How to Make AI Artists Feel Guilty in a Good Way? Designing Integrated Sustainability Reflection Tools (SRTs) for Visual Generative AI

Yuanyang Ren[*] and **Aswath Sivakumaran**[†] and **Emma Niemelä**[‡] and **Petra Jääskeläinen**[§]
MID, EECS, KTH Royal Institute of Technology, Stockholm, Sweden
yren@kth.se, aswaths@kth.se, emmanie@kth.se, ppja@kth.se

Abstract

AI can be energy intensive, and artists currently lack access to empowering information. With growing concerns of climate change and calls for environmental sustainability, there is a real need to explore strategies to communicate sustainability information to artists using generative AI, given its increasing presence and widening accessibility. This paper presents an exploratory Research-through-Design study (including design-informing survey, design prototyping, user testing) of integrating sustainability reflection features into generative AI systems, and provides preliminary knowledge of the design characteristics that can be leveraged, including artists' experiences of them. This paper finds that granular, relatable data visualizations and informed use of colors are effective in communicating about energy consumption. Furthermore, artists were positive towards "feeling bad" in the process of becoming aware of their impacts, and called for systems that could provide them low-energy settings during exploratory stages of the artistic process.

Introduction

Artificial Intelligence (AI) is computationally intensive (Dhar 2020) with increasing power demands (Mehonic and Kenyon 2022). However, the complex and often black-box nature of AI makes it difficult for people to understand the environmental impact of their work. The recent rapid and increasing use of accessible Generative AI (G-AI) tools warrant investigation into the sustainability aspects of such technology. This paper explores how to design **Sustainability Reflection Tools (SRTs)** for **Visual Generative AI (VG-AI)**. This paper serves as an exploratory study into transparentizing the environmental impact of AI art generators and discusses how designers (and users) of such tools can take steps to addressing the environmental sustainability problems glooming in the horizon.

State of Research: SRTs for Generative AI

Self-reflection for sustainability has been widely researched in HCI (Kefalidou et al. 2015), with a particular focus on studies and tools that attempt to promote more pro-environmental energy consumption habits. Strategies have included; information displays through smart monitors (Froehlich, Findlater, and Landay 2010), personalized information delivery (Mankoff et al. 2007), or computer games (Bang, Torstensson, and Katzeff 2006). Furthermore, well-established methodologies such as Life Cycle Assessment (LCA) and other critical frameworks (Grover, Emmitt, and Copping 2019) have been explored to analytically make sense of sustainability of technologies.

However, despite this diversity of research and tools development in HCI, the current state of SRTs in the context of AI - and particularly G-AI - is lacking both in terms of research and practically usable tools. Existing research focuses on building tools for the more technically inclined (Anthony, Kanding, and Selvan 2020), or is more generic (Lacoste et al. 2019) rather than geared towards specific applications. Simultaneously, current research has brought up environmental sustainability concerns in the specific context of G-AI (Jääskeläinen, Pargman, and Holzapfel 2022a; Jääskeläinen, Pargman, and Holzapfel 2022b; Bender et al. 2021), although attempts at addressing these concerns practically are scarce. Currently there are no SRTs aimed specifically for G-AI, or AI artists. However, as discussed previously, current SRTs are not suitable for non-technical end users (who in this case may be professional artists, or any individuals engaging in image-making using generative tools) and not necessarily well-versed in AI, Computer Science, or Environmental studies. The combination of the increasing energy demands of AI and G-AI tools becoming more prevalent and widely accessible (regardless of technical skill or available hardware) warrants the necessity of empowering as many stakeholders along the line as possible to take control (or at least be informed) of their environmental impact created while using these tools.

SRTs and The Complexity of Behavior Change for Sustainability

Multiple models exist to promote pro-environmental behavior. In this paper we have employed the most used "information" model (Froehlich, Findlater, and Landay 2010), and augmented it using colors, pictograms, and data visualizations. However, research surrounding what shapes pro-environmental or pro-sustainability behavior is not clear

[*] 1^{st} author: prototype design, created & performed user tests.
[†] 2^{nd} author: survey design, co-created user tests, co-writer.
[‡] 3^{rd} author: contributed to user tests.
[§] 4^{th} author: project supervisor, co-writer.

(Kollmuss and Agyeman 2002). Further, no long-term studies about the effects pertaining to SRTs or similar reflection tools exist, while short-term studies (mostly on carbon calculators) are heavily criticized, lack empirical evidence, and are mostly inconclusive (Biørn-Hansen, Barendregt, and Andersson 2020). This is primarily due to the simplistic viewing of change towards sustainability when in reality behavior change is complex (Brynjarsdottir et al. 2012; Strengers 2014) and influenced by various factors, such as the context, prior knowledge, feelings, culture, etc. However, many SRT related studies build on an underlying assumption that presenting information is enough to result in behavior change, painting a picture of humans as more consciously rational agents than they likely are. However, studies of eco-feedback tools have shown some promise (Holmes 2009) and warrant further investigation of what specific conditions and factors are successful (or not) in facilitating behavior change, in specific use practices. This exploratory empirical study *we do not argue or aim for long-term behavior change*, but rather focused on developing knowledge on these underlying factors that lay ground on behavior change for sustainability in context of VG-AI through asking: (1) How do certain design characteristics (colors, symbols, infographics) and strategies relate to the effectiveness of *communicating* sustainability-related information?, (2) What kind of quality v. impact trade-offs are users willing to make?, (3) What insights/themes can we draw from participants' experiences in the user tests to inform future research and development in VG-AI SRTs?

Methods

To address our research questions, we used an exploratory Research-through-Design (RtD) approach (Zimmerman and Forlizzi 2014) that involved a design-informing survey to map user insights relating different aspects of sustainability information representation through **Survey**, designed a **Prototype**, and performed a **User Study** in an exploratory setting.

Survey Questionnaire To Obtain General Insights of VG-AI Users

An online survey was distributed through social media and personal networks, targeted broadly at people who have used VG-AI systems. The survey consisted of 21 questions assessing attitudes and experience of color, pictographic associations, and data visualization. Additionally, demographic data (age, gender, current residence) was collected to ensure specificity and clarity of data due to the non-random convenience sampling (Gideon 2012). Background information was collected regarding respondents attitudes around environmental sustainability and familiarity around AI art generators [1], competence in art/design[2] (as our target audience was non-specific), and color blindness declaration to get an overview of factors that may influence how people experienced the evaluated aspects. Prior artistic experience was of interest, since respondents with experience in art might be more critical of color and pictograms. Eventually, data was gathered from 26 respondents, mostly located in Europe (n_1=20), followed by India (n_2=3) and New Zealand (n_3=2). Majority had some art/design competency (Advanced knowledge (student or professional)=8, Hobbyist=4, Some basic knowledge=9), 5 reported having no competency. Majority were aged 25-34 (n_4=18), the rest were between 18-24 (n_5=5) and 35-44 (n_6=2). The questions asked in the survey gathered insights specifically on these aspects (see also Fig. 1): (1) **Color associations**[3] in relation to energy consumption levels; (2) **Pictogram associations** in relation to environmental sustainability; (3) **Preferences and experience of data visualization** style for visualizing energy consumption[4].

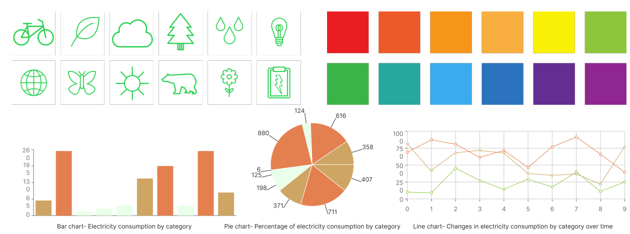

Figure 1: Colors (top-left), Pictograms (top-right), Data visualization styles (bottom).

Design Process and Prototype

A Figma prototype[5] of an AI art generator with SRT augmented features was developed (see Fig. 2). The prototype was based on Nightcafe[6] with UI modifications made with the intent to inform the user about the energy consumption of their usage. The SRT features were informed by the survey data. The prototype consists of a predetermined path for users to take during the user study.

User Study

A 3 stage user-study was conducted to understand participants' perception of AI art generators with integrated SRTs: (1) Using Nightcafe to familiarize themselves with the base UI; (2) Using our prototype while following a set of instructions and thinking aloud (Martin and Hanington 2012) to understand their thought process, emotional states, and perception; (3) Interview to evaluate specific aspects of the prototype's SRT design qualities. The interview included

[1] 5pt Likert scale: 1=Strongly Disagree (SD), 2=Disagree (D), 3=Neutral (N), 4=Agree (A) 5=Strongly Agree (SA)

[2] For simplicity, will refer to both as *art* in the rest of the paper.

[3] While color associations have been studied in the past (Elliot and Maier 2007), the lack of consistency in color relations (Adams and Osgood 1973), and lack of color study around eco-associations motivated us to include this question. Furthermore, approaching these questions from an exploratory RtD perspective, we wanted the design to rely and be informed by empirical data.

[4] Data visualizations do not contain real measurement data - we only focused on exploring visualization strategies.

[5] Prototype can be accessed at this link.

[6] Nightcafe is a well known multi-model online AI art generator.

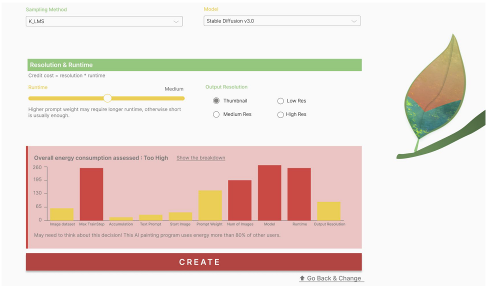

Figure 2: The SRT Prototype

questions that can broadly be divided into 3 areas of investigation: *Attention, Knowledge/Information,* and *Feeling/Emotion*. The following areas were evaluated on a Likert scale of 1(lowest) to 5 (highest): (A) Effectiveness in drawing *attention* to (UTQ_1), providing *information* about (UTQ_2) , and (UTQ_3) increasing *understanding* of the environmental impact; Anticipated effectiveness in *influencing* behavior change towards reduced consumption (UTQ_4); *Feeling* about different design elements, and emotions experienced during the whole interaction (UTQ_7); Evaluating the trade off between quality (1) or sustainability (5) (UTQ_8). 6 participants were involved in the user tests. P_1 and P_6 is from Italy, P_2 from Sweden, P_3, P_4, and P_5 from Finland. P_1 and P_3 identify as male, the rest as female. P_1, P_2, P_3, P_4, and P_6 are professional artists, while P_5 is a art hobbyist.

Results & Analysis

Survey Results

Color Associations Emphasize Visual Culture Rather than Nature 72% (n_{10}=18) reported some shade of green represented low power consumption and 60% of them (n_{11}=12) reported choosing green as they have seen it depict eco-friendliness. Only 17% (n_{12}=3) associated it with nature - e.g. *nature is green*. One respondent specifically wrote the associations were a *"cultural agreement"*. 88% (n_{15}=22) associated red with high power consumption and justified the choice by association to danger/concern/alarm (n_{17}=9), or a negative environmental labels (n_{18}=6). Similar to green, very few participants associated it with natural phenomena such as heat/warmth/fire (n_{19}=3). Respondents were split between oranges (n_{20}=10) and yellow (n_{21}=11) for middle level power consumption 47% associating orange (n_{22}=5) or yellow (n_{23}=5) with being used as a *middle color* between green and red, with some specifically recalling the orange in traffic lights (n_{24}=3). Interestingly, some picked green (n_{25}=5), including the respondent who reported being red-green colorblind; the reason for picking this color however is not entirely clear with with 2/5 responding that it was the middle/neutral color. When analyzing the results, it is evident that majority of the respondents had constructed an association through exposure to certain type of visual culture (majority reported seen it used in a similar context in society, in contrast to having seeing it in natural environments). This indicates, that color associations are dynamically constructed in contemporary culture - and rely less on natural representations.

Pictograms Rely on Symbolic Association to Nature Pictogram of a leaf was among the top 3 choices to represent sustainability due to their association with nature/environment (n_{25}=18). This is in stark opposition to color discussed in the previous section. Symbols closely associated with natural elements (flower, polar bear, sun, butterfly, drops, tree, leaf) were picked more often (37 votes for 8 pictograms) than man-made elements (energy, globe, light bulb, bicycle) (9 votes for 4 pictograms). The cloud pictogram was the only nature pictogram that was never picked, perhaps due to its greater association with *cloud computing* or *cloud storage*; there is however, no clearly discernible reason.

Preference for Granularity in Data Representations Though the bar chart was chosen by almost half the sample (n_9=12), all respondents (irrespective of graph style) noted that their choices were due to (1) ease of quick parsing, and (2) ease of understanding where and when power consumption has spiked/sunk. This aligns with literature where people want more granular transparency (Padgett et al. 2008), i.e., simply informing users about low/high consumption is not enough. Thus, future design patterns could include the provision of high granularity or customization to increase effectiveness of communicating sustainability-related information.

Insights from the User Testing

Colors and Data Visualization Are More Important than Pictograms All participants immediately understood what the colors indicated, P_6 reported the colors were all *"very universal"*. The use of red also made participants think more about their actions and made them feel bad about their high energy usage. Color was also the highest rated[7] design characteristic, both individually (Bringing Attention=4.34, Informing=4.5, Increasing Awareness=4.5, Effective in potentially changing behavior=4.5) and on average ($\mu = 4.46$); followed by data visualization (Bringing Attention=4.5, Informing=4.17, Increasing Awareness=4.17, Effective in potentially changing behavior=4.5, $\mu = 4.34$). Participants consistently reported that the data visualization *"[was] the most interesting"* (P_3), *"[made me] want to do better"* (P_2), *"gives me the curiosity to discover precisely which [choice] will change what."* (P_1). However, all participants reported wanting data to be relatable *(ref. Information Should be Relatable)*. The pictogram was barely noticed by participants and the least important (Bringing Attention=3, Informing=3.5, Increasing Awareness=3.5, Effective in potentially changing behavior=2.67, $\mu = 3.17$), with only P_4 ($\mu_{symbols} = 4.5$) and P_5 ($\mu_{symbols} = 4.75$) rating it high. P_1 said he *"was not really impressed"* and *"did not give attention to that"*, P_2 said about the leaf: *"didn't notice the...flower withering"*, P_6 *"forgot about them"*, and P_4 had

[7]All scores are out of 5. Those with recurring decimals are rounded up.

to be shown the pictogram again when asked to rate it as they forgot it completely.

Information Should be Relatable As discussed in *Colors and Data Visualization Are More Important than Pictograms*, participants were highly interested in seeing the data of their usage and choices. However, participants expressed dismay and confusion about the values provided: *"500kW....I don't know if it's a lot..."* (P_2), *"they [kW values] can be really abstracted. Okay. 100 kilowatts. So what, what does it mean?* (P_3), *"I would like to have some data ... [like] keeping your television on for two hours.."* (P_1), *"how is it in relation to all the energy consumption ... if I make a video to YouTube..."* (P_4). Providing data in relatable terms similar to Huang, O'Neill, and Tabuchi would allow for the data to be more effective in informing participants about their consumption. Participants also provided other ideas for more effective communication, such as implementing a virtual *bot* that informs you of the effects of your choices and corresponding parameters (P_3), or using emoticons (P_6).

Evoking "Negative" Emotions for Good All participants reported feeling bad about having high energy consumption being reported in the prototype. However, they all reported that this *bad feeling* was positive as it made them want to go back and change their outputs to be more *green*. P_2 compared the experience of getting red to *loosing a game*. P_1 provided the analogy of someone asking him to pickup litter from the ground: even if he feels bad about it, its for a good reason.

Systems Could Provide Low-Energy Settings for Experimentation and Exploration Phases of Creative Endeavors All participants (except P_3) wanted to go back and change their outputs to lower the energy use, though P_3 had expressed that he was limited by a static prototype. Although participants remained divided about preferring quality or sustainability ($\mu = 2.5$), they all suggested they would use lower consuming settings during initial experimentation or drafts, and use the highest quality settings for final deliverables. This insight also indicates that future work on VG-AI SRTs could take into consideration the various stages of creative work (Jääskeläinen, Pargman, and Holzapfel 2022a) and tailor capabilities and offerings to those specific stages. However, artists expressed concern over this method of working due to the randomness of VG-AI generation, and the experimental nature of creative work, *viz.* the randomness of digital noise inherent to G-AI is at odds with other more "controllable tools", such as Photoshop. However, optimistically all participants reported that they would want to consume less energy after using our prototype and became more informed about their power consumption, which shows promise regarding future work in SRTs.

Discussion & Conclusion

In this paper, we have provided preliminary knowledge on designing SRTs for VG-AI through an empirical design research study that included surveys, prototyping of an SRT tool, and user tests. To summarize the results, colors and data representations were more important than symbols in communicating sustainability information, and there was a clear connection of color association to contemporary visual culture. Pictograms, in contrast, relied on symbolic association to nature - and participants referred to ones that had a stronger association to nature. When it came to data representation, users preferred granular and relatable representation, such as comparing the consumption to X hours of watching YouTube videos. Participants also experienced feeling "bad", but described these emotions as something positive - and were willing to experience them in a process of becoming more aware of the sustainability implications of their work. Our study showed the artistic community responding positively to addressing environmental sustainability aspects of their practice - and that will be important going forward with the research agenda concerning SRTs. Further, these positive attitudes may help people to be, for example, more receptive to presented information, or more willing to change their behaviors; it may even be a predicament that needs to be met for SRTs designs to have net positive effects on the users. To acknowledge some of the limitations of this study, we would like to return to the complex nature of behavior change briefly discussed in *SRTs and The Complexity of Behavior Change for Sustainability*. Firstly, it is challenging to confirm if SRTs would in fact inflict behavior change without longitudinal studies with ethnographic and contextual observations of artists' work practices. Furthermore, we can anticipate that there are several factors outside the scope of SRTs that incentivize and drive people towards more or less sustainable ways of using VG-AIs. For example, it has been pointed out that higher socio-economic position likely enables people to use these systems to a greater extent, as they are primarily emerging in the Global North (Jääskeläinen, Holzapfel, and Åsberg 2022). Furthermore, one of the important tensions that should be acknowledged going forward with SRTs is that the underlying assumption that humans are rational agents and will change their behavior when prompted with information is weakly grounded. Thus, we argue that more research should be directed towards emotional aspects of SRTs and how we might inflict emotional experiences that commit users to certain practices on a deeply personal level - including the question: ***How to make artists feel guilty in a good way?***

Acknowledgements

This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program – Humanities and Society (WASP-HS) funded by the Marianne and Marcus Wallenberg Foundation (Grant 2020.0102).

References

[Adams and Osgood 1973] Adams, F. M., and Osgood, C. E. 1973. A Cross-Cultural Study of the Affective Meanings of Color. *Journal of Cross-Cultural Psychology* 4(2):135–156.

[Anthony, Kanding, and Selvan 2020] Anthony, L. F. W.; Kanding, B.; and Selvan, R. 2020. Carbontracker: Tracking and Predicting the Carbon Footprint of Training Deep Learning Models. arXiv:2007.03051 [cs, eess, stat].

[Bang, Torstensson, and Katzeff 2006] Bang, M.; Torstensson, C.; and Katzeff, C. 2006. The PowerHhouse: A Persuasive Computer Game Designed to Raise Awareness of Domestic Energy Consumption. In IJsselsteijn, W. A.; de Kort, Y. A. W.; Midden, C.; Eggen, B.; and van den Hoven, E., eds., *Persuasive Technology*, Lecture Notes in Computer Science, 123–132. Berlin, Heidelberg: Springer.

[Bender et al. 2021] Bender, E. M.; Gebru, T.; McMillan-Major, A.; and Shmitchell, S. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. Virtual Event Canada: ACM.

[Biørn-Hansen, Barendregt, and Andersson 2020] Biørn-Hansen, A.; Barendregt, W.; and Andersson, D. 2020. Introducing Financial Data and Groups in a Carbon Calculator: Issues With Trust and Opportunities for Social Interaction. In *Proceedings of the 7th International Conference on ICT for Sustainability*, 11–17. Bristol United Kingdom: ACM.

[Brynjarsdottir et al. 2012] Brynjarsdottir, H.; Håkansson, M.; Pierce, J.; Baumer, E.; DiSalvo, C.; and Sengers, P. 2012. Sustainably unpersuaded: how persuasion narrows our vision of sustainability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, 947–956. New York, NY, USA: Association for Computing Machinery.

[Dhar 2020] Dhar, P. 2020. The carbon impact of artificial intelligence. *Nature Machine Intelligence* 2(8):423–425. Number: 8 Publisher: Nature Publishing Group.

[Elliot and Maier 2007] Elliot, A. J., and Maier, M. A. 2007. Color and Psychological Functioning. *Current Directions in Psychological Science* 16(5):250–254. Publisher: SAGE Publications Inc.

[Froehlich, Findlater, and Landay 2010] Froehlich, J.; Findlater, L.; and Landay, J. 2010. The design of eco-feedback technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1999–2008. Atlanta Georgia USA: ACM.

[Gideon 2012] Gideon, L., ed. 2012. *Handbook of Survey Methodology for the Social Sciences*. New York, NY: Springer New York.

[Grover, Emmitt, and Copping 2019] Grover, R.; Emmitt, S.; and Copping, A. 2019. Reflecting on sustainability: coproducing a critical framework for sustainable design in the architectural studio. *Higher Education Pedagogies* 4(1):41–63. Publisher: Routledge _eprint: https://doi.org/10.1080/23752696.2019.1624589.

[Holmes 2009] Holmes, T. 2009. *Eco-visualisation: Combining art and technology to reduce energy consumption*. Thesis, University of Plymouth. Accepted: 2013-11-20T14:22:02Z.

[Huang, O'Neill, and Tabuchi 2021] Huang, J.; O'Neill, C.; and Tabuchi, H. 2021. Bitcoin Uses More Electricity Than Many Countries. How Is That Possible? *The New York Times*.

[Jääskeläinen, Holzapfel, and Åsberg 2022] Jääskeläinen, P.; Holzapfel, A.; and Åsberg, C. 2022. Exploring More-than-Human Caring in Creative-Ai Interactions. In *Nordic Human-Computer Interaction Conference*, 1–7. Aarhus Denmark: ACM.

[Jääskeläinen, Pargman, and Holzapfel 2022a] Jääskeläinen, P.; Pargman, D.; and Holzapfel, A. 2022a. On the Environmental Sustainability of Ai Art(s). In *Eighth Workshop on Computing within Limits 2022*. LIMITS.

[Jääskeläinen, Pargman, and Holzapfel 2022b] Jääskeläinen, P.; Pargman, D.; and Holzapfel, A. 2022b. Towards sustainability assessment of artificial intelligence in artistic practices. arXiv:2210.08981 [cs].

[Kefalidou et al. 2015] Kefalidou, G.; Skatova, A.; Shipp, V.; and Bedwell, B. 2015. The Role of Self-Reflection in Sustainability. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, MobileHCI '15, 1030–1033. New York, NY, USA: Association for Computing Machinery.

[Kollmuss and Agyeman 2002] Kollmuss, A., and Agyeman, J. 2002. Mind the Gap: Why do people act environmentally and what are the barriers to pro-environmental behavior? *Environmental Education Research* 8(3):239–260. Publisher: Routledge _eprint: https://doi.org/10.1080/13504620220145401.

[Lacoste et al. 2019] Lacoste, A.; Luccioni, A.; Schmidt, V.; and Dandres, T. 2019. Quantifying the Carbon Emissions of Machine Learning. arXiv:1910.09700 [cs].

[Mankoff et al. 2007] Mankoff, J.; Matthews, D.; Fussell, S. R.; and Johnson, M. 2007. Leveraging Social Networks To Motivate Individuals to Reduce their Ecological Footprints. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*, 87–87. ISSN: 1530-1605.

[Martin and Hanington 2012] Martin, B., and Hanington, B. M. 2012. *Universal methods of design: 100 ways to research complex problems, develop innovative ideas, and design effective solutions*. Beverly, MA: Rockport Publishers, digital ed edition.

[Mehonic and Kenyon 2022] Mehonic, A., and Kenyon, A. J. 2022. Brain-inspired computing needs a master plan. *Nature* 604(7905):255–260. Number: 7905 Publisher: Nature Publishing Group.

[Padgett et al. 2008] Padgett, J. P.; Steinemann, A. C.; Clarke, J. H.; and Vandenbergh, M. P. 2008. A comparison of carbon calculators. *Environmental Impact Assessment Review* 28(2):106–115.

[Strengers 2014] Strengers, Y. 2014. Smart energy in everyday life: are you designing for resource man? *Interactions* 21(4):24–31.

[Zimmerman and Forlizzi 2014] Zimmerman, J., and Forlizzi, J. 2014. Research Through Design in HCI. In Olson, J. S., and Kellogg, W. A., eds., *Ways of Knowing in HCI*. New York, NY: Springer. 167–189.