

Designers Should Evaluate Their Work. You say those are scissors you are running with, but do they even cut? Author post-print copy

Sander Hermesen^{a*}

^aUtrecht University of Applied Sciences, Utrecht, Netherlands

*Corresponding author e-mail: sander.hermesen@hu.nl

Abstract: Designed artefacts for health and wellbeing presuppose an actual impact on people's health or wellbeing; similarly, proposed improvements to design processes presuppose improved design processes. Unfortunately, design(research)ers often do not evaluate the efficacy of their designed artefacts and processes. This paper argues that it is essential to evaluate the assumed effects of designs and processes, to avoid common pitfalls and encourage the growth of knowledge, as a responsibility towards the users of our designs, and to enable an exchange with adjacent fields. To encourage doing so, the paper offers recommendations for evaluation strategies that fit designerly methods.

Keywords: Evaluation, Design process, Design Efficacy

1. Designers don't evaluate

Designers often display an ambiguous relationship with evaluating their designs. Although much-cited design researchers such as Papanek (1991, p. 13–15) and Norman (2007) stress the importance of investigating the effects of designed artefacts on their users, this investigation more often than not fails to take place.

This is especially true for designs that aim for behaviour change, which is often the case when designing for health and wellbeing, and sustainability. Designed artefacts (products, services, communication designs) displayed at well-known outlets such as graduation exhibitions, design (research) conferences and funding platforms for creative projects, give the impression that the sky is the limit: design appears capable of totally transforming our lives. However, upon closer scrutiny, the presented designs are often hardly more than a hypothesis waiting for a test. Design processes and theoretical frameworks are even less often subject to rigorous evaluations and are often assumed true without testing.

This paper argues that this lack of evaluation is detrimental not only to the development and the position of the field of design (research), but also to the health and wellbeing of the users of the many designed artefacts, and to the development of and integration with adjacent fields. Finally, some suggestions for evaluation approaches that fit designerly methods are introduced.

2. Reasons (not) to evaluate

In the previous section, I claimed that many designed artefacts and processes remain without evaluation. But is this really the case? A systematic review of evaluative approaches in design research work would be beyond the scope of this positioning paper, but a brief look at the *design for next health*-track of the previous EAD conference confirms the statement. Of the 47 presented papers, 29 did not contain any form of evaluation of the designed artefacts or proposed frameworks whatsoever. Those that did mostly used methods that are easily susceptible to bias, such as case studies (9 papers), expert reviews (2 papers) and focus groups (1 paper). An interesting research question to answer would be why this is the case. When asked, designers often mention a lack of time or budget, or the scope of the commission. However, one suspects that it is more a matter of unwritten practices: evaluation is just not something designers *do*. Is thoughtful, rigorous scrutiny of the effects of one's work simply too *boring* for a field that would rather run with scissors?

All this is a shame, because there are many arguments to support rigorous evaluation. Firstly, it is our responsibility towards our own field. When we do not look critically at the efficacy of our works, we are likely to fall prey to common errors available to people who depend on gut feelings to inform their work: unsubstantiated intuitions, an overly rational or irrational view of human behaviour and cognition, and an over-reliance on rationality or perceived utility to drive behaviour change (Kelly & Barker, 2016). When we do not evaluate our work, we do not contribute to the growth of knowledge in the field of design and do not learn to avoid common pitfalls.

Secondly, and very much connected with the first argument, is the designers' responsibility towards the users of the designed artefact. (Potentially) ineffective designs can foster unwarranted hopes that health and environmental problems have an easy solution, or even inflict damage. This issue is most visible in the field of digital health solutions, colloquially known as the 'new wild wild west' (Lobelo et al., 2016), but certainly not uncommon for other fields of design for behaviour change.

Thirdly, evaluating makes it possible to exchange our findings with the behavioural and medical sciences – where rigorous evaluations are routinely expected for every intervention – and to ensure our place at the table where new developments in scientific research and evaluation are being shaped. Design (research) is certainly not the only area currently looking for ways to improve its methodological rigour. In health-related disciplines, the role of its gold standard of evaluation, the randomised controlled trial, is being discussed as it is too slow, too costly, and too inflexible (Mohr et al., 2015). Similarly, the current standard in process evaluations (Moore et al., 2015) does not take the benefits of approaches such as research through design, action research, or participatory design into account. It would be immensely valuable when design (research) would take part in the scientific discussions about the merit of new approaches in evaluation, not in the least because this would be an excellent opportunity to promote the use of designerly methods in other fields.

3. How to evaluate?

Evaluation of our designs and design processes should have a high validity and a low risk of bias. But the previous paragraph shows it would be unhelpful to simply copy current best practices of other disciplines. Our evaluation methods should take the holistic, context-driven and system-aware nature of design research into account.

When using quantitative data for evaluation, we can make use of recent developments in rapid research designs (McCallum, Rooksby, & Gray, 2017). Single case studies, for instance, rely on gathering many data points from a single individual, to not only test the efficacy of a design, but also

contextual influences - what works for whom at what moment (Dallery, Cassidy, & Raiff, 2013). Another promising approach is Multiphase Optimization Strategy (MOST; Collins, Murphy, & Strecher, 2007), which integrates evaluation into design phases and allows for different iterations and potential intervention components tested in small experiments.

Qualitative research is where design research is already showing its forte. However, there is still room for improvement. A relatively new method that minimizes demand characteristics and provide opportunities to involve the use context is Ecological Momentary Assessment (also known as experience sampling, cf. Stone, 2017). Finally, a great opportunity lies in using and developing qualitative methods for evaluating design processes, such as Systematic Process Analysis (Hall, 2006).

All in all, there is much to be gained in more rigorous evaluation of our design work. If we are going to run with scissors, we might as well make sure they are sharp.

References

- Collins, L. M., Murphy, S. A., & Strecher, V. (2007). The Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART). *American Journal of Preventive Medicine*, 32(5), S112–S118. doi:10.1016/j.amepre.2007.01.022
- Dallery, J., Cassidy, R. N., & Raiff, B. R. (2013). Single-Case Experimental Designs to Evaluate Novel Technology-Based Health Interventions. *Journal of Medical Internet Research*, 15(2), e22. doi:10.2196/jmir.2227
- Hall, P. A. (2006). Systematic process analysis: when and how to use it. *European Management Review*, 3(1), 24–31. doi:10.1057/palgrave.emr.1500050
- Kelly, M. P., & Barker, M. (2016). Why is changing health-related behaviour so difficult? *Public Health*, 136, 109–116. doi:10.1016/j.puhe.2016.03.030
- McCallum, C., Rooksby, J., & Gray, C. M. (2018). Evaluating the Impact of Physical Activity Apps and Wearables: Interdisciplinary Review. *JMIR mHealth and uHealth*, 6(3), e58. doi:10.2196/mhealth.9054
- Mohr, D. C., Schueller, S. M., Riley, W. T., Brown, C. H., Cuijpers, P., Duan, N., ... Cheung, K. (2015). Trials of Intervention Principles: Evaluation Methods for Evolving Behavioral Intervention Technologies. *Journal of Medical Internet Research*, 17(7), e166. doi:10.2196/jmir.4391
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., ... Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, 350(mar 19 6), h1258–h1258. doi:10.1136/bmj.h1258
- Lobelo, F., Kelli, H. M., Tejedor, S. C., Pratt, M., McConnell, M. V., Martin, S. S., & Welk, G. J. (2016). The Wild Wild West: A Framework to Integrate mHealth Software Applications and Wearables to Support Physical Activity Assessment, Counseling and Interventions for Cardiovascular Disease Risk Reduction. *Progress in Cardiovascular Diseases*, 58(6), 584–594. doi:10.1016/j.pcad.2016.02.007
- Norman, D. (2007). *Industrial Design: Claims Without Substance*. Retrieved on 26 September 2019 from https://www.jnd.org/dn.mss/industrial_design_c.html (Archived by WebCite at <http://www.webcitation.org/72icgEZVG>).
- Papanek, V. (1991). *Design for the Real World: Human Ecology and Social Change*. Revised second edition. London, UK: Thames & Hudson. ISBN: 0 500 27358 8
- Stone, A. (2017). Ecological Momentary Assessment and Experience Sampling. In: *The Palgrave Handbook of Survey Research*, pp. 563–571. doi:10.1007/978-3-319-54395-6_64

About the Author:

Sander Hermesen is senior researcher at the Research Group Crossmedial Communication in the Public Domain (Publab) at the Utrecht University of Applied Sciences; his work focuses on evaluating designs for behavioural change and increasing their efficacy, and on developing models and toolkits that make insights from the behavioural sciences available for designers.