's performance as well as the designer's ability to adapt their expectations to the presented stimuli are key factors enabling this process.

R4.1.1. Evaluation influenced by perception of platform performance. The first way that designers acknowledged that the search results retrieved by the platform matched their expectations was to refer to the search itself as good (e.g., "*I think the search is good*" or "*it kind of works*"), which can be linked to an assessment of the platform's performance. By contrast, their evaluation could be motivated by an assessment of the specific results returned, which might be "*the kind of thing I was looking for*", be something they liked (e.g., "*Oh, there's a lamp shade I like*"), or have particular desirable features such as the shape or size. The 'goodness' of parts can also be attributable to features of the design problem or the designer's current idea, such as a part being able to fit inside a kitchen counter, referencing the household context of the design prompt. These examples demonstrate how designers expressed their evaluation of search results as matching their expectations using rationale around platform performance and specific features and relevance of results.

R4.1.2. Designers may adapt expectations to search results. Another way that designers evaluated search results as matching their expectations was to first *adapt* their initial expectations to the parts returned, which may have appeared in a different form or context than originally searched for. This scenario differs from the evaluation of a result as unexpected, which would involve a search outcome that was incorrect, according to the designer's expectations (e.g., a flat tabletop instead of a rectangular can). Instead, these examples demonstrate scenarios where the retrieved part was 'correct' and the designer could understand why it was returned, but also identified unsuitable or irrelevant features. This pathway is explored to understand how designers rationalized overcoming these features to apply the retrieved results to their current design context. To represent this scenario with an example, two different participants conducted a new search by keyword for a "hinge", for which various hinges were returned. After one participant (P15) initially identified "*these are hinges for these* ***doors on the cabinets***", they adapted their expectations for a more contextually relevant hinge (e.g., attaching a lid to a container) to conclude "*I'm guessing that would work*". Similarly, another participant (P4) verified "*this is a hinge*", but then noted "***it's quite small*** … ***it's more of a cabinet hinge***", before conceding that they would "*take it*". In a third example, a result from the

search term "trashcan" retrieved something that "*might be a bit large for a household*" but that the participant (P3) could still "*probably work with*". Across these examples, even though the parts were what they expected (i.e., a hinge part returned for the search for "hinge"), specific features such as the size and original context of the part presented initial barriers to their acceptance. However, these examples demonstrate that designers are importantly able to overcome this initial fixation and adapt their expectations.

3.2.2 Designers' acceptance of unexpected stimuli

The second pathway for which we explore designer rationale is the selection of unexpected inspirational stimuli, corresponding to 21/235 of all searches. As we showcase through qualitative insights from the following examples, there is an opportunity for unexpected stimuli to introduce exciting and beneficial design features during ideation. Several reasons for accepting an unexpected result were found including: (1) it introduced a desirable, but unanticipated design feature, (2) it fulfilled a searched for purpose, in a different way, and (3) the designer satisfied for a result, even though it did not meet their expectations.

R4.2.1. Unexpected stimuli introduce potentially desirable features. The first way designers expressed rationale for selecting an unexpected result retrieved by the search platform was that it introduced a desirable, but previously unanticipated feature to their concept. In two cases, designers were inspired to add wheels to their designs, though this is not what they initially sought from their search. Participant P8, looking for different forms of containers through a part-based search with high functional similarity and low appearance similarity to a container lid, received parts including the set of wheels shown in Figure 8a. These were returned by the search tool because lids and wheels are visually dissimilar but share a common functional context in object assemblies including containers. Discovering the wheels, participant P8 noted: "*Well now that I see it, I think it may be a good idea to have the unit movable, so I think castors would be something useful*". The resulting influence on their design can be seen in Figure 8b, displaying that the wheels were subsequently added to the base of their disposal unit.

In a second instance, participant P7, when looking for "*something similar to this drawer*" using a workspace-based search, was returned chair wheels (Figure 9a). The search tool, recognizing visual similarity of the drawer to the seat in the chair assembly, returned chair wheels due to their shared context with the seat. After first remarking, "*well that's kind of funny*", the chair wheels were added to their design (Figure 9b) after acknowledging, similar to P8: "*Now we can add wheels to this and make it mobile, which is good!*"

In both examples, retrieved wheels introduced an unanticipated feature to their designs, i.e., mobility. In the first example, wheels from an analogically "near-field" (as defined by Fu et al. (2013b)) object assembly (a different kind of container) were added, which may represent a more obvious transfer of unexpected stimuli to the design. The second example is striking as it demonstrates how even unintentional stimuli from a "far-field" domain (a chair) can be effectively applied towards introducing a desirable, but unanticipated feature to the design. The use of contextually unrelated stimuli is also relevant to the next rationale discussed.

R4.2.2. Unexpected stimuli differently fulfill the same searched for purpose. The second rationale designers provided for using an unexpected stimulus was that it fulfilled the same purpose originally intended, but in a different way. Participant P4, upon retrieving three tabletop results (e.g., Figure 10a) when searching for a lid to place on a rectangular trashcan found that "*Nonetheless, it's actually fitting what I'm looking for exactly*". In this example, although the object did not match what was searched for, its visual form suited the designer's needs for a cover they could scale to the size of their trashcan. In a similar example, Participant P7 searched for a "can" and was given a round base of a candle holder, as shown in Figure 10b. While expressing that this is not what they were looking for, and that it was at the incorrect scale, they also stated, "*This one is maybe promising, I can maybe make it bigger … this looks like it has an opening*". Despite the size of the result, an acknowledged ability to scale it to the correct size made it useable to the designer. Finally, when looking for cylindrical shapes, Participant P14 was returned a chair seat (e.g., Figure 10c). This result was identified as being potentially useful because reorientation could be used such that, "*worst case, I can flip it … if I don't find anything, I can work with this shape which is resembling something that I might be looking for*." Object transformations, including rescaling and reorientation, were thus identified as methods enabling the use of unexpected parts to fulfill designers' intended purposes.

R4.2.3. Designers satisfied for unexpected stimuli. A final reason designers expressed for accepting unexpected stimuli was as a result of satisficing for a part. Two distinct scenarios were observed: in the first, designers' search results included previously rejected parts. Encountering these may have strengthened the belief that a more relevant match did not exist in the database. Secondly, even when acknowledging that a result is "*not quite what I was looking for*" (P15), the result was accepted. These examples suggest that designers can tolerate an acceptable threshold of accuracy when using inspiration-retrieval tools.

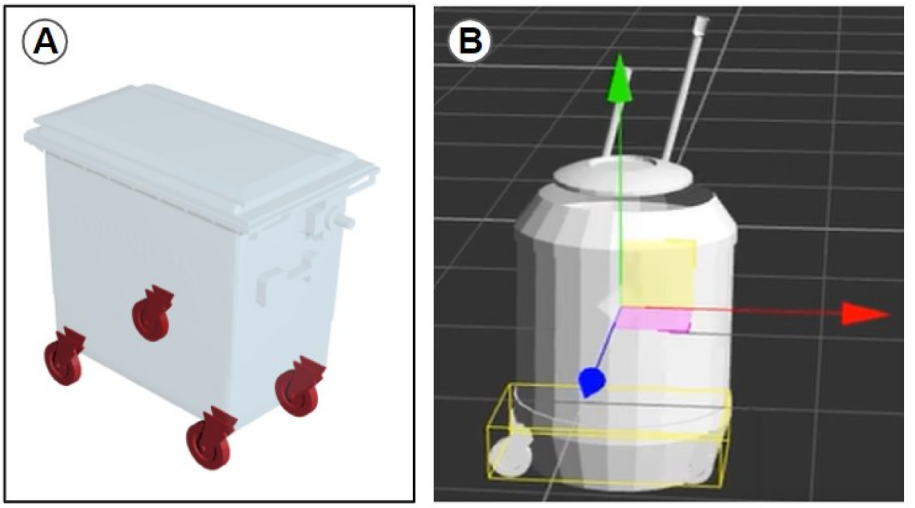

Figure 8 ***Example of unexpected results introducing an unanticipated desirable feature (P8)****:* ***(a)*** *Unexpected wheel results returned by search platform and* ***(b)*** *addition of part to P8's design*

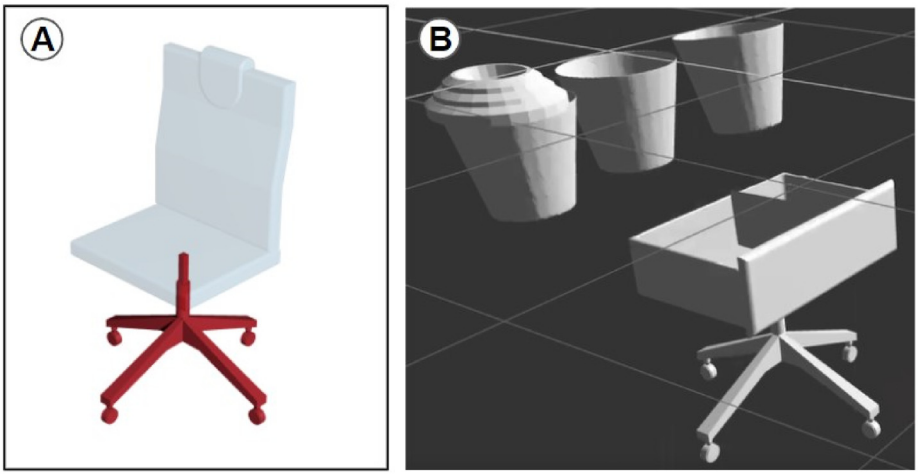

Figure 9 ***Example of unexpected results introducing an unanticipated desirable feature (P7)****:* ***(a)*** *Unexpected wheel results returned by search platform and* ***(b)*** *addition of part to P7's design*

3.2.3 Designers' rejection of unexpected results

The most frequent pathway designers followed was the rejection of unexpected search results, accounting for 134/235 searches. While beneficial outcomes of unexpected stimuli were observed, it is desirable for more results to meet expectations and be accepted, and thus important to uncover rationale for this pathway. Of these searches, 59 results were not evaluated with accompanying verbal data, but classified as unexpected if results were then rejected (as defined

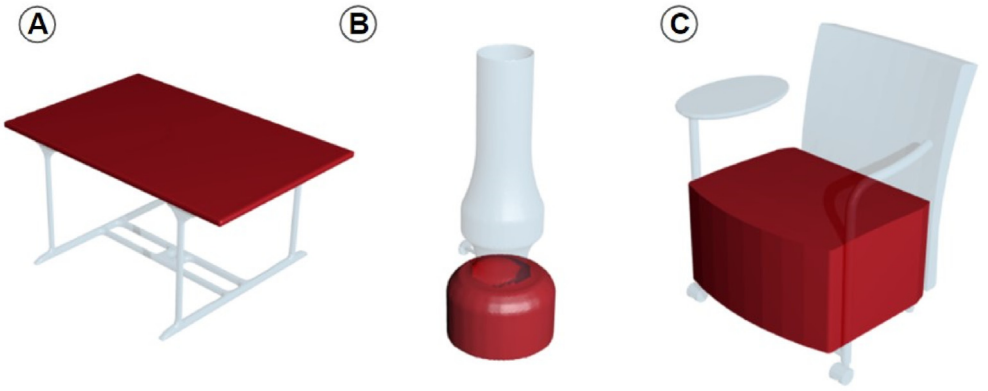

Figure 10 ***Examples of unexpected results that fulfill purposes of intentionally searched for parts****:* ***(a)*** *Tabletop scaled down to fit top of trashcan* ***(b)*** *Candle holder base scaled up to serve as can, and* ***(c)*** *Chair seat reoriented to container*

in Table 2). Of the remaining 75 searches, designers stated or described why the results did not meet expectations before not engaging with results further. When describing why results were unexpected and then rejected, two main reasons emerged, which can help improve AI-based support systems.

R4.3.1. Designers anticipated specific results in mind. Designers provided rationale for their evaluation of results as unexpected and rejection of results by indicating that their initial intention was not met. Most results were evaluated as not meeting the specific intention of the designer by being either "wrong" or "close", both prompting additional searches. In one notable example, when searching by keyword for a "trashcan", participant P5 stated "*Ok, it's not what I was thinking, but that is a trashcan, for sure*". Thus, even if the search provided a correct outcome, if a designer's goal is specific in their mind, results may still be rejected. This specificity of imagined results may influence the selection of results since, accounting for Result **R4.2.3**, designers were also observed to satisfice for and accept less desirable results.

R4.3.2. Limitations of platform and its expected use. While expectedness of results could be attributed to good platform performance (Result **R4.1.1**), unexpectedness could result from not understanding how the platform operates. Participant P8, for example, stated "*I can't really figure out how this is functionally similar or how the software determines that*" or for a different search, "*I'm trying to figure out why that might have happened*". Evaluation of retrieved results is connected to understanding how the platform functions and can impact how the examples provided are perceived and used. This finding is especially relevant when engaging with novel AI-based systems, which may not be familiar to users. Other reasons expressed by the designers in our study refer to specific features and limitations of the platform used, which may not be as generalizable. These include the platform's tendency to retrieve the same

results multiple times (when previously rejected) or the missing support for general shapes and forms as opposed to specific objects.

4 Discussion

This paper investigates how designers search for inspirational stimuli when using an AI-enabled multi-modal search platform. In the design exploration task conducted, participants with either novice (graduate students) or expert (working professionals) levels of design experience searched for 3D-model parts using three modalities of search to inspire solutions to a given design challenge. By eliciting think-aloud descriptions of their interactions with the search platform, further insight into their definition, evaluation, and selection of the retrieved stimuli, and the rationale underlying these behaviors, are studied. Revisiting the research questions initially posed to introduce the aims of the present work, the main contributions made are summarized and discussed in this section.

4.1 Search input modalities result in different search outcomes

The first comparison made in this work is of the use of different search modalities to support search activities, behaviors, and pathways. Search activity was found to differ across designers, where keyword search was associated with significantly higher frequency than part and workspace searches. Differences in how designers evaluated search results can help explain the lower frequency of workspace searches made: across designers, workspace compared to keyword searches had a higher likelihood of being evaluated as unexpected. This difference can be ascribed to limitations in the search platform in recognizing the designer's search intent, as well as the designer's ability to define and expect what they were looking for when using a less intuitive search modality. In early observations about example or image-based search, Hearst (2009) identified a limitation in the searcher being required to know about the visual properties of the image searched for, which can limit search for new images. Similarly, searching with workspace inputs that rely on appearance similarity measures may produce results that are difficult to anticipate.

Through an examination of search pathways, we further demonstrate how the evaluation of workspace search results as unexpected is especially true for new searches. When continuing to search for a desired part, the same effect of modality on evaluation of results was not observed, such that neither continuing part nor workspace searches were significantly more likely to be unexpected. Continuing a search with any input may be useful during search. Sarkar and Chakrabarti discuss how stimuli referred to as "triggers" can influence designers' search of the solution space (Sarkar & Chakrabarti, 2008). Referencing O'Day and Jeffries (O'Day & Jeffries, 1993), one trigger that may motivate a switch in search strategy is the encounter with something that

introduces a new way of thinking about the problem at hand Continued search using any input can facilitate encounters with stimuli that "trigger" new searches.

4.2 Expertise impacts designers' search for inspirational stimuli

Secondly, we examined how expertise level may influence how designers search. While professionals and students were not found to differ by search activity, i.e., the frequency of use of keyword, part, and workspace searches, they did differ by search behaviors followed. Expertise is suggested not to affect how often search modalities are used, but how search results are evaluated and selected. Notably, students were found to be more likely to evaluate results as unexpected, and to ultimately reject more results from inclusion in their designs. These behaviors suggest that students may fixate more on finding their originally intended results and demonstrate less openness to incorporating unexpected parts into their design ideas. Students are expected to have less experience with design and working with AI-assisted design tools, which may make them more prone to relying on their own experience and internal stimuli (Gonçalves et al., 2016). Less experience also affects novice designers' tendency to reflect on how inspiration sources can impact their designs, thereby limiting the adaption of unexpected stimuli to their designs (Gonçalves et al., 2013). These findings also reinforce Gonçalves et al.'s results on expert designers' greater ability to absorb and adapt detailed information from stimuli compared to novices (2014), and Cross's argument that experts more readily seek a diversity of information to support their design process (2004).

Through investigating specific search pathways, such as the relationship between how new vs. continuing searches were evaluated, professionals were found to evaluate more new search results as expected than students. This can be attributed to professional designers exhibiting broader expectations for parts, allowing them to consider more results as expected without continued search and exploration. This interpretation supports previous work by Gonçalves et al. (2014), Cross (2004), and Cai et al. (Cai, Do, & Zimring, 2010) that professional designers seek to extract detailed information from diverse inspirational sources. Thus, a relationship between their initial search inputs and the retrieved results may have been more immediately inferred. Our findings contrast professional designers' broad expectations with novice designers' relatively narrower expectations. Relatedly, Cai et al.'s findings suggest that novice designers found value in stimuli for their connection to familiar knowledge. If search results did not immediately meet expectations, designers' ability to recognize the connection between retrieved results to their initial search input may have been limited. Students thus proceeded to conduct more continued searches, on average. While the aim of this work was to

specifically investigate search processes, these findings can be more broadly applied to the role of expertise on the ability to use and extract meaning from inspirational design stimuli.

4.3 Rationale underlying less explored search pathways

An interesting finding in this paper was the uncovering of search results that were evaluated by designers as expected or unexpected. Think-aloud transcription data was examined to understand the rationale behind the evaluation of search results and the uses of unexpected stimuli. The evaluation of new search results as expected was linked to a positive assessment of the performance of the search platform itself or of specific features of the retrieved results. As Cascini et al. (2010) propose, the consideration of expected behavior of products is needed from both the perspectives of the product user and designer. As discovered in Result **R4.1.2**, initial fixation to specific part features or object contexts could importantly be overcome by adapting expectations. This may be especially true when working in a CAD environment, in contrast to a physical environment, where parts may be easily adjusted in scale and isolated from their original context.

Several examples from this study challenge whether the aim of the search platform should be to support the retrieval of inspirational stimuli that users interacting with it expect. Indeed, desirable design outcomes, such as the introduction of new design features during idea generation, can occur as a result of the discovery of initially unintended search results. Given the large proportion of results that were not what designers expected (156/235), 135 of which were rejected and unused towards continued idea generation, one area for further exploration is how to encourage designers to similarly leverage information when derived unexpectedly. Through examples underlying Result **R4.2.2**, object transformations were found to assist designers' ability to discover usefulness from unexpected sources of inspiration. Reorientation has specifically been proposed in prior research as a strategy to aid creative object reuse (Olteteanu & Shu, 2018). Damen and Toh (Damen & Toh, 2019) have found that information designers evaluate as helpful is not necessarily used during idea generation. They additionally suggest that designers are able to effectuate readily available information sources (i.e., make use of existing resources), even those that may not evidently influence an outcome (Damen & Toh, 2021). These strategies may help overcome the motivations designers expressed for rejecting unexpected results, explored in Sec. 3.2.3, by overcoming design tool limitations and specific expectations held in mind. These findings recommend that, while continuing to improve computational definitions of similarity relationships according to designers' needs and expectations is important, methods to promote designers' adaptation of expectations and ready use of available stimuli can also be beneficial.

4.4 Limitations and future work

This paper presents the results of a design exploration task in which participants, consisting of designers with a range of design experience, interacted with a multi-modal search platform. Methodologically, three main limitations and opportunities for future work exist. Firstly, across two studies (the first described in prior work (Kwon et al., 2022) and the second in the present), participants of both novice and expert level design experience found the search platform's novel modalities difficult to use. Despite some observed benefits of encountering unexpected results, continued work in the development of this and other search platforms can be done towards improving retrieval accuracy. This may be achieved through the exploration of different sets of inspirational stimuli and definitions of appearance and function-based similarity that are more intuitive to designers. More generally, the results presented in this paper, especially regarding search activities and behaviors, may be heavily influenced by features of the search platform used and the design stimuli returned. Despite this limitation, we present findings that can be adaptable to use of other design tools, such as comparisons between novice and expert designers. Through investigation of pathways, we also explore how search results are engaged irrespective of their content. Secondly, in the design exploration task completed, approx. 15 min. were allotted to search for parts. While most participants reached an impasse in their search and design activity by this time, prior work by Moss et al. (2011) has shown how incidental information provided at the point of impasse can be beneficial for problem solving. Continued design ideation after receiving new stimuli following an impasse can therefore be studied. Finally, participants were tasked with searching for parts to inspire solutions to the given design problem. These instructions were specified to promote search activity, which was the focus of the present work, rather than to encourage and assess idea generation. Thus, the extent to which designers worked on developing a single or multiple final design ideas varied, limiting our ability to assess the impact of stimuli on design activity. Future work can link how the stimuli discovered as a result of different search processes and modalities can contribute to specific design outcomes. For instance, unexpected search results may lead to more novel design features.

5 Conclusion

The main contribution made by this work is to deepen an understanding of how designers search for inspirational stimuli. This aim was achieved through a think-aloud design exploration task where designers used an AI-enabled multi-modal search platform developed for this task. Search modality and designer expertise were factors found to influence the process of searching for design inspiration. By contrasting the uses of a more familiar mode of search (by keyword) with more novel modes of search (by 3D-model part and 3D-modeling workspace inputs), we found that modality affected how designers interacted with retrieved results. When searching by keyword, more

results were expected than workspace search results, and accepted than part search results. While these differences can be partially attributable to limitations of the current system, we suggest that designers may have difficulty defining their search intent and forming expectations for results when searching based on visual and functional relationships. Improved understanding of how designers perceive and seek inspiration in terms of these less explored modalities can help support the further development of multi-modal design tools. The role of expertise was also examined by comparing behaviors of student and professional designers. Professionals generally had broader expectations for search outcomes than students, who tended to reject and evaluate more results as not meeting their initial expectations. Increased design expertise was associated with greater openness to potential sources of inspiration and reduced fixation to intended results. This difference reveals both how expertise influences the use of increasingly prevalent AI-enabled design tools as well as how the process of becoming inspired may engage prior experience. Search modality and expertise were factors found to impact design behavior when engaging with an AI-enabled platform for inspiration discovery. Our study supports continued research to understand and improve designers' interactions with AI-based design tools and the relationship between the inspiration designers seek and effectively use.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the National Science Foundation under grant 2145432 − CAREER. The authors would like to thank Dr. Forrest Huang for his contribution to the multi-modal search platform employed in this work.

References

Arora, R., Darolia, I., Namboodiri, V. P., Singh, K., & Bousseau, A. (2017). SketchSoup: Exploratory ideation using design sketches. *Computer Graphics Forum, 36*, 302−312.

Athukorala, K., Głowacka, D., Jacucci, G., Oulasvirta, A., & Vreeken, J. (2016). Is exploratory search different? A comparison of information search behavior for exploratory and lookup task. *J. the Association for Information Science and Technology, 67*, 2645−2651.

Biskjaer, M. M., Christensen, B. T., Friis-Olivarius, M., Sille, J., Abildgaard, J., Lundqvist, C., et al. (2020). Fixation or inspiration? A meta-analytic review of the role of examples on design processes. *International Journal of Technology and Design Education, 30*, 101−125.

Borgianni, Y., Rotini, F., & Tomassini, M. (2017). Fostering ideation in the very early design phases: How textual, pictorial and combined stimuli affect creativity. *Proc. of the 21st International Conference on Engineering Design, ICED17*, 139–148.

Botella, M., Zenasni, F., & Lubart, T. (2018). What are the stages of the creative process? What visual art students are saying. *Frontiers in Psychology, 9*.

Cai, H., Do, E. Y-L., & Zimring, C. M. (2010). Extended linkography and distance graph in design evaluation: An empirical study of the dual effects of inspiration sources in creative design. *Design Studies, 31*, 146–168.

Cascini, G., Del Frate, L., Fantoni, G., & Montagna, F. (2010). Beyond the design perspective of gero's FBS framework. In J. S. Gero (Ed.), *Design Computing and Cognition '10* (pp. 77–96), Springer.

Chan, J., Dow, S. P., & Schunn, C. (2015). Do the best design ideas (really) come from conceptually distant sources of inspiration? *Design Studies, 36*, 31–58.

Chan, J., Fu, K., Schunn, C., Cagan, J., Wood, K., & Kotovsky, K. (2011). On the benefits and pitfalls of analogies for innovative design: Ideation performance based on analogical distance, commonness, and modality of examples. *Journal of Mechanical Design, 133*, 081004.

Chaudhuri, S., & Koltun, V. (2010). Data-driven suggestions for creativity support in 3D modeling. *ACM Transactions on Graphics, 29*, 1–10.

Cross, N. (2004). Expertise in design: An overview. *Design Studies, 25*, 427–441.

Damen, N., & Toh, C. (2019). Looking for inspiration: Understanding the information evaluation and seeking behavior of novice designers during creative idea generation. *Proc. of the Design Society: International Conference on Engineering Design, 1(1)*, 1793–1802.

Damen, N., & Toh, C. (2021). Reflections on designing in the wild: How theories of design information manifest in practice. In *Proc. ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Virtual, Online. Aug. 17–19* V006T06A041.

Do, E. Y-L (2005). Design sketches and sketch design tools. *Knowledge-Based Systems, 18*, 383–405.

Eckert, C., & Stacey, M. (2000). Sources of inspiration: A language of design. *Design Studies, 21*, 523–538.

Fu, K., Cagan, J., Kotovsky, K., & Wood, K. (2013a). Discovering structure in design databases through functional and surface based mapping. *Journal of Mechanical Design, 135*, 031006.

Fu, K., Chan, J., Cagan, J., Kotovsky, K., Schunn, C., & Wood, K. (2013b). The meaning of "Near" and "Far": The impact of structuring design databases and the effect of distance of analogy on design output. *Journal of Mechanical Design, 135*, 021007.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45–56.

Goel, A. K., Vattam, S., Wiltgen, B., & Helms, M. (2012). Cognitive, collaborative, conceptual and creative — four characteristics of the next generation of knowledge-based CAD systems: A study in biologically inspired design. *Computer-Aided Design, 44*, 879–900.

Gonçalves, M., Cardoso, C., & Badke-Schaub, P. (2013). Through the looking glass of inspiration: Case studies on inspirational search processes of novice designers. In *Proc. of the IASDR conference 2013, Tokyo, Japan, Aug. 26-30*.

Gonçalves, M., Cardoso, C., & Badke-Schaub, P. (2014). What inspires designers? Preferences on inspirational approaches during idea generation. *Design Studies, 35*, 29–53.
Gonçalves, M., Cardoso, C., & Badke-Schaub, P. (2016). Inspiration choices that matter: The selection of external stimuli during ideation. *Design Science, 2*.
Goucher-Lambert, K., & Cagan, J. (2019). Crowdsourcing inspiration: Using crowd generated inspirational stimuli to support designer ideation. *Design Studies, 61*, 1–29.
Goucher-Lambert, K., Gyory, J. T., Kotovsky, K., & Cagan, J. (2020). Adaptive inspirational design stimuli: Using design output to computationally search for stimuli that impact concept generation. *Journal of Mechanical Design, 142*, 091401.
Goucher-Lambert, K., Moss, J., & Cagan, J. (2019). A neuroimaging investigation of design ideation with and without inspirational stimuli—understanding the meaning of near and far stimuli. *Design Studies, 60*, 1–38.
Han, J., Shi, F., Chen, L., & Childs, P. R. (2018). "The combinator–A computer-based tool for creative idea generation based on a simulation approach". *Design Science, 4*.
Hearst, M. (2009). *Search user interfaces*. Cambridge University Press.
Herring, S. R., Chang, C. C., Krantzler, J., & Bailey, B. P. (2009). "Getting inspired! Understanding how and why examples are used in creative design practice". In *Proc. Of SIGCHI Conference on human Factors in computing systems, CHI'09* (pp. 87–96). New York, NY, USA: Association for Computing Machinery.
Hua, M., Han, J., Ma, X., & Childs, P. (2019). "Exploring the effect of combinational pictorial stimuli on creative design performance". *Proc. of the Design Society: International Conference on Engineering Design, 1*(1), 1763–1772.
Jansson, D., & Smith, S. (1991). Design fixation. *Design Studies, 12*, 3–11.
Jiang, S., Hu, J., Wood, K. L., & Luo, J. (2022). Data-driven design-by-analogy: State-of-the-Art and future directions. *Journal of Mechanical Design, 144*, 020801.
Jiang, S., Luo, J., Ruiz-Pava, G., Hu, J., Christopher, L., & Magee. (2021). Deriving design feature vectors for patent images using convolutional neural networks. *Journal of Mechanical Design, 143*, 061405.
Jiang, S., Luo, J., Ruiz-Pava, G., Hu, J., & Magee, C. L. (2020). A convolutional neural network-based patent image retrieval method for design ideation. In *Proc. ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Virtual, Online. Aug. 17–19* V009T09A039.
Kazi, R. H., Grossman, T., Cheong, H., Hashemi, A., & Fitzmaurice, G. (2017). DreamSketch: Early stage 3D design explorations with sketching and generative design. In *Proc. Of the 30th annual ACM symposium on user interface software and technology, UIST '17* (pp. 401–414). New York, NY, USA: Association for Computing Machinery.
Kittur, A., Yu, L., Hope, T., Chan, J., Lifshitz-Assaf, H., Gilon, K., et al. (2019). Scaling up analogical innovation with crowds and AI. *Proceedings of the National Academy of Sciences, 116*, 1870–1877.
Kwon, E., Huang, F., & Goucher-Lambert, K. (2022). Enabling multi-modal search for inspirational design stimuli using deep learning. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 36*, E22.
Linsey, J. S., Clauss, E. F., Kurtoglu, T., Murphy, J. T., Wood, K. L., & Markman, A. B. (2011). An experimental study of group idea generation

techniques: Understanding the roles of idea representation and viewing methods. *Journal of Mechanical Design, 133*, 031008.
Linsey, J. S., Wood, K. L., & Markman, A. B. (2008). Modality and representation in analogy. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 22*, 85–100.
Lucero, A. (2012). Framing, aligning, paradoxing, abstracting, and directing: How design mood boards work. Proc. of the designing interactive. In *systems conference, DIS'12* (pp. 438–447). New York, NY, USA: Association for Computing Machinery.
Luo, J., Sarica, S., & Wood, K. L. (2021). Guiding data-driven design ideation by knowledge distance. *Knowledge-Based Systems, 218*, 106873.
Malaga, R. A. (2000). The effect of stimulus modes and associative distance in individual creativity support systems. *Decision Support Systems, 29*, 125–141.
Marchionini, G. (2006). "Exploratory search: From finding to understanding". *Communications of the ACM, 49*, 41–46.
Menezes, A., & Lawson, B. R. (2006). How designers perceive sketches. *Design Studies, 27*, 571–585.
Moss, J., Kotovsky, K., & Cagan, J. (2011). "The effect of incidental hints when problems are suspended before, during, or after an impasse". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 140–148.
Murphy, J., Fu, K., Otto, K., Yang, M., Jensen, D., & Wood, K. (2014). "Function based design-by-analogy: A functional vector approach to analogical search.". *Journal of Mechanical Design, 136*, 101102.
Nagel, J. K., & Stone, R. B. (2012). A computational approach to biologically inspired design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 26*, 161–176.
O'Day, V. L., & Jeffries, R. (1993). Orienteering in an information landscape: How information seekers get from here to there. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, CHI '93* (pp. 438–445). Association for Computing Machinery.
Olteteanu, A. M., & Shu, L. H. (2018). Object reorientation and creative performance. *Journal of Mechanical Design, 140*, 031102.
Sarkar, P., & Chakrabarti, A. (2008). The effect of representation of triggers on design outcomes. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 22*, 101–116.
Sartori, J., Pal, U., & Chakrabarti, A. (2010). A methodology for supporting 'transfer' in biomimetic design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 24*, 483–506.
Siangliulue, P., Chan, J., Gajos, K. Z., & Dow, S. P. (2015). Providing timely examples improves the quantity and quality of generated ideas. In *In Proc. of the 2015 ACM SIGCHI Conference on Creativity and Cognition, C&C '15* (pp. 83–92). Association for Computing Machinery.
Sio, U. N., Kotovsky, K., & Cagan, J. (2015). Fixation or inspiration? A meta-analytic review of the role of examples on design processes. *Design Studies, 39*, 70–99.
Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research and Evaluation, 9*, 4.
Sutcliffe, A., & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers, 10*, 321–351.
Toh, C. A., & Miller, S. R.. (2014). The impact of example modality and physical interactions on design creativity. *Journal of Mechanical Design, 136*, 091004.

Tseng, I., Moss, J., Cagan, J., & Kotovsky, K. (2008). The role of timing and analogical similarity in the stimulation of idea generation in design. *Design Studies, 29*, 203–221.

Vasconcelos, L. A., Cardoso, C. C., Sääksjärvi, M., Chen, C-C., & Crilly, N. (2017). Inspiration and fixation: The influences of example designs and system properties in idea generation. *Journal of Mechanical Design, 139*, 031101.

Wallace, S., Le, B., Leiva, L. A., Haq, A., Kintisch, A., Bufrem, G., et al. (2020). *"Sketchy: Drawing Inspiration from the Crowd." Proc. of the ACM on Human-Computer Interaction, 4*, 1–27.

Zhang, Z., & Jin, Y. (2020). An unsupervised deep learning model to discover visual Similarity Between Sketches for Visual Analogy Support. In *Proc. ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Virtual, Online. Aug. 17–19* V008T08A003.

Zhang, Z., & Jin, Y. (2021). Toward computer aided visual analogy support (CAVAS): Augment designers through deep learning. In *Proc. ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Virtual, Online. Aug. 7–19* V006T06A057.

Vattam, S., Wiltgen, B., Helms, M. E., Goel, A. K., & Yen, J. (2011). DANE: Fostering creativity in and through biologically inspired design. In T. Taura, & Y. Nagai (Eds.), *Design creativity 2010* (pp. 115–122). London: Springer.

Evaluating the efficacy and effectiveness of design methods: A systematic review and assessment framework

Philip Cash, Northumbria University, UK

Jaap Daalhuizen, DTU Technical University of Denmark, Denmark

Paul Hekkert, Delft University of Technology (TU Delft), Netherlands

The increasingly transdisciplinary context of design, where designers collaborate with other disciplinary and domain experts, means there is a growing need to evidence the effectiveness of design methods. We address this need in two ways. First, we propose a 'chain of evidence', from motivation to claims, operationalising this in a systematic assessment framework. Second, we systematically review current design method research. Our results reveal that while all links in the chain of evidence are reported across the literature and best practices can be identified, no individual paper either reports all links or consistently achieves best practice. Our framework and results demonstrate the need for standards of evidence in this area, with implications for design method research, development, education, and practice.

Keywords: design methods, design tools, design methodology, design research, systematic review

The increasingly transdisciplinary context of design, where designers collaborate with other disciplinary and domain experts, means there is a growing need to evidence the effectiveness of design methods. Key questions are, what do design methods claim to achieve and what evidence are these claims based on? This comes in addition to method's traditionally central role in design; helping to shape, describe, teach, and explain our discipline (Roozenburg & Eekels, 1995; van Boeijen, Daalhuizen, & Zijlstra, 2020).

Design methods, and their associated claims of improved performance, have reached this status by translating research insights and best practices into real-world impact via education and practice (Blessing & Chakrabarti, 2009; Cantamessa, 2003; Daalhuizen, Person, & Gattol, 2014). The proposal and study of new methods arguably forms *the* central pillar of design research

Corresponding author:
Philip Cash
philip.cash@northumbria.ac.uk

www.elsevier.com/locate/destud
0142-694X *Design Studies* **88** (2023) 101204
https://doi.org/10.1016/j.destud.2023.101204

(impact). Yet, in contrast with almost all similarly important research impact mechanisms (e.g., interventions in education (Levin & O'Donnell, 1999) or health (Gottfredson et al., 2015; Grimes & Schulz, 2002)), there is no consolidated procedure for assessing design methods and their supporting evidence (Gray, 2022). This undermines the very foundations of design research impact and design practice's credibility in contributing to social transformation.

Design methods capture key procedural knowledge to provide '*a formalised representation of a design activity, which functions as a mental tool to* support *designers in achieving a goal, in relation to the circumstances and resources available*' (Daalhuizen, Timmer, van der Welie, & Gardien, 2019). In this form, methods significantly contributed to the emergence of design research as a field in the 1960s and have continued to grow in importance across domains (Jones, 1977; Roozenburg & Eekels, 1995; van Boeijen et al., 2020). However, they have also formed the subject of heated debate, exacerbated by conflicting accounts of how methods impact practice (if at all) (Daalhuizen & Cash, 2021; Dorst, 2008; Jones, 1992; Wallace, 2011), leading to a critical gap between the importance and credibility of design methods. This gap is rooted in the multifaceted nature of rigorous method evaluation (Daalhuizen & Cash, 2021, Figure 3; Gericke, Eckert, & Stacey, 2017, sec. 4.4) and poses a grave challenge to design research (Lloyd, 2019; Meyer & Norman, 2020), as well as leaving numerous disciplines — ranging from health to engineering — reliant on methods built on ambiguous evidence and lacking transparency. Thus, there is a vital need to better understand how design methods and their associated claims can be assessed.

To address this need we build on the recent work of Daalhuizen and Cash (2021), who defined a basic understanding of 'good' method content,[1] and thus provide a foundation for assessing methods more generally. Taking this as a starting point, we first develop a framework for assessing method development, reporting, claims, and supporting evidence, before using this to systematically review and analyse current method research. Throughout, we focus on methods that involve human activities or interaction and thus cannot be evaluated in isolation, in contrast to more technical and/or computational tools whose efficacy can be directly assessed. Our assessment framework and review provide a basis for grounding the current debate on method credibility. This substantially extends research towards a wider *Theory of Design Methods* and has significant implications for method research, development, education, and practice.

1 Background

To assess design methods, it is first necessary to clarify a common understanding of what methods in design are and subsequently how research proposing methods and reporting their associated claims can be evaluated.

1.1 Methods in design

Design methods broadly serve to embody understanding of design work (Bucciarelli, 1994; Stappers & Sanders, 2005), direct design outcomes (Araujo, 2001), and shape design education, skill development, and practices (Andreasen, 2011; Kunrath, Cash, & Kleinsmann, 2020). They do this by supporting belief formation and cognition via an interaction between method content and method user, which we call 'method use' (Daalhuizen & Cash, 2021). This interaction can range from almost complete offloading (e.g., via an algorithmic, computer supported method such as a design structure matrix (Pektaş & Pultar, 2006)) to cognitively intensive (e.g., via a heuristic principle such as satisficing (Simon, 2019)), or any combination thereof (Daalhuizen, 2014). Further, the specific focus of this support and the associated claims is as varied as design itself (Cash, Valles Gamundi, Echstrom, & Daalhuizen, 2022; Kumar, 2013; van Boeijen et al., 2020). For example, claims can include statements about efficacy (e.g., the brainstorming method leads to a greater number of and more diverse ideas compared to using competing methods or no method at all) or effectiveness (e.g. the brainstorming method can be successfully used by multidisciplinary teams in corporate organizations to generate ideas), as well as other direct and indirect outcomes (e.g. the brainstorming method leads to better ideas and also shared understanding in multidisciplinary teams). Thus, the interaction between method content and method user (for simplicity, and to avoid confusion with product user, we refer to the method user throughout as 'designer') provides a common foundation for understanding method impact.

In this context, it is possible to trace a logical link between a method's support for individual belief formation and cognition to at least: i) taskwork processes and effects on outputs and artefacts (e.g. by fostering creativity (Chulvi, Mulet, Chakrabarti, López-Mesa, & González-Cruz, 2012)), and simultaneously ii) teamwork processes and effects on team and environment (e.g. by fostering shared understanding or affect (Vaajakallio & Mattelmäki, 2014)). As highlighted above, methods have been claimed to impact all aspects of the design process, with the ultimate contribution to more successful design work and outcomes (Lewrick, Link, & Leifer, 2018, 2020). However, current approaches to evaluating method impact typically focus on method's effects at the artefact (outcome) or organisation level (Andreasen, Thorp Frey & Dym, 2006; Hansen, & Cash, 2015), and thus provide little insight into the relationship between the basic method/designer interaction and the ultimate impact. This is particularly problematic because design methods are typically heuristic in nature, and thereby enhance chances of producing desirable outcomes, rather than guaranteeing them. This means that good methods that are used properly might be discarded based on undesirable outcomes.

This disconnect between method function and impact evaluation lies at the heart of methodological debate; perhaps best illustrated by Christopher Alexander's dual role as both a founding father of the methods movement and one of its greatest critics (Alexander, 1971). Central to his criticism was the overemphasis on method development at the expense of understanding of how methods work and — most importantly — how they contribute to better design practices and real-world impact. As illustrated in this section, this challenge is as vital today as it was in Alexander's time. Hence, there is a pressing need to better align understanding of method function and impact in the discussion and assessment of methods. This is essential to fostering a more mature and robust research culture around method development, testing, and dissemination.

1.2 Evaluating methods

Current evaluation efforts largely focus on criteria related to consistency and overall impact in terms of design outcomes (Frey & Dym, 2006; Vermaas, 2016). This neglects the key method/designer interaction (see Section 1.1) (Daalhuizen & Cash, 2021; Daalhuizen et al., 2019). For example, a prototypical evaluation framework in the current design literature is the validation square (Seepersad et al., 2006; Vermaas, 2016). This deals with the internal coherence and consistency of method content in terms of evaluating algorithm-like procedures that require information and resources to be processed to produce design outputs, independent of both designer and purpose. While internal consistency is essential to methods' function, the recent work of Daalhuizen and Cash (2021) demonstrates that content is processed by designers with respect to a method's purpose or goal, and therefore content, purpose, and designer response are related. Hence, Daalhuizen and Cash (2021) expand the scope of method evaluation by treating both the internal conceptual coherence of a method and the interaction between the method and the designer.

Daalhuizen and Cash (2021) operationalise this understanding in four key properties of 'good' method content. These broadly align with conceptualisations of 'good' theory and efficacy (Flay et al., 2005; Wacker, 2008), and 'good' artefacts and effectiveness (Araujo, 2001; Daalhuizen, 2014), and include (from Daalhuizen and Cash (2021, Figure 3)).

- **Defined**: The major content variables are logically complete, coherent (i.e. not conflicting), and unambiguously described; and the domain of operation is clear i.e. designers understand in what context(s) the method will perform as described;
- **Predictable**: The internal structure of the method is understandable and predictable i.e. designers can predict how altering one variable will impact the

other variables; and the outcome(s) of interacting with the content is understandable and predictable within the domain of operation;
- **Useable**: The method is accessible, understandable, and credible to the designer;
- **Desirable**: The outcome(s) of interacting with the method is appropriate and valuable.

Based on these properties, Daalhuizen and Cash (2021) predict that method performance will be negatively impacted by incomplete reporting of the content, conflict between the content elements, and conflict between the content and designer/use context. Some of these properties have been operationalized, for example by Tromp and Hekkert (2016) who define process quality, process efficiency, and design quality as key measures to evaluate effect-driven design methods. The four properties of 'good' method content broadly complement prior work on method development by elaborating Daalhuizen et al.'s (2019) discussion of types of methodological elements and their internal hierarchy, as well as Gericke et al.'s (2017) distinction between method evaluation and validation.

More generally, these properties provide an initial basis for evaluating the content of design methods, which broadly mirrors assessment frameworks in related fields such as software engineering (Kitchenham et al., 2002; Kitchenham, Dyba, & Jorgensen, 2004) and prevention science (Flay et al., 2005; Gottfredson et al., 2015). However, this only treats the reporting of the method itself and thus neglects the other major aspect affecting the credibility of methods i.e. the robustness of the supporting evidence for methodological claims (Cash, Daalhuizen, & Hay, 2022; Gottfredson et al., 2015; Kitchenham et al., 2004; Prochner & Godin, 2022). For example, Tromp and Hekkert (2016) argue for the need to combine qualitative and quantitative evidence to credibly evaluate a method, while Vermaas (2016) critiques expert justifications and highlights robust empirical evidence as essential to credible method validation. Similarly, Olewnik and Lewis (2005) discuss how valid decision support methods should be logical, use reliable information, and not bias the designer, and use this as a basis for evaluating the evidence used to support method claims. Thus, while Daalhuizen and Cash's (2021) framework complements prior design research on method development, method content and evaluation (Olewnik & Lewis, 2005; Tromp & Hekkert, 2016; Vermaas, 2016), this must be contextualised with respect to its supporting research claims.

1.3 Evaluating research claims

Fundamental to evaluating research claims, as well as communicating their significance to practitioners, is understanding how they are supported by evidence derived from the research and development process (Flay et al., 2005).

However, while the importance of assessing evidence has been highlighted by several authors (Cash, 2018; Cross, 2012; Reich, 2010) there are few current standards for evidence evaluation in design research (Cash, 2018). This further contrasts with fields developing interventions akin to design methods, including software engineering (Kitchenham et al., 2004), management (Bansal & Corley, 2011), education (Levin & O'Donnell, 1999), and all areas of health (Glasgow & Emmons, 2007; Grimes & Schulz, 2002). Despite this deficit in design research, there is a remarkable degree of commonality across fields, especially in relation to evaluating interventions influenced by interaction with a method user (as with design methods; Levin & O'Donnell, 1999). Hence, it is possible to build on these standards in developing an understanding of claims supporting design methods.

In this context, two works are particularly relevant. First, Gottfredson et al. (2015) describe standards for evaluating the research and development process, from theory to 'effective' recommendations for practice. Second, Grimes and Schulz (2002) describe standards for evaluating evidence for causal research claims (e.g. relating an intervention to increased performance). Together these highlight the need to not only evaluate the intervention itself (as in Section 1.2), but also: i) the rationale for the specific intervention based on prior research and real-world practice; ii) the development of the intervention, including the key decisions taken and how these are grounded in research; and iii) the nature of the claims being made, their extent and scope of relevance (e.g. contextual bounding), and the evidence that is being used to support these. Coupled with the properties of 'good' method content outlined in Section 1.2, these provide a basis for developing an assessment framework for design methods.

2 *Building a systematic assessment framework for design methods*

Bringing together the literature from Section 1 it is possible to build a basic lens for understanding the assessment of design methods from conception to application. This forms a logical chain of evidence from initial insights regarding the need, through development, to claimed impact, anchored around the method content (Daalhuizen & Cash, 2021). We conceptualise this chain with respect to five major links as illustrated in Figure 1: method motivation (Section 2.1), method nature (Section 2.2), method development (Section 2.3), method content (Section 2.4), and method claims (Section 2.5). Weaknesses in any of these links diminishes the strength of the whole chain, and ultimately the design method itself. This forms the conceptual basis for our assessment framework, which we operationalise in Table 1 (and detail in Sections 2.1–2.5).

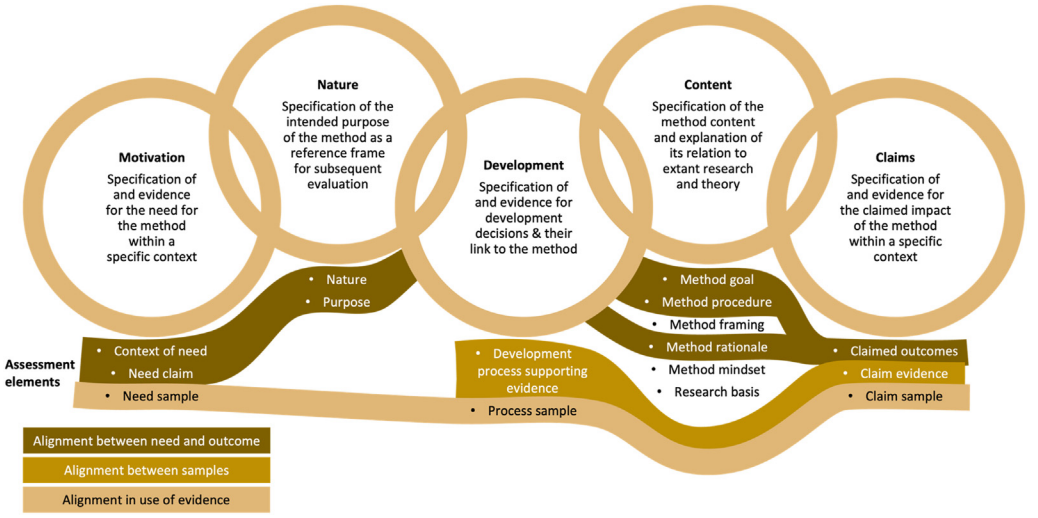

Figure 1 The chain of evidence supporting the proposal of a design method with assessment elements listed below each link, and alignments between elements indicated by the coloured threads

Important to note here, is that no assessment framework can exhaustively capture every aspect of method research and the associated ecosystem of method producers, promoters, users, and wider 'market' (for lack of a better term) of methods available in the wild. Therefore, we explicitly limit our focus to evaluating the robustness of the research supporting the proposal of a method and hence an understanding of its likely reliability in fulfilling its claims. This focus on the endogenous elements of method assessment implies two important exclusions relevant to the broader discussion of design methods.

First, the 'good' artefact property of desirability goes far beyond the scope of endogenous assessment criteria. Just as with products, desirability is also linked to users' perceptions of value or novelty, marketing, branding, and many other exogenous factors beyond the content of a method and its supporting evidence (many of which are little acknowledged in academic reporting, including in our sample of papers). Second, and following the same logic, we also consider method adoption, popularity, or other similar assessment criteria beyond the scope of endogenous assessment. Again, such outcomes are highly dependent on exogenous factors including marketing as well as the producers and promoters of a method and can often completely overshadow endogenous factors in user decision making (just as with products and their marketing). Hence, we exclude exogenous factors related to the method ecosystem or adoption unless there are specific claims being made about this in the reporting of the method itself.

With these exogenous exclusions in mind, Figure 1 illustrates the chain of evidence together with the endogenous assessment elements associated with each

link. Further, it highlights alignments between these elements that can strengthen or weaken the overall chain. For example, alignment between the need sample and the claim sample ensures that claims are being recontextualised in the same general group from which the motivation for the method emerged. For details of each assessment element see Table 1.

Table 1 A proposed systematic assessment framework for design methods, from need to impact (see Figure 1), building on method content theory (Daalhuizen & Cash, 2021)

Element	*Operationalisation*	*Description based on literature*
Method motivation		
Context of need	Practice Education Research Not Reported	*Multiple choice*: Where does the need for the method originate? This includes the major sources of insight in design research following Cash (2018, 2020)
Need claims	Efficacy Effectiveness Dissemination Not reported	*Multiple choice*: What is the extent of the need claim? This includes efficacy (validity), effectiveness (scale of impact), and dissemination (scope of applicability) following Daalhuizen and Cash (2021) and Gottfredson et al. (2015)
Need sample	Students Practitioners Both Not reported	*Selection*: From what sample is evidence for the need derived? This includes the major sub-samples typically recognised in design research following Atman et al. (2007)
Method nature		
Nature	Principle Approach Strict method Tool Template	*Multiple choice*: What is the nature of the method? This ranges from generic principles that do not necessarily say what to do to templates describing the structure of a specific output (Lewrick et al., 2020; van Boeijen et al., 2020)
Purpose	Support practice Support education Support both Not reported	*Selection*: What is the general purpose of the method? This includes the major impact areas typically recognised in design research (Meyer & Norman, 2020; Zielhuis et al., 2022)
Method development		
Development process supporting evidence	Expert practitioner opinion Research through design Single case Multi-case Experiment RCT Not reported	*Selection*: What approach has been used to generate evidence to inform method development? This ranges from expert opinion to Randomised Controlled Trails (RCTs)) adapted from the generic evidence hierarchy by Grimes and Schulz (2002). While RCTs are not expected for development they are included for completeness in terms of levels of evidence.
Process sample	Students Practitioners Both	*Selection*: What sample has been used to inform method development e.g. via prototyping with students? Again, following Atman et al. (2007)
Method content		
Method goal	Specific goal(s) Prioritisation or hierarchy Not reported	*Selection*: What is the goal(s) of the method? This includes the specific goal the method is to contribute to, its scope, and degree of flexibility building on Lee, Bobko, Earley, and Locke (1991)
Method procedure	Specific steps How to complete steps Purpose of steps	*Selection*: What are the steps in the method? This includes the structural knowledge about a specific way to reach a goal building on Roozenburg and Eekels (1995)

(*continued on next page*)

Table 1 (*continued*)

Element	*Operationalisation*	*Description based on literature*
Method framing: i) Context	Description … with explicit boundaries Not reported	*Selection*: In what context can the method be applied? This includes several dimensions: i) organisational and environmental context of use, ii) task or type of action involved, and iii) positioning or relation to the wider design process, building on various descriptions of method staging (Andreasen, Thorp Hansen, & Cash, 2015; Badke-Schaub, Daalhuizen, & Roozenburg, 2011; Gericke et al., 2017)
Method framing: ii) Task	Description … with explicit boundaries Not reported	
Method framing: iii) Positioning	Description … with explicit boundaries Not reported	
Method framing: Prerequisites	Required competences Required materials Required resources Required knowledge	*Multiple choice*: What is needed to stage the method? This includes the prerequisites or resources necessary for successful use (Andreasen et al., 2015; Badke-Schaub et al., 2011; Gericke et al., 2017)
Method rationale	Goal(s) success criteria Goal(s) end conditions Rationale for above criteria Support to reflect on progress/goal completion Not reported	*Multiple choice*: What is the performance-goal relationship for the method? This includes how to evaluate if you have succeeded, when to end, and how to reflect on progress with respect to the method's goal(s) in its specific domain and context of use. Again, building on Lee et al. (1991)
Method mindset	Values and beliefs Working principles Not reported	*Multiple choice*: What is the required method mindset? This includes descriptions of underlying values and beliefs and basic working principles building on Andreasen (2003)
Research basis	Logical speculation or inductive reasoning References to past findings Existing conceptual arguments Existing models, diagrams, or figures Existing theory	*Selection*: What is the basis for the components incorporated in the method? This includes the logic behind translation from grounding research to method content, and the associated implicit or explicit expectations for its performance, building on Cash (2020) and Colquitt and Zapata-Phelan (2007)
Method claims		
Claimed outcomes	Efficacy Effectiveness Dissemination	*Multiple choice*: What is the extent of the outcome claim? Again, following Daalhuizen and Cash (2021) and Gottfredson et al. (2015)
Claim evidence *	Expert practitioner opinion Research through design Single case Multi-case Experiment RCT Not reported	*Selection*: What approach has been used to generate evidence to inform method claims? Again, following Grimes and Schulz (2002) * Iterated for each claim
Claim sample *	Students Practitioners Both	*Selection*: From what sample is evidence for the claim derived? Again, following Atman et al. (2007) * Iterated for each claim

2.1 Method motivation: why is the method needed?

Evaluation of the validity and impact of a method logically builds on a foundational understanding of the need(s) that it addresses. Without this, success claims become detached from context. Thus, specification and evidence for need(s) forms the first link in our chain (Figure 1).

Three elements are required to understand need. First, where does the need for the method originate: practice, education, and/or research (following the major sources of insight typically found in design research (Cash, 2018, 2020))? This provides an initial context essential to understanding the later validity of causal claims (Wacker, 2008).

Second, what is the claimed extent of the need: efficacy (i.e., the validity of the need), effectiveness (i.e., the scale of the need), and/or dissemination (i.e., the generalisability of the need) (adapted from Gottfredson et al. (2015) and Daalhuizen and Cash (2021))? This provides the basis for understanding the extent of causal claims (Wacker, 2008).

Third, from what sample is evidence for the need derived: students and/or practitioners (following the major sub-samples typically found in design research (Atman et al., 2007; Cash, Isaksson, Maier, & Summers, 2022))? This again informs understanding of the context of the need and provides a first point of alignment across the chain. For example, if a need is based in practice but evidence is drawn from student samples during method development or testing then potential weaknesses in the chain could emerge (Figure 1). Together, these three elements serve to evaluate the specificity and evidence for the method need. While these do imply a gap in the wider method market (i.e., the need is at least relevant to the sample reported in the motivation) they primarily deal with the claims made in the research itself and thus generalisation to the whole method ecosystem and systematic positioning of originality and market research are not the focus here.

2.2 Method nature: what type of method is it?

Evaluation of method performance requires a reference frame for its intended purpose. For example, some methods are intended to provide general, abstract guidelines that are contextually adapted and applied by the designer (e.g., some versions of design thinking (Brown, 2008)), while others intended to provide a more constrained framework that will lead to repeatable outcomes across designers (e.g., many versions of the design structure matrix (Pektaş & Pultar, 2006)). Such differences can substantially impact how method claims should be evaluated in terms of their generalisability (Wacker, 2008). Thus, specification of method type forms the second link in our chain (Figure 1).

Two elements are required to understand type. First, what is the nature of the method? Despite the widespread use of the term 'method', this has substantially different meanings across the design research literature. Cutting across this variation, the key criteria affecting evaluation of causal claims is generalisability i.e., their sensitivity to method user and context (Wacker, 2008). Therefore, we build upon Daalhuizen et al.'s (2019) logic for method categorisation based on the scale of the cognitive support offered. Here, we contend that the greater the support the less the (particularities of the) designer and context will impact method outcomes. Hence, we identify five distinct types of methods commonly found in the literature (Lewrick, Link, & Leifer, 2020; van Boeijen et al., 2020): i) values and principles that guide overall work (e.g., principles of sustainable design or user-centered design), ii) approaches that structure a whole process (e.g. the Vision in Design method; van Dijk & Hekkert (2011), or Product Development Process; Ulrich and Eppinger (2008), iii) strict methods that structure sequences of tasks (e.g. mind mapping or Quality Function Deployment (QFD)), iv) tools that support tasks (e.g. PrEmo; Desmet (2018), or CAD software), and v) templates that structure the output of actions (e.g. business model canvas; Osterwalder, Pigneur, Oliveira, and Ferreira (2011) or the eco-design strategy wheel; van Boeijen et al. (2020)). This provides the basis for understanding the generality, abstraction, and contextual bounding of a method with respect to design work (Wacker, 2008).

Second, what is the general purpose of the method (as opposed to the specific goal, see Section 2.4): supporting practice and/or education (following the major impact areas in design research (Meyer & Norman, 2020; Zielhuis, SleeswijkVisser, Andriessen, & Stappers, 2022))? This provides a frame of reference for testing of causal claims (Wacker, 2008), and a second point of alignment across the chain. For example, a need may emerge from practice, but a method may be directed towards education, and then tested in practice and/or education. As such, clarification of alignment between these contexts and samples is needed to avoid weakness in the chain. Together, these two elements serve to evaluate the specific method type, providing a reference frame for the rest of the chain (Figure 1).

2.3 Method development: how is the method designed?

Evaluation of developmental robustness builds on an understanding of the evidence used to support key choices during the development of a method (Gottfredson et al., 2015; Vermaas, 2016). Thus, specification and evidence for development decisions forms the third link in our chain (Figure 1).

Two elements are required to understand development. First, what approach has been used to generate evidence to inform method development decisions? As this typically deals with the use of evidence to support causal claims, we

follow a widely accepted hierarchy of evidence in this context (Grimes & Schulz, 2002). This provides the basis for understanding the support for key decisions that might impact method content or claims (Gottfredson et al., 2015).

Second, what sample has been used to inform method development: students and/or practitioners (again following Atman et al. (2007))? This again informs understanding of the context of the development and provides a third point of alignment across the chain. While there are many other details required for replicability, these two elements serve to evaluate specificity and evidence for method development.

2.4 Method content: how does the method work?

Evaluation of method efficacy is underpinned by understanding of method content (Daalhuizen & Cash, 2021; Seepersad et al., 2006), as well as how this has been translated from basic research (Gottfredson et al., 2015). This is a key feature of design methods as an output of design *research* — as opposed to design *practice*. Thus, specification of method content and its basis in design research insights forms the fourth link in our chain (Figure 1).

Two elements are required to understand content. First, what does the method comprise? Here, we build directly on the model of method content proposed by Daalhuizen and Cash (2021), which includes the five main components listed below. This provides the basis for understanding the key concepts and relationships that define how a method functions, as a foundation for explanation, prediction, and causal claims (Wacker, 2008). Notably, these reflect the reported content of the method itself and thus provide several points of alignment with the wider supporting research across the chain, as highlighted and exemplified with the Contextmapping method (listed in van Boeijen et al., 2020) below.

- **Method Goal**: the explicit description of the goals and their prioritization a method aims to help achieve through method use (related to Sections 2.1 and 2.2). For example, the contextmapping method describes its goal as helping designers create solutions that fit people's needs.
- **Method Procedure**: the explicit description of the structural activities involved in the proper use of the method and their relative chronological and logical ordering (related to Section 2.2). For example, the contextmapping method describes its procedure in three steps for collecting and communicating user insights: preparation and sensitizing, generative assignments, and analysis and ideation.
- **Method Framing**: the explicit description of the scope of use setting (including context, task, and positioning) and its implications and prerequisites for method use. For example, the contextmapping method described

that it is to be preferably used in the pre-concept stage and as part of co-design or co-creation processes.

- **Method Rationale**: the explicit description of the performance-goal relationship and motivations underlying the goals of the method (related to Section 2.1). For example, the contextmapping method describes that it helps uncover users' latent knowledge that helps to empathize with intended users.
- **Method Mindset**: the explicit description of the set of values, principles, underlying beliefs, and logic that inform method use. For example, the contextmapping method describes that a core value is that the people designers design for are the expert of their own experiences and that designers should respect this.

Second, what is the basis for the components incorporated in the method? This follows the idea that methods embody research-based understanding of design work (Daalhuizen, 2014; Gray, 2022; Stappers & Sanders, 2005). Hence, the translation from research insight to a functional method constitutes some form of test (even if only implicitly). Therefore, it is possible to follow the acknowledged schema for assessing such translation provided by Colquitt and Zapata-Phelan (2007), and recently applied in design research by Cash (2020). This provides the basis for understanding the robustness of the research logic underpinning method content. Together, these two elements serve to evaluate the specificity and research grounding of the method content.

2.5 Method claims: what is the Method's impact?

Evaluation of claimed impact builds on the robustness of evidence for both the basic functionality of a method, as well as its ability to address its intended need/purpose in context (Daalhuizen & Cash, 2021; Gottfredson et al., 2015). These form a logical specification of and counter point to the claimed needs (Section 2.1). Thus, specification and evidence for claims forms the final link in our chain (Figure 1).

Three elements are required to understand claims. First, what is the extent of the outcome claim(s): efficacy (i.e., the functional predictability and validity of the method), effectiveness (i.e., the usability and scale of its impact in context), and/or dissemination (i.e. the desirability and general uptake) (adapted from Gottfredson et al. (2015) and Daalhuizen and Cash (2021))? This provides the basis for understanding the specific nature of the claims being made, providing a reference frame for subsequent testing methods and evidence (Gottfredson et al., 2015).

Second, what approach has been used to generate evidence to inform each method claim? Due to the typical focus on causal claims (i.e., the method improves …) we again follow the hierarchy of evidence provided by Grimes and

Schulz (2002). This provides the basis for understanding the robustness of the support for each claim (Gottfredson et al., 2015).

Third, from what sample is evidence for the claim derived: students and/or practitioners (again following Atman et al. (2007))? This again informs understanding of the context of the claim and provides a final point of alignment across the chain. Together, these three elements serve to evaluate the specificity and evidence for the method claims. As with motivation, while these do imply a contribution to the wider method market (i.e., the method is at least relevant to the degree reported in the claim) they primarily deal with the claims made in the research itself and thus generalisation to the whole method ecosystem and evaluation of potential uptake or market buy-in are beyond the scope of these elements.

Bringing together the criteria discussed in this section, we can propose an assessment framework for design methods, which systematically evaluates each link in the chain of evidence (Figure 1). Table 1 provides an overview of this assessment framework, detailing each link, their major elements, operationalisation, and description based on literature.

3 Method

To provide a foundation for field development and ground current debate, we conducted a systematic review of recent research proposing a design method, following the updated PRISMA guidelines (Page et al., 2021). Here, our intention was not to characterise the whole history of methods research, but rather establish the current state of design methods research as a basis for moving forward as a field. Given this aim the review comprised three main phases: (1) identification of design research proposing a method via search of major design journals, (2) screening the records and eligibility assessment according to pre-defined inclusion and exclusion criteria, and (3) inclusion of eligible records in the final sample for further review and analysis. The PRISMA flow diagram summarising our process is presented in Figure 2.

While the PRISMA approach was developed primarily as a guide for meta-analyses and systematic reviews of health interventions, such as clinical trials, the 27-item checklist and flow diagram are applicable for other types of systematic reviews (Page et al., 2021). This provides a standard approach, which has been applied in several recent reviews of design research (e.g., see Hay et al. (2017)). In this paper, we focus on current research proposing new methods bounded by two main criteria. First, we only consider methods published in recognised design research journals for two main reasons: i) these represent the 'best' of the field with more stringent peer review and acceptance criteria than either conferences or books (especially with respect to the provision of

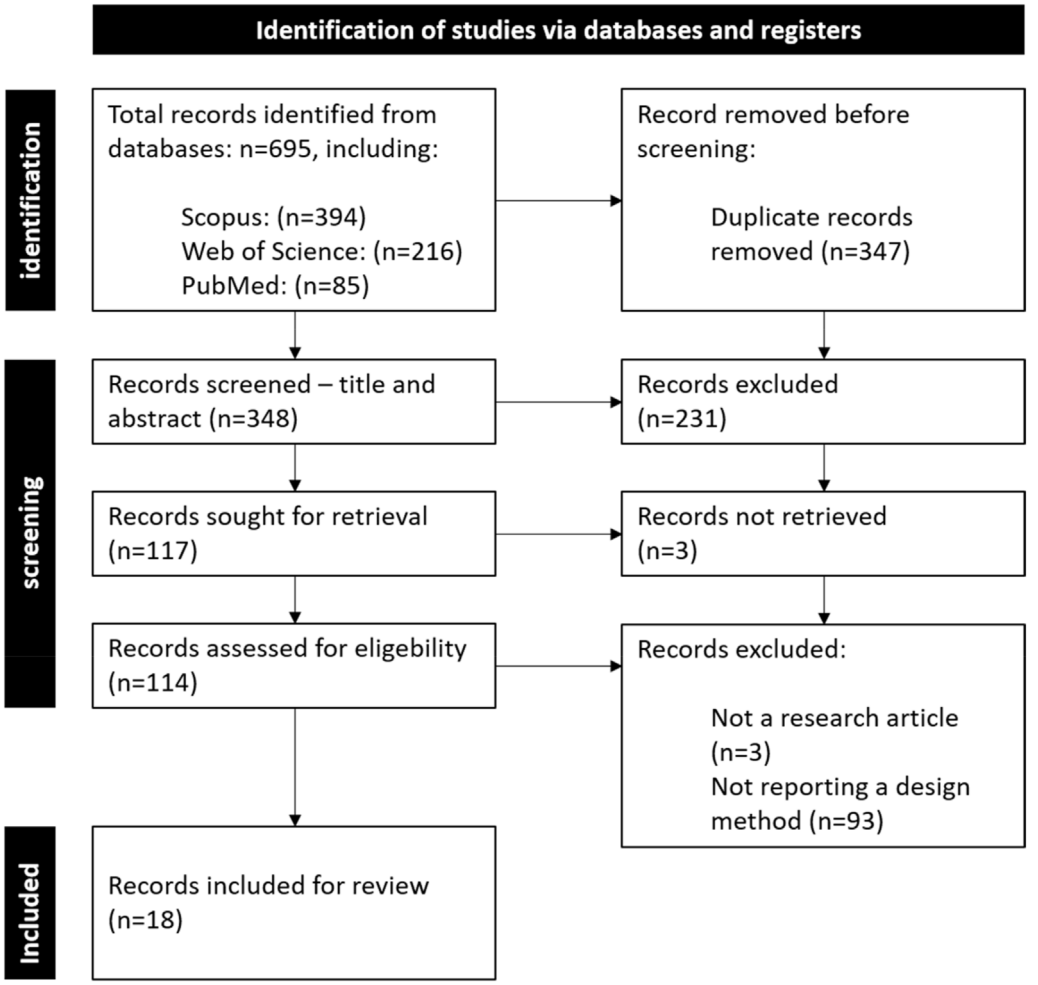

Figure 2 The PRISMA flow diagram specific to the literature review presented in this paper, based on Page et al. (2021)

supporting evidence) and ii) these reflect a bounding of the design research domain recognised within the field itself (Cash, 2018; Gemser, de Bont, Hekkert, & Friedman, 2012). Notably, we exclude textbooks and method repositories (where most method descriptions are found) because these primarily focus on dissemination and hence do not typically report underlying research and evidence (Figure 1). Second, we only considered methods published in 2020, due to these representing the most current still unaffected by the potential limitations of the COVID-19 pandemic (e.g., limiting data access).

3.1 Article selection process

The search was conducted based on the journal list developed by Gemser et al. (2012), which is widely recognised in the community as identifying top design

research journals. Again, as our focus is on establishing the current best case, we deliberately focus on a limited set of journals recognised as leaders in the field. Due to this focus, we were able to review every paper published in 2020 across all six journals summarised in Table 2. This resulted in a total of 695 records identified.

Following this, records were screened by a trained research assistant based on the title and abstract. Here, 231 records were excluded as immediately out of scope (e.g., reporting general design research and not including the proposal of a new method), not retrievable (n = 3) or not being a research article (n = 8). A second round of screening examined the remaining 106 records, which were screened based on the full text. Again, records were excluded if they did not propose a specific method that fell within the basic definition summarised in Table 2 (e.g., reporting a conceptual framework that could be applied by a designer but was not claimed as a specific method or 'methods' that were purely computational and thus eliminated interaction between method and designer). This resulted in a final selection of 18 records. The full list of inclusion and exclusion criteria applied can be found in Table 2.

3.2 Article analysis process

To analyse the 18 records selected for the review, we used a multi-part coding process. First, we coded the records for background information like application area, research context, etc. Second, we analysed selected records according to the elements of our proposed assessment framework (Table 1). The first round of analysis was performed by a trained research assistant. Training happened iteratively, starting with an explanation of the assessment framework after which the research assistant analysed three records, noting down any doubts that arose during analysis. After this, one of the authors and the research assistant discussed the analysed records and resolved all instances of doubt. After training, the research assistant analysed all records, again marking any instances of doubt. The second round of analysis was performed

Table 2 Review inclusion and exclusion criteria

Criteria	*Explanation*
Quality	Publication in top general design research outlets as recognised by the design research community (Gemser et al., 2012): *Design Studies, Design Issues, Journal of Engineering Design, International Journal of Design, The Design Journal, Journal of Design Research*
Publication date (2020)	Most recent publication date unambiguously unaffected by potential data access limitations introduced by the COVID-19 pandemic
Proposal of a method	Proposal of a contribution corresponding to the definition: '*a formalised representation of a design activity, which functions as a mental tool to support designers in achieving a goal, in relation to the circumstances and resources available*' (Daalhuizen et al., 2019).

by one of the authors during which all instances marked in the previous round were analysed and resolved.

4 Results

In this section we first provide an overview of the reviewed records before detailing the outcomes of our analysis.

4.1 Descriptive summary

Of the initial set of 695 records 18 were found to contain a method proposal. These were distributed across the surveyed journals: *Design Studies* (3), *Design Issues* (0), *Journal of Engineering Design* (8), *International Journal of Design* (3), *The Design Journal* (4), *Journal of Design Research* (0). Further, they targeted a wide range of design practices, including, ideation (3), development (10), modelling (2), and evaluation (3). Similarly, they covered an array of contexts from manufacturing and engineering design to participatory and co-design applied to everything from products to food or shared spaces. An overview of all the reviewed records is provided in Appendix Table A. Given the spread of the reviewed records and distribution across outlets we conclude that this sample provides a credible foundation for evaluating current research across the field.

4.2 Assessing current design method research

Analysing the selected records in detail revealed the proposal of 18 methods, with some methods containing multiple elements (e.g., a method and a tool) resulting in a total of 25 methodological elements. Upon inspection, the evidence presented within a record typically addressed all included elements, as such we take the 18 records as our primary level of analysis. Evaluating these provided important findings across the chain of evidence (Figure 1).

4.2.1 Method motivation

In this link of the chain (Figure 1) we evaluated the context of the need, the need claim, and the supporting sample. The overall results for each of these evaluation elements is summarised in Figure 3a–c. Taken together, the major insight in this link was that while 14 records claimed a valid or specific scale of need (Figure 3b) emerging from a specific context (practice, research, or both) (Figure 3a), only 3 reported the sample supporting these claims (Figure 3c). Further, there was a general lack of maturity in need claims with only 3 records claiming anything beyond validity, neglecting the scale and applicability of the need. Together these leave significant gaps in the initial link in most records and point to the potential for much clearer articulation of the scope of need and the evidence supporting this.

4.2.2 Method nature

In this link we evaluated the nature and purpose of the method type, as summarised in Figures 3d and 3e. The reviewed records presented a total of 25 methodological artefacts, which included 2 principles, 7 approaches, 13 strict methods, 1 tool, and 2 templates (Figure 3d). Notably 7 records proposed two methodological artefacts as part of a wider proposal (e.g. a strict method and a supporting tool). While all artefacts could be consistently characterised based on Table 1 the nomenclature varied significantly across records. Further, 2 records did not report the intended target of the proposal. Thus, while most records were clear in their articulation of a specific purpose these results reveal a critical need for alignment in terminology across the design literature.

4.2.3 Method development

In this link we evaluated the development process, supporting evidence, and its associated sample, as summarised in Figures 3f and 3g. While we would not expect the full scale to be utilised in the context of method development (e.g., RCTs are rarely—if ever in the design context—used for developmental purposes) our results reveal a critical deficit in the chain of evidence. Specifically, 12 records did not report on the evidence used as the basis for their development process (Figure 3f). For those records that did report on evidence used, one reported a research through design study, two a single case study, two a multi-case study, and one an experiment. This again reveals a critical lack of maturity in reporting practices with regards to the development process. Thus, despite the acknowledged importance of understanding the basis for developmental decision making in traditional design, this is currently not reported in the context of method development.

4.2.4 Method content

In this link we evaluated the various aspects of method content as summarised in Figure 3h–p. Here, while all records provided some explanation of their research basis there was significant inconsistency and incompleteness with respect to the reporting of the method content itself. Specifically, no single record described all aspects of method content, with several elements notably neglected. For example, 10 records did not provide goal rationale (Figure 3n) and 7 did not provide specific goals (Figure 3h). Further, 13 records did not describe the underlying method mindset (Figure 3o) and 7 did not even report the prerequisites needed to complete the proposed method (Figure 3m). Again, this reveals a lack of maturity in reporting and highlights the need for further examination of what is needed to understand and use method proposals to increase consistency in this critical link. It is surprising that even for perhaps the most basic element of methodological proposals – their core content – reporting is so often lacking.

4.2.5 Method claims and evidence

In this link we evaluated the claimed outcomes, their supporting evidence, and associated samples, as summarised in Figure 3q–s. These results revealed two important insights. First, while 13 records reported on the efficacy of their proposals and 5 reported effectiveness, none reported on dissemination (Figure 3q). Second, of the 13 records making efficacy claims, 12 provided some form of evidence (Figure 3r) but only 6 explicitly defined the sample from which this was derived (Figure 3s). However, it is notable that 2 records reported claims on both efficacy and effectiveness, including the evidence and samples used. Hence, while examples of more complete reporting can be found in the reviewed literature there is a critical lack of clarity in reporting the supporting evidence and sample for method claims.

4.3 Summarising current design method research

Figure 3 provides a summary of the results from across the chain of evidence; however, one critical insight should be highlighted based on the detailed results reported in this section. When the reviewed literature is taken as a whole, there are major deficits in reporting in almost every link in the chain of evidence (Figure 3). Yet, for all links at least one record was identified that clearly reported the assessment elements (Table 1), with individual records often providing complete and clear reporting of one or more specific links (however, no record covered all links). Hence, within the reviewed records there is overall recognition and reporting of all links in the chain of evidence, and evidence for the emergence of possible best practices in each specific link. Thus, while there is an evident lack of consistency and maturity across the reviewed records there is also evidence for the possibility of improvement and the overall potential value to be derived from structuring the reporting of method proposals.

5 Discussion

We set out to better understand how design methods and their associated claims can be assessed and subsequently ground current debate on method credibility in design research. In answer to this, our results revealed that even though some papers are quite complete in reporting specific links in the chain of evidence (Figure 1), overall reporting is quite incomplete and never complete (in a single paper) across all links in the chain. Taking two of the most complete works as good examples, Stylidis, Wickman, and Söderberg (2020) propose a method for ranking attributes of perceived product quality while Y. Lee, Breuer, and Schifferstein (2020) propose a set of food design tools. Both works report to a moderate/high degree of maturity on the method itself, its development, evidence supporting its efficacy and the need, yet do not report the sample for the need nor on the dissemination of the method. Yet these examples are exceptions with most cases offering relatively incomplete reporting, reflected in the high numbers of 'not reported' results (red bars)

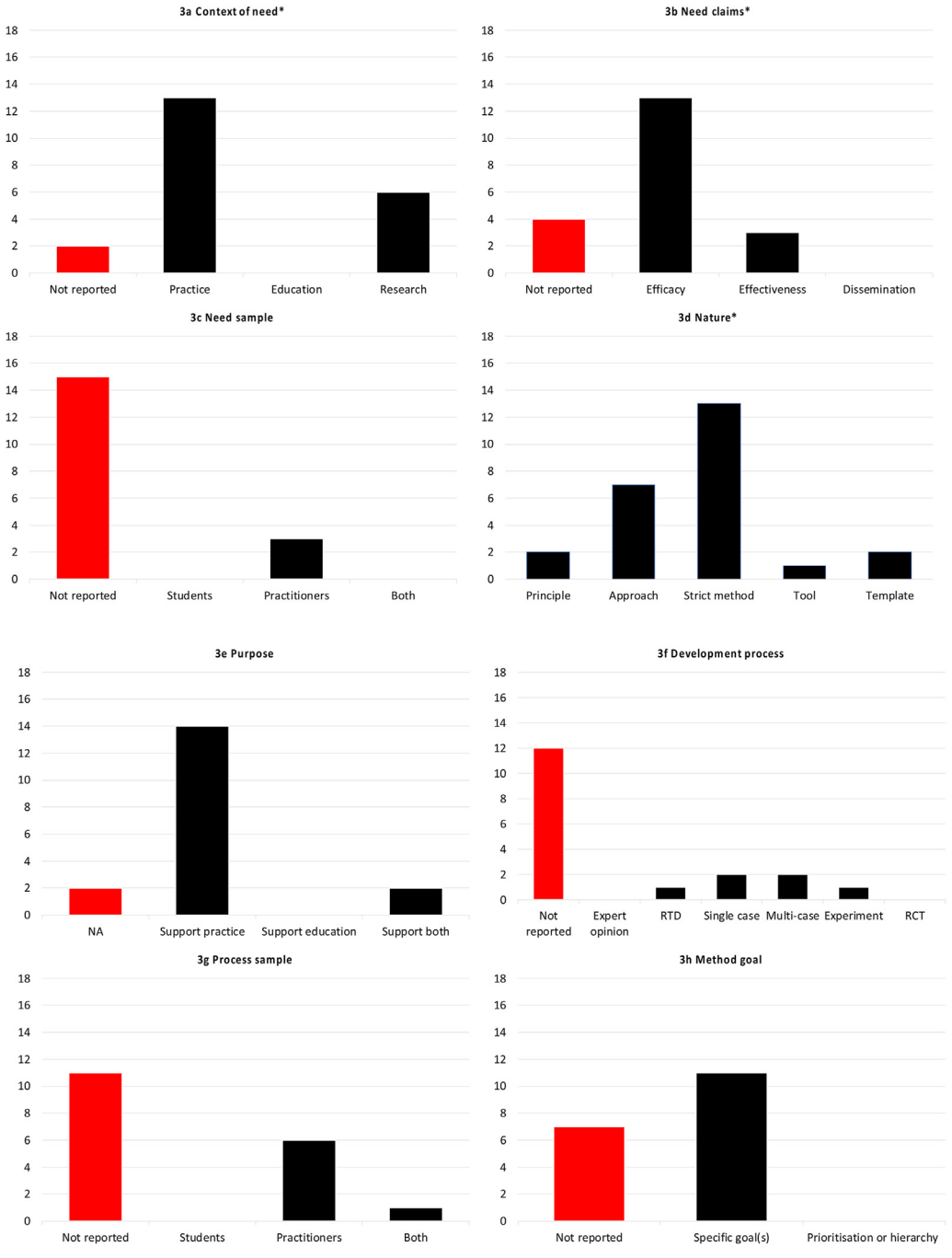

Figure 3 (a–s) Summary of results across the chain of evidence. Note: all y-axis reflect number of records and ∗ denotes multiple choice assessment

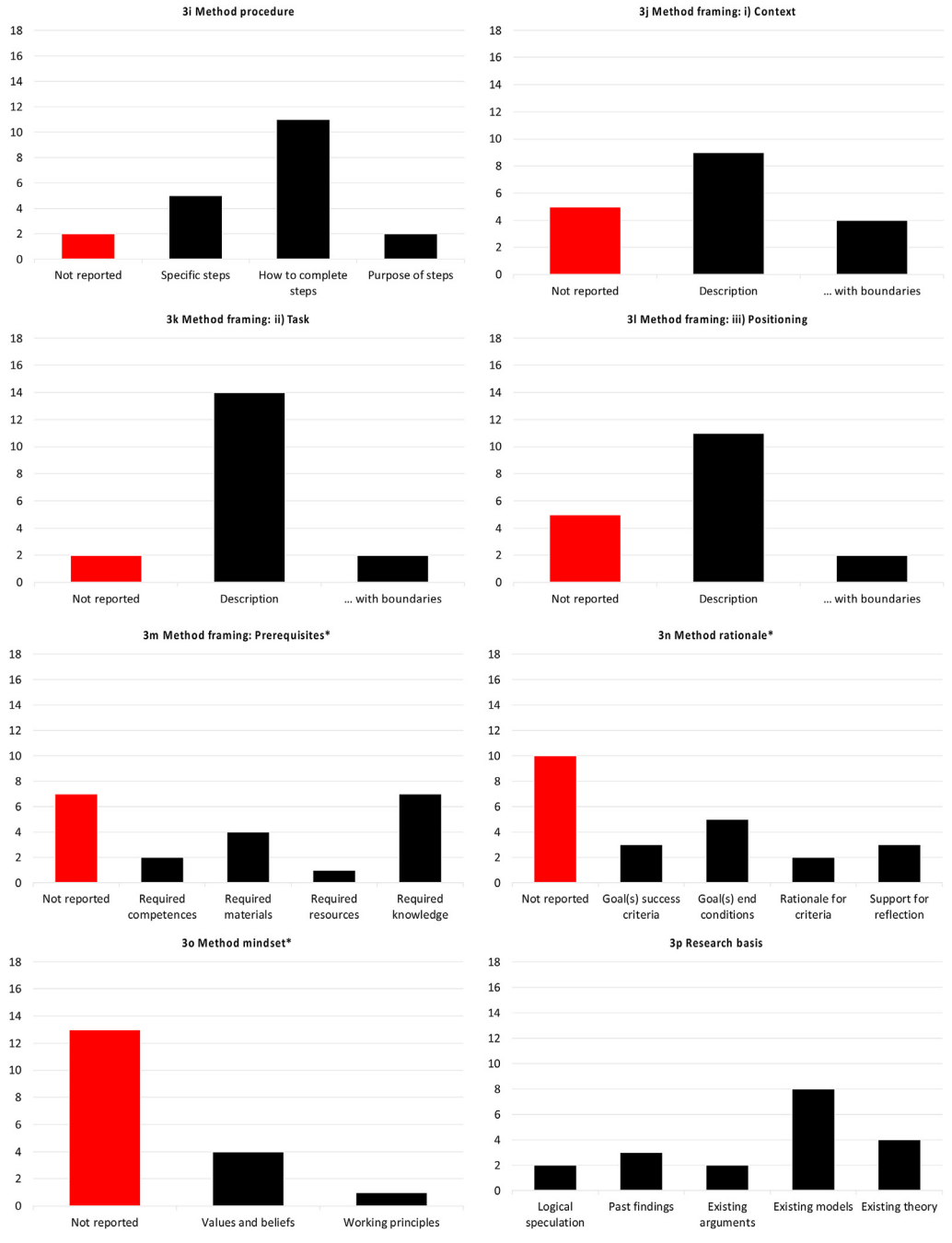

Figure 3 (continued)

Evaluating design methods

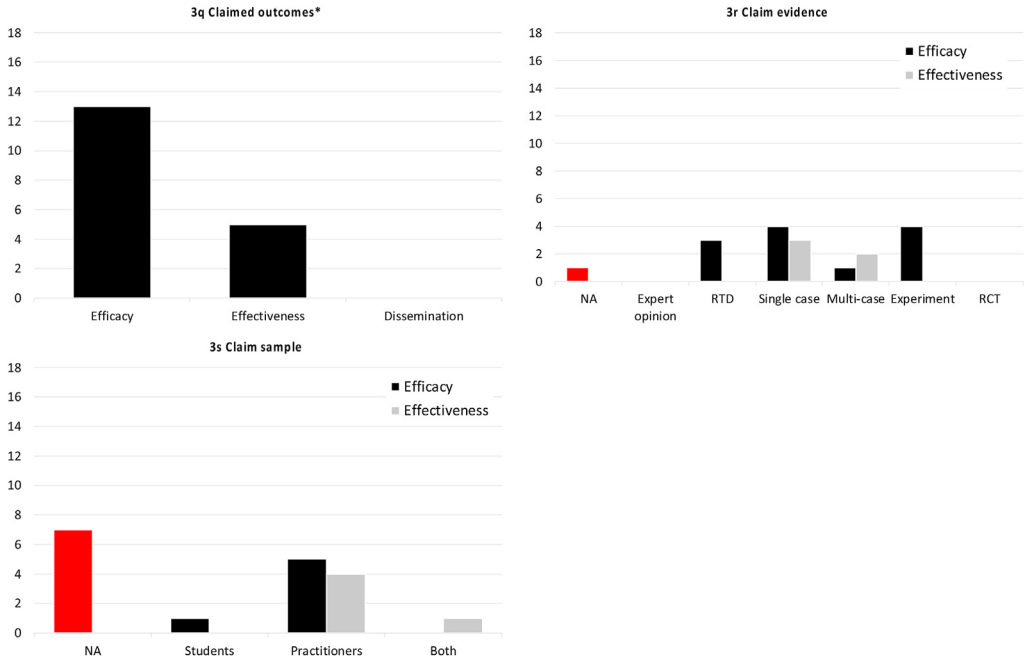

Figure 3 (continued)

in Figure 3a–s. Overall, most of the reviewed papers do not provide the information and/or evidence crucial to understanding the content and potential impact of the method proposals and their development. In short, the chain of evidence is often broken, with key links missing in most papers. Further, these missing links are also typically not recognised in the limitations of these papers.

Our findings reveal three major shortcomings highlighting the importance and potential value of the proposed assessment framework for supporting self-reflection, method assessment, and research evaluation: i) all links were found when looking across the whole set of reviewed papers, but never complete in individual papers; ii) for every link, potential 'best' practices were found, yet most papers did not report or only superficially addressed each link; and iii) where links were missing or otherwise incompletely addressed such deficits were typically not reported in the limitations. However, it is important to recognise that these results are perhaps not surprising given the lack of standards for method reporting in design research. Therefore, the positive exceptions highlighted (in completing most of the chain, in fully addressing specific links, and in reflecting on relevant limitations) should be applauded rather than the less complete examples criticized (see Table A for examples). Thus, our results demonstrate the need for clear and complete standards of

evidence in method research, as well as a more general need to take concerted action towards developing the maturity of the field.

5.1 Limitations

Before discussing implications, it is important to consider the main limitations of this work. First, the extent of the conceptual framework. The proposed framework reflects a first step in assessing the whole evidence chain from need to claimed impact; but does not deal with many other secondary factors that can impact methods, such as embodiment, facilitation, staging, context of deployment etc. and explicitly excludes interactions with exogenous factors that can impact users' perceptions of novelty, value, or wider uptake, such as the method producer/promotor, its branding/marketing, and other positioning and promotion within the method market and wider ecosystem. However, given the lack of maturity in this area this still provides an important contribution as evidenced by the findings reported in Figure 3, and a foundation on which market facing and other exogenous assessment elements could be built.

Second, the extent of the review. The review reflects current research in core design research journals. This serves its purpose as we aimed to assess the 'best' current practices, which are fostered by rigorous peer review and editorial oversight provided by the journal outlets. However, the reality is that most methods are only proposed in books or conferences, which vary wildly in peer review, editorial oversight, and ultimately quality and content. Thus, while our work provides an important foundation for moving the field forward, we need more work to really understand to what degree the 'best' represents the rest as well as how this might have changed over time, and how standards can be effectively deployed in the face of such varied outlets.

5.2 Implications and future research

Our work has several implications for method research, development, education, and practice. First, in terms of method research and theory, our evaluation, particularly of method content, highlights the potential for cross-cutting theory to scaffold maturation of the research field. Specifically, our findings show how increasing our understanding of the phenomenon underlying design methods and their use can directly contribute to improvements in development and reporting. While some elements were more consistently reported (such as aspects of method content), many elements are still treated implicitly or not reported at all (Figure 3). This inconsistency in reporting—even in leading design research journals—also emphasises the need for more systematic and consistent peer review. In this context, our work could provide a general guide for reviewers confronted with manuscripts involving method development, and points to the need for further investigation of quality criteria in this context, linking to the call for action in the recent Design Research Notes

initiative (Cash, Isaksson, et al., 2022). Hence, key future research questions include how method content interacts with the designer (as user of a method) in context, how to understand the adaption necessary in translating and embodying method content in practice, how the staging of methods might impact outcomes, and how method research can be consistently reviewed.

Second, in terms of wider method research, our work highlights the need for focused study of exogeneous factors affecting perceptions of methods and their uptake. Specifically, there is a need to better understand how users perceive methods (including evaluation of their novelty, relevance, and value in context) and how this relates to method content, its reporting, dissemination, and marketing. Further, there is a need to examine what other exogeneous factors impact method adoption and how these relate to the elements in Figure 1. For example, little is known about the impact of the method producer, promotor, or branding and marketing efforts on user uptake. Further, designers don't typically have time to try out multiple methods for a given project/task, and hence need support in identifying and applying the most relevant methods to their context, yet research on this meta level is, to the authors' knowledge, lacking. As such, a logical follow-up to this work is an examination of the evidence and reasons for method adoption amongst method users across contexts, including the impact of third-party actors in the method ecosystem, such as design consultancies or university marketing teams. Hence, key future research questions include how to understand exogenous factors impacting method perception and adoption and how method research can be tailored to engage with these without compromising quality.

Third, in terms of method development, there is currently no framework that establishes what good methods are and how to develop and report them. As such, our proposal for a chain of evidence (Figure 1) could form a foundation for developing good practice in this context. A second aspect of this is to acknowledge the contextual nature of method use, and to report what skills are required to properly use the method and in what contexts the method is best applied – and in what contexts we better refrain. Further, the systematic assessment framework can be used 'in reverse' as a checklist when planning, developing, and reporting methods. However, it is important to acknowledge that this is not a meta-method for method development, and that such an approach is an area for further research. Hence, key future research questions include how to understand and report on method development processes balancing generic and context specific elements and how best practices might be further developed in this area.

Finally, in terms of societal, practice, and educational impact, methods often target key challenges and include corresponding large-scale impact claims. Our findings highlight the need for a more in-depth discussion of what constitutes quality of methods and evidence in this context, and what types of research

infrastructure are needed to be able to achieve such quality standards reliably. Both policy makers and industry, for example, are more and more asking for evidence-based methods (Alonso et al., 2020; Design Council, 2020), which are proven to be actionable and effective. This is critical if design methods are to be held up alongside methods developed in related fields such as engineering or health. We need to show non-design sectors that deal with major societal challenges that our methods can—and do—make a difference and that they deliver the promises we make. It is therefore crucial that future research continues to examine how methods are adapted and applied, how implementation can be understood in and across contexts, and how we can offer compelling evidence of both methodological rigour and impact to diverse stakeholders.

6 *Conclusions*

We set out to better understand how design methods and their associated claims can be assessed and subsequently ground current debate on method credibility in design research. In doing so, we first developed a systematic assessment framework for design methods (Table 1) built on a logical chain of evidence from initial insights regarding the need to claimed impact (Figure 1). Based on this, we reviewed all papers published in 2020 in leading design research journals. Specifically, we examined whether papers that report new methods provide the information necessary to define and evaluate the proposed method and its development process, as well as support the claims associated with this.

Our results revealed that while all links in the chain of evidence are reported across the literature and best practices can be identified for each link, no individual paper either reports all links or consistently achieves best practice. While these findings might not be surprising—due to the lack of current standards of evidence in this area—they highlight the potential value of our proposed assessment framework and point to critical implications for maturing this central pillar of design research (impact).

Ultimately, we started with the question, 'what's in a claim', and can conclude with the realisation, that while the answer is complex and multifaceted (Table 1) it is also tractable in design research (with many positive examples of good—yet patchy—practice across the field). Our work thus provides a foundation for evaluating method research, demonstrates the need for clear and complete standards of evidence in this area, and highlights directions for future method research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix.

Table A Overview of papers included in the systematic review

Journal	*Citation*	*Title*	*Example of complete reporting in links:*
Design Studies	Hoolohan and Browne (2020)	Design thinking for practice-based intervention: Co-producing the change points toolkit to unlock (un)sustainable practices	• Motivation • Development • Content • Claims: effectiveness
	Strömberg, Pettersson, and Ju (2020)	Enacting metaphors to explore relations and interactions with automated driving systems	• Development • Claims: efficacy
	Pirinen and Tervo (2020)	What can we share? A design game for developing the shared spaces in housing	• Development
International Journal of Design	Karana, Barati, and Giaccardi (2020)	Living Artefacts: Conceptualizing Livingness as a Material Quality in Everyday Artefacts	
	Woo and Lim (2020)	Routinoscope: Collaborative Routine Reflection for Routine-Driven Do-It-Yourself Smart Homes	• Content • Claims: efficacy • Claims: effectiveness
	(Y. Lee, Breuer, & Schifferstein, 2020)	Supporting Food Design Processes: Development of Food Design Cards	• Development • Claims: efficacy • Claims: effectiveness
Journal of Engineering Design	Paparistodimou, Duffy, Whitfield, Knight, and Robb (2020)	A network science-based assessment methodology for robust modular system architectures during early conceptual design	• Claims: effectiveness
	Rigger, Vosgien, Shea, and Stankovic (2020)	A top-down method for the derivation of metrics for the assessment of design automation potential	• Claims: effectiveness
	Bashir and Ojiako (2020)	An integrated ISM-MICMAC approach for modelling and analysing dependencies among engineering parameters in the early design phase	• Motivation
	Eddy, Krishnamurty, Grosse, and Steudel (2020)	Early design stage selection of best manufacturing process	
	(D. Lee, Pan, & Fang, 2020)	Improving early stage system design under the uncertainty in reliability-wise structure	• Motivation
	Stylidis et al. (2020)	Perceived quality of products: a framework and attributes ranking method	• Development • Content • Claims: efficacy
	Stief, Dantan, Etienne, Siadat, and Burgat (2020)	Product design improvement by a new similarity-index-based approach in the context of reconfigurable assembly processes	
	Barravecchia, Mastrogiacomo, and Franceschini (2020)	The player-interface method: a structured approach to support product-service systems concept generation	

(*continued on next page*)

Table A (*continued*)

Journal	*Citation*	*Title*	*Example of complete reporting in links:*
The Design Journal	Güneş (2020)	Extracting Online Product Review Patterns and Causes: A New Aspect/Cause Based Heuristic for Designers	
	Rodgers, Mazzarella, and Conerney (2020)	Interrogating the Value of Design Research for Change	
	Celikoglu, Krippendorff, and Ogut (2020)	Inviting Ethnographic Conversations to Inspire Design: Towards a Design Research Method	• Development
	Jacobs et al. (2020)	Made-Up Rubbish: Design Fiction as a Tool for Participatory Internet of Things Research	• Claims: efficacy

Notes

1. 'Method content' is here deliberately separated from 'method use', to stress that methods can have all the proper ingredients yet fail to deliver when used improperly or under the wrong circumstances.

References

Alexander, C. (1971). The state of the art in design methods. *DMG Newsletter, 5*(3), 3–7.

Alonso, M. B., van der Bijl-Brouwer, M., Hekkert, P., Hummels, C., Kraal, J., Krul, K., et al. (2020). *Sleutelmethodologieën (KEM's) voor missiegedreven innovatie*.

Andreasen, M. M. (2003). Improving design methods usability by a mindset approach. In U. Lindemann (Ed.), *Human behaviour in design*. (209–218). Berlin Heidelberg: Springer.

Andreasen, M. M. (2011). 45 Years with design methodology. *Journal of Engineering Design, 22*(5), 293–332.

Andreasen, M. M., Thorp Hansen, C., & Cash, P. (2015). *Conceptual design: Interpretations, mindset and models*. Springer.

Araujo, C. S. (2001). *Acquisition of product development tools in industry: A theoretical contribution*. Denmark: Lyngby.

Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., & Saleem, J. (2007). Engineering design processes: A comparison of students and expert practitioners. *Journal of Engineering Education, 96*(4), 359–379.

Badke-Schaub, P., Daalhuizen, J., & Roozenburg, N. (2011). Towards a Designer-Centred Methodology: Descriptive considerations and prescriptive reflections. In B. Herbert (Ed.), *The future of design methodology* (pp. 181–197). Springer.

Bansal, P., & Corley, K. (2011). From the editors: The coming of age for qualitative research. *Academy of Management Journal, 54*(2), 233–237.

Barravecchia, F., Mastrogiacomo, L., & Franceschini, F. (2020). The player-interface method: A structured approach to support product-service systems concept generation. *Journal of Engineering Design, 31*(6), 331–348.

Bashir, H., & Ojiako, U. (2020). An integrated ISM-MICMAC approach for modelling and analysing dependencies among engineering parameters in the early design phase. *Journal of Engineering Design, 31*(8–9), 461–483.

Blessing, L. T. M., & Chakrabarti, A. (2009). *DRM: A design research methodology*. London: Springer.

van Boeijen, A., Daalhuizen, J., & Zijlstra, J. (2020). *Delft design guide: Perspectives - models - approaches - methods* (2nd ed.). BIS Publishers BV.
Brown, T. (2008). Design thinking. *Harvard Business Review, 2008*(June), 1–9.
Bucciarelli, L. L. (1994). *Designing engineers*. Cambridge, Massachusetts: MIT Press.
Cantamessa, M. (2003). An empirical perspective upon design research. *Journal of Engineering Design, 14*(1), 1–15.
Cash, P. (2018). Developing theory-driven design research. *Design Studies, 56*(May), 84–119.
Cash, P. (2020). Where next for design research? Understanding research impact and theory building. *Design Studies, 68*, 113–141.
Cash, P., Daalhuizen, J., & Hay, L. (2022). Editorial: Design research notes. *Design Studies, 78*, 101079.
Cash, P., Isaksson, O., Maier, A. M., & Summers, J. D. (2022). Sampling in design research: Eight key considerations. *Design Studies, 78*, 101077.
Cash, P., Valles Gamundi, X., Echstrom, I., & Daalhuizen, J. (2022). Method use in behavioural design: What, how, and why? *International Journal of Design, 16*(1), 1–21.
Celikoglu, O. M., Krippendorff, K., & Ogut, S. T. (2020). Inviting ethnographic conversations to inspire design: Towards a design research method. *The Design Journal, 23*(1), 133–152.
Chulvi, V., Mulet, E., Chakrabarti, A., López-Mesa, B., & González-Cruz, C. (2012). Comparison of the degree of creativity in the design outcomes using different design methods. *Journal of Engineering Design, 23*(4), 241–269.
Colquitt, J. A., & Zapata-Phelan, C. P. (2007). Trends in theory building and theory testing: A five-decade study of the academy of management journal. *Academy of Management Journal, 50*(6), 1281–1303.
Cross, N. (2012). Editorial. *Design Studies, 33*(1), 1–3.
Daalhuizen, J. (2014). Method Usage in design: How methods function as mental tools for designers *(PhD)*. TU Delft.
Daalhuizen, J., & Cash, P. (2021). Method content theory: Towards a new understanding of methods in design. *Design Studies, 75*, 101018.
Daalhuizen, J., Person, O., & Gattol, V. (2014). A personal matter? An investigation of students' design process experiences when using a heuristic or a systematic method. *Design Studies, 35*(2), 133–159.
Daalhuizen, J., Timmer, R., van der Welie, M., & Gardien, P. (2019). An architecture of design doing: A framework for capturing the ever-evolving practice of design to drive organizational learning. *International Journal of Design, 13*(1), 37–52.
Design Council. (2020). *Making life better by design: Design Council's story of impact*.
Dorst, K. (2008). Design research: A revolution-waiting-to-happen. *Design Studies, 29*(1), 4–11.
Eddy, D. C., Krishnamurty, S., Grosse, I. R., & Steudel, M. (2020). Early design stage selection of best manufacturing process. *Journal of Engineering Design, 31*(1), 1–36.
Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., et al. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*(3), 151–175.
Frey, D. D., & Dym, C. L. (2006). Validation of design methods: Lessons from medicine. *Research in Engineering Design, 17*(1), 45–57.

Gemser, G., de Bont, C., Hekkert, P., & Friedman, K. (2012). Quality perceptions of design journals: The design scholars' perspective. *Design Studies, 33*(1), 4–23.
Gericke, K., Eckert, C., & Stacey, M. (2017). What do we need to say about a design method? *Proceedings of the International Conference on Engineering Design, ICED'17*, 101–110.
Glasgow, R. E., & Emmons, K. M. (2007). How can we increase translation of research into practice? Types of evidence needed. *Annual Review of Public Health, 28*(1), 413–433.
Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., Sandler, I. N., et al. (2015). Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: Next generation. *Prevention Science, 16*(7), 893–926.
Gray, C. (2022). Languaging design methods. *Design Studies, 78*, 101076.
Grimes, D. A., & Schulz, K. F. (2002). An overview of clinical research: The lay of the land. *The Lancet, 359*(9300), 57–61.
Güneş, S. (2020). Extracting online product review patterns and causes: A new aspect/cause based heuristic for designers. *The Design Journal, 23*(3), 375–393.
Hay, L., Duffy, A., McTeague, C., Pidgeon, L., Vuletic, T., & Grealy, M. (2017). A systematic review of protocol studies on conceptual design cognition: Design as search and exploration. *Design Science, 3*, e10.
Hoolohan, C., & Browne, A. L. (2020). Design thinking for practice-based intervention: Co-Producing the change points toolkit to unlock (un)sustainable practices. *Design Studies, 67*, 102–132.
Jacobs, N., Markovic, M., Cottrill, C. D., Edwards, P., Corsar, D., & Salt, K. (2020). Made-up rubbish: Design fiction as a tool for participatory internet of things research. *The Design Journal, 23*(3), 419–440.
Jones, J. C. (1977). How my thoughts about design methods have changed during the years. *Design Methods and Theories: Journal of DMG and DRS, 11*(1).
Jones, J. C. (1992). *Design methods*. John Wiley and Sons.
Karana, E., Barati, B., & Giaccardi, E. (2020). Living artefacts: Conceptualizing livingness as a material quality in everyday artefacts. *International Journal of Design, 14*(3), 37–53.
Kitchenham, B. A., Dyba, T., & Jorgensen, M. (2004). Evidence-based software engineering. In *26th international conference on software engineering* (pp. 273–281).
Kitchenham, B. A., Pfleeger, S. L., Pickard, L. M., Jones, P. W., Hoaglin, D. C., El-Emam, K., et al. (2002). Preliminary guidelines for empirical research in software engineering. *IEEE Transactions on Software Engineering, 28*(8), 721–734.
Kumar, V. (2013). *101 design methods - a structured approach for driving innovation in your organization*. New Jersey: John Wiley & Sons.
Kunrath, K., Cash, P., & Kleinsmann, M. (2020). Designers' professional identity: Personal attributes and design skills. *Journal of Engineering Design, 31*(6), 297–330.
Lee, C., Bobko, P., Earley, C. P., & Locke, E. A. (1991). An empirical analysis of a goal setting questionnaire. *Journal of Organizational Behavior, 12*(6), 467–482.
Lee, Y., Breuer, C., & Schifferstein, H. N. J. (2020). Supporting food design processes: Development of food design cards. *International Journal of Design, 14*(2), 51–64.

Lee, D., Pan, R., & Fang, G. (2020). Improving early stage system design under the uncertainty in reliability-wise structure. *Journal of Engineering Design, 31*(10), 485–507.
Levin, J. R., & O'Donnell, A. M. (1999). What to do about educational research's credibility gaps? *Issues in Education, 5*(2), 177–229.
Lewrick, M., Link, P., & Leifer, L. (2018). *The design thinking playbook: Mindful digital transformation of teams, products, services, businesses and ecosystems*. John Wiley & Sons.
Lewrick, M., Link, P., & Leifer, L. (2020). *The design thinking toolbox: A guide to mastering the most popular and valuable innovation methods*. John Wiley & Sons.
Lloyd, P. (2019). You make it and you try it out: Seeds of design discipline futures. *Design Studies, 65*, 167–181.
Meyer, M. W., & Norman, D. (2020). Changing design education for the 21st century. *She Ji, 6*(1), 13–49.
Osterwalder, A., Pigneur, Y., Oliveira, M. A. Y., & Ferreira, J. J. P. (2011). Business model generation: A handbook for visionaries, game changers and challengers. *African Journal of Business Management, 5*(7), 22–30.
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). *The PRISMA 2020 statement: An updated guideline for reporting systematic reviews*. *The BMJ*. BMJ Publishing Group.
Paparistodimou, G., Duffy, A., Whitfield, R. I., Knight, P., & Robb, M. (2020). A network science-based assessment methodology for robust modular system architectures during early conceptual design. *Journal of Engineering Design, 31*(4), 179–218.
Pektaş, S. T., & Pultar, M. (2006). Modelling detailed information flows in building design with the parameter-based design structure matrix. *Design Studies, 27*(1), 99–122.
Pirinen, A., & Tervo, A. (2020). What can we share? A design game for developing the shared spaces in housing. *Design Studies, 69*, 100941.
Prochner, I., & Godin, D. (2022). Quality in research through design projects: Recommendations for evaluation and enhancement. *Design Studies, 78*, 101061.
Reich, Y. (2010). My method is better. *Research in Engineering Design, 21*(3), 137–142.
Rigger, E., Vosgien, T., Shea, K., & Stankovic, T. (2020). A top-down method for the derivation of metrics for the assessment of design automation potential. *Journal of Engineering Design, 31*(2), 69–99.
Rodgers, P. A., Mazzarella, F., & Conerney, L. (2020). Interrogating the value of design research for change. *The Design Journal, 23*(4), 491–514.
Roozenburg, N. F. M., & Eekels, J. (1995). *Product design: Fundamentals and methods*. Chichester: John Wiley Sons.
Seepersad, C. C., Pedersen, K., Emblemsvåg, J., Bailey, R., Allen, J. K., & Mistree, F. (2006). The validation square: How does one verify and validate a design method?. In *Decision making in engineering design* (pp. 305–316) ASME.
Simon, H. A. (2019). *The sciences of the artificial*. Cambridge: [M.I.T. Press.
Stappers, P. J., & Sanders, E. B. N. (2005). Tools for designers, Products for Users: The role of creative design techniques in a squeezed-in design process. In *The proceedings of the international conference on planning and design: Creative interaction and sustainable development* (pp. 1–16).
Stief, P., Dantan, J.-Y., Etienne, A., Siadat, A., & Burgat, G. (2020). Product design improvement by a new similarity-index-based approach in the context

of reconfigurable assembly processes. *Journal of Engineering Design, 31*(6), 349–377.
Strömberg, H., Pettersson, I., & Ju, W. (2020). Enacting metaphors to explore relations and interactions with automated driving systems. *Design Studies, 67*, 77–101.
Stylidis, K., Wickman, C., & Söderberg, R. (2020). Perceived quality of products: A framework and attributes ranking method. *Journal of Engineering Design, 31*(1), 37–67.
Vaajakallio, K., & Mattelmäki, T. (2014). Design games in codesign: As a tool, a mindset and a structure. *CoDesign, 10*(1), 63–77.
Vermaas, P. E. (2016). A logical critique of the expert position in design research: Beyond expert justification of design methods and towards empirical validation. *Design Science, 2*(May), e7.
Wacker, J. G. (2008). A conceptual understanding of requirements for theory-building research: Guidelines for scientific theory building. *Journal of Supply Chain Management, 44*(3), 5–15.
Wallace, K. (2011). Transferring design methods into practice. In H. Birkhofer (Ed.), *The future of design methodology*. (239–248). London: Springer.
Woo, J., & Lim, Y. (2020). Routinoscope: Collaborative routine reflection for routine-driven do-it-yourself smart homes. *International Journal of Design, 14*(3), 19–36.
Zielhuis, M., Sleeswijk Visser, F. S., Andriessen, D., & Stappers, P. J. (2022). Making design research relevant for design practice: what's in the way? *Design Studies, 78*, 101063.