Healthcare and Medical Devices, Vol. 51, 2022, 239–247
https://doi.org/10.54941/ahfe1002123

Considerations and Strategies for Operationalizing Heuristic Evaluation Work

Helen J.A. Fuller[1], Timothy Arnold[1,2], Kyle D. Maddox[1], and Kathleen Adams[1]

[1]Human Factors Engineering, Clinical Informatics and Data Management Office, Office of Health Informatics, Veterans Health Administration, Department of Veterans Affairs, Washington, DC, USA
[2]University of Michigan, College of Pharmacy, Ann Arbor, MI, USA

ABSTRACT

Heuristic evaluation (HE) is a popular usability inspection method that allows expert evaluators to document usability problems with the design of interfaces. The relatively low resource requirements and high utility of identifying usability issues and suggesting dimensions along which to correct them have made HE highly scalable. However, there are several limitations and concerns when operationalizing HE for a large enterprise, and education and communication along with combining HE with other usability techniques may greatly increase the utility of the work.

Keywords: Heuristic evaluation, Usability, Human factors, Healthcare, Systems engineering, Systems redesign, Risk management

INTRODUCTION

We are part of the Office of Health Informatics Human Factors Engineering group (OHI HFE) within the Veterans Health Administration (VHA), Department of Veterans Affairs (VA) that regularly uses Heuristic Evaluation (HE) to evaluate healthcare interfaces, including clinical templates in the electronic health record, for usability concerns. HE is an inspection-based usability method where practitioners review a system interface against a preselected set of design criteria (Nielsen and Molich, 1990). The activity of performing HE and delivering findings is generally referred to as a heuristic evaluation (HE). HEs often include both usability findings and ratings to prioritize those findings (Cook and Herout, 2015). Critical and careful reviews of information systems' interfaces and related interaction components are important to promote successful system operation (Smith and Aucella, 1983).

Clinical templates intended for enterprise release undergo HE review by OHI HFE as part of a VHA approval process. OHI HFE also utilizes HE as an inspection method for system interfaces already implemented. The approval process requires OHI HFE evaluation of the system interface for usability, safety, and design concerns, but OHI HFE chooses the method to use for the evaluation.

Two key indicators of user experience (UX) maturity are having UX expertise available to evaluate systems and the appropriate use of proper methods and techniques to achieve the organization's goals (Chapman and Plewes, 2014). It is important to critically review the effectiveness of methods used. The relatively low resource requirements and high utility of identifying usability issues and suggesting dimensions along which to correct them have made HE highly scalable. However, there are several limitations and concerns when operationalizing HE for a large enterprise. In this piece, we will present lessons learned from our group's HE work, including suggestions for addressing challenges we have faced. We hope to offer guidance—or at least empathy—for others seeking to operationalize HE.

SOME BENEFITS OF HE

HE can be both usable and useful in many circumstances. In general, HE is cost effective and easy to perform by practitioners with basic human factors training (Fu et al., 1998). It also can be performed at any time—or multiple times—during project development, including on an early prototype of the design, and the output is immediately useful to the designer (Stanton et al., 2017). Identifying usability issues early in the design process generally means they will be less expensive to fix; because the design is less finalized at this time, designers are often less resistant to changes.

HE findings may be informed by a different set of biases then other inspection and evaluation approaches (Fu et al., 1998). When used in ensemble with other approaches, HEs may help balance out biases throughout the design process. In the OHI HFE group, we have developed templates that make it more efficient to perform HEs and present findings. This form of standardization also makes it easier to compile HE findings and track program and development team responses.

We have observed that HE findings are also relatively easily understood by the groups that request a usability evaluation. The rating component helps these groups understand potential impacts and prioritize solutions. They can also give stakeholders language to use when discussing solutions with developers. This combination of ease of use and utility also makes HE easy to advocate for, and OHI HFE often receives requests for HE work from groups we have worked with in the past.

Beyond the usability findings themselves, a key benefit of HE is that it provides a structure for examining a system and a framework for facilitating a conversation. Following the process for an HE can help scope an evaluation and frame the questions that the team will ask. The initial framework HEs provide helps define the rules of engagement and lowers barriers to project momentum.

When HFE practitioners present findings from an HE, they may bring up paradoxes and encounter the need for trade-off decisions, and the resulting discussion can actively engage stakeholders and subject matter experts (SMEs) in the decision-making process (Pritchett and Strong, 2016). This may help to support interprofessional team decisions, promote coproduction of solutions, and engender buy-in from all parties. Facilitation of the evaluation

and discussion by an HFE practitioner can ease tensions and provide a more neutral space for decision-making.

The structure of HEs creates a learning opportunity for all participants. Developers and stakeholders learn about usability techniques and design best practices, but HFE practitioners may also acquire valuable knowledge about the context of use and constraints of the tools and systems under evaluation. In addition, it may help build relationships between practitioners and stakeholders and facilitate the emergence of participatory design communities.

KEY LIMITATIONS OF HE

While HE can be useful and powerful if deployed correctly, there are some key limitations to using HE for usability evaluations. Some limitations—as well as advantages—have been discussed previously (e.g., Novick and Hollingseed, 2007). In our experiences, HE can cause the practitioner to focus on interface details and ignore greater structural issues related to work systems and their match to overall operator goals. Thus, it is important to consider a full complement of contextual analysis and usability methods when determining the best strategy for design and evaluation.

Limited Focus of the Evaluation

The structure of HEs may cause reviewers to evaluate what is present but ignore what might be missing. A reviewer might find that the information presentation is correct but fail to note that not enough information or the right information is included. For example, a template for medication education may meet all design heuristics for presenting information to the user but may completely fail to account for some less common situations in which a clinician might want to record medication education.

Put another way, HE may promote a surface-level review rather than encouraging the reviewer to consider the foundational structure of the design and the translation from goal into product. Performing an HE does not necessarily answer the question of whether the design meets the underlying goal. In fact, there is a tendency for HFE practitioners to assume that the design is correct and that one need only consider the usability of the interface when conducting an HE.

This points to the need for SME involvement in HE. Nielsen (1992) found that evaluators who had expertise in both usability and the type of interface being evaluated found more problems than evaluators who had expertise in either area alone. In industry, an HFE practitioner is often paired with an SME for reviews. This may be especially important for more specialized domains such as healthcare; an HFE practitioner is unlikely to have the necessary background to fully understand the domain and context of use. While stakeholders and developers may be able to offer some SME input, there is often a strong underlying motivation to push forward the proposed solution that may cause people to fail to consider the possibility of varying experiences and to overlook local capabilities.

Not a Stand-Alone Usability Method

The tendency of HE to focus an evaluation on the implementation of a design rather than looking at whether the correct thing was designed points to a need to consider HE as part of a suite of usability methods. There is great value in conducting early observation and systems studies to gather a baseline understanding of the current system as well as design work such as contextual analysis, environmental scans, and conceptual modeling during the design phase. In particular, conceptual modeling can help us to ensure that we are designing the right thing (Johnson and Henderson, 2002), and this can also aid in evaluating the design using HE. In addition, in many cases it is important to include representative end users in human factors evaluations such as usability tests and cognitive walkthroughs.

As HFE practitioners, we sometimes encounter the belief that usability is a box to check or that a design can be made usable simply by performing an HE. HFE practitioners can play a role in offering education to design team members to help them understand the benefits and limitations of different types of usability findings. As Nielsen and Molich (1990) pointed out, not all usability "problems" identified through HE will cause issues in actual use; the decision of whether and how to address a problem through redesign requires discussion and trade-off analysis with the entire team.

Fostering communities that include nonpractitioners in longitudinal participatory design efforts may promote an understanding of the importance of an ensemble of approaches throughout an integrated lifecycle. Nonpractitioners may have more opportunities to help HFE practitioners understand and collect important patient outcome measures. ISO 9241-210 describes the importance of a multidisciplinary design team, including stakeholders and user groups (ISO, 2010). Learning about the nuances of usability findings may help teams to make more informed design decisions based on usability findings as well as garner support for additional usability work. While some usability work is likely better than no usability work, it is important to promote selection of study methodologies based on the concerns we are addressing and to use findings appropriately.

Oversimplification of Human Factors Work

Because HE is generally recognizable, understandable, and perceived as useful by customers and stakeholders, there is a danger that they may misinterpret it as being the main type of usability work that HFE practitioners can offer. HEs are relatively low-cost, so they may represent a large proportion of projects for an HFE group. Encountering HEs repeatedly may increase the visibility of HE and perpetuate this assumption among customers. This may lead to a feedback loop where demand for HEs increases and HFE practitioners become increasing comfortable with performing HEs. It can be important for practitioners to educate customers and stakeholders on alternative human factors evaluation methods and advocate for their use when appropriate.

This leads us to a broader, perhaps philosophical, debate around human factors and usability methodologies. Again, it seems likely that some usability testing is better than no usability testing for most projects. The rise of lower

resource usability methods such as HE and rapid usability testing (e.g., Russ et al., 2010) make it possible to conduct some usability work over a larger number of projects with the same resources.

It is also important to consider the usability of human factors approaches and the findings they generate to people who are not human factors experts. Because the output of an HE is relatively usable to the design team, the overall impact of doing an HE may be greater than utilizing other human factors methods, even ones that could identify more—or more severe—usability issues.

Perhaps the right mindset is to view HE as an introduction to human factors for design teams and to gradually add and then integrate more complex approaches as they begin to understand the value of human factors work. Given the many different approaches to usability work, a combination or hybrid approach may yield the greatest value to clinician users and patients. While teaching about human factors and usability, it is also important to listen to the customer to learn about their needs and values. Integrating participatory human-centered design (HCD) practices into the daily and weekly work of interprofessional teams is a step towards understanding challenges while encouraging acceptance and adoption. Providing designers with a human factors framework, usability vocabulary, and evaluation experience may lead to a blending of values and offer increased awareness of the benefits of HCD practices.

FURTHER CHALLENGES AND SOME POSSIBLE SOLUTIONS

There are some additional challenges—or at least considerations—for an HFE group seeking to maximize the effectiveness of HE. Some of these relate to the limitations mentioned above, and others seek to maximize the benefits described previously.

Selection of Heuristics

Many different sets of heuristics have been used in a variety of settings (Nielsen, 1994). While some sets of heuristics are general, others are tailored for use in specialized areas. For example, Zhang and colleagues (2003) suggested usability heuristics for evaluating patient safety of medical devices, and Miller and colleagues (2018) matched heuristics to clinical decision support recommendations. In our practice in healthcare informatics, we often find ourselves relying on a subset of the Nielsen (2020) heuristics but also sometimes identify usability issues that do not map readily to those heuristics. Our experience seems to support Hermawati & Lawson's (2016) finding that some industries may require domain-specific heuristics.

Navigating Ratings Systems

Most HEs include a rating of severity for each finding, though different authors have recommended different scales and anchor phrases. The rating is generally based on a combination of severity, probability, and persistence (Zhang et al., 2003). The number of severity levels used has varied between groups and over time. Nielsen (1992) described two levels (major and minor),

but more recent HE papers tend to list five levels (e.g., Zhang et al., 2003). In our practice, we generally use three severity levels: minor, moderate, and serious (Cook and Herout, 2015). Unfortunately, a 3-point scale can quickly become a 1-point scale if there is a very high barrier to awarding a "serious" rating and a disinclination to give a "minor" rating out of fear a finding will be dismissed.

Indeed, ratings may be biased by the reviewer's belief about how the rating will be interpreted. While ratings are intended to assist with prioritizing redesign efforts, in practice they sometimes are viewed as binary, with moderate and serious ratings receiving attention and minor ratings being ignored. This points to another opportunity for education on the part of the practitioner to teach the design team the meaning of the ratings system and to help them understand how multiple minor issues may contribute significantly to users' dissatisfaction and loss of confidence with a system. An option for improvement with the ratings system might be to add "timing to implementation" as a dimension to address findings that are serious but difficult to implement in the near-term. In addition, involving the design team in the prioritization discussions and even considering participatory design of a ratings system are areas for future exploration.

We have observed a few categories of customer responses to findings when they do not plan to make changes. The first is technical—that a recommended change simply is not possible. Nielsen described asking evaluators to keep in mind basic technical limitations of the system and to not include criticism related to those aspects (Nielsen, 1992). While we generally follow this practice, there are times when there is value in documenting when technical limitations of the system make it impossible to follow design best practices, to have as a record in case those limitations change and to advocate for that change. There may also be clinical reasons to avoid a change, and some things may need to be done in a certain way from a business perspective. SME coreview and trade-off discussions with the design team could help us arrive at better recommendations. Finally, there are times the design team rejects a change because they do not believe the design will cause a problem. In this case, combining HE with other usability methods such as user testing could help to further inform decision-making.

Promoting Upstream Analysis

As discussed previously, HE tends to focus on design details, sometimes at the expense of understanding the efficacy of the overall system. How then do we capture what is missing or promote evaluation of what might be missed during an HE? A first step to addressing this gap is to actively ask what should be in the design rather than merely evaluating what is there. This points to a need to engage in activities such as evaluation of context of use, environmental scans, and conceptual modeling at the beginning stages of development. This early involvement has the added benefit of identifying changes at a time when they are more likely to be addressed; once work has been done and structure put in place, redirection becomes less likely.

While it would be ideal to have HFE practitioners involved in the early stages of a design, this is not always possible or practical. To evaluate the design at a later stage, the HFE practitioner might ask intake questions and perhaps complete a validation step to try to match the underlying information model to the users' needs. These questions should address what—if anything—has been done in the way of environmental scan, user analysis, or other systems studies and may be asked through an email intake form and/or a kickoff call.

Ideally, a committee including stakeholders and SMEs would meet to consider proposed designs prior to development. This should include formal evaluation and documentation of how well the product under evaluation mirrors the process it is attempting to support. HFE practitioners could facilitate these discussions or provide materials to help guide the process.

Early analysis helps with identifying data objects that are being captured or that should be captured that support the overall goals of the system. This may help in designing safer and more efficient work systems while collecting information for understanding the effects and measuring the value of human factors work.

Promoting a Productive Atmosphere

The structure of an HE and the way that findings are commonly presented can sometimes promote an adversarial atmosphere, where the person who led the design and development of the system under evaluation feels defensive when hearing about problems or shortcomings with the design. It is important to develop a presentation structure and techniques to limit this dynamic, because it can be difficult to regain a sense of community and return to productive discussion when this happens.

Good communication skills can help encourage cooperation during the presentation of HE findings. It is also important for the HFE practitioner to recognize that they may not have the full contextual picture of how a system will be used or what sociotechnical constraints may exist. This should help them to present findings with humility and compassion and collaborate in arriving at a solution.

Our group has identified some concrete strategies to promote a positive dynamic. Prior to the presentation of findings, we send a copy of the HE results for review. During the discussion, we start with sharing strengths of the design. Besides starting the discussion on a positive note, this also helps the designers to better understand and identify design choices that work well so that these will be propagated in the future. When presenting findings, we try to describe the problem identified and what a good outcome should achieve rather than telling the customer how they should change their design. This generally leads to a discussion where the HFE practitioner works together with the customer to achieve the desired outcome, resulting in greater buy-in, better collaboration, and often improved design recommendations.

Some Final Thoughts

The iterative nature of the HCD process is one of its strengths. In this spirit, we are never fully satisfied with our own work process, including HE, and are always seeking improvement.

We might look to mission engineering methods to ensure we are creating the right capabilities and work products for our healthcare system. Rather than a myopic focus on how we do HE and whether we are meeting all guidelines, perhaps we should take a step back and consider why we do HE and how we add value with our work. The process and conversation surrounding an HE are often as important as the final product. They help to create partners of the entire design team and provide an opportunity to convert others to an HCD way of thinking. Discussing design trade-offs offers a chance to educate people who will continue to make design decisions, and this ultimately may provide more value than any specific usability recommendations.

Our work has also taught us the importance of collaboration with a multidisciplinary group. Each team member can offer valuable, often irreplaceable, perspectives. From clinical knowledge to point-of-care workflow information to implementation strategies, HFE practitioners can gain as much from the team as the team can from them.

Finally, it is important to consider the organization's operations tempo in any human factors work. This issue is partially addressed by gathering the right team players so that necessary knowledge and decision authority are available in real time. HFE practitioners need to remain mindful of organizational constraints, including policy, technology, and timelines, and present recommendations that will be both useful and usable.

By accounting for the strengths as well as the limitations of HE and choosing complementary methods as applicable, by thoughtfully constructing the design and review teams, and by remaining mindful of constraints, we can use HE to identify usability issues and areas for design improvement. Through constant reexamination of our own human factors efforts, we can work to balance different methodologies and optimize the application of HE in usability work.

ACKNOWLEDGMENT

We thank patients and staff at VHA for inspiring this work, OHI HFE practitioners for their contribution to the discussion, and everyone in the VHA Office of Health Informatics for their commitment to Veterans. The authors declare no conflict of interest. The opinions expressed in this article are those of the authors and do not necessarily represent those of the Department of Veterans Affairs.

REFERENCES

Chapman, L. and Plewes, S. (2014). A UX maturity model: Effective introduction of UX into organizations. In *International Conference of Design, User Experience, and Usability* (pp. 12–22). Springer, Cham.

Cook, A. and Herout, J. (2015). Developing a usability ranking system for findings in health information technology products. In *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care* (Vol. 4, No. 1, pp. 23–28). Sage India: New Delhi, India: SAGE Publications.

Fu, L., Salvendy, G. and Turley, L. (1998). Who finds what in usability evaluation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 42, No. 19, pp. 1341–1345). Sage CA: Los Angeles, CA: SAGE Publications.

Hermawati, S. and Lawson, G. (2016). Establishing usability heuristics for heuristics evaluation in a specific domain: Is there a consensus?. *Applied Ergonomics*, 56, pp. 34–51.

ISO 9241–210. (2010). *Ergonomics of human system interaction - part 210: Human-centred design for interactive systems*. International Organization for Standardization, Switzerland.

Johnson, J. and Henderson, A. (2002). Conceptual models: begin by designing what to design. *Interactions*, 9(1), pp. 25–32.

Miller, K., Capan, M., Weldon, D., Noaiseh, Y., Kowalski, R., Kraft, R., Schwartz, S., Weintraub, W.S. and Arnold, R. (2018). The design of decisions: Matching clinical decision support recommendations to Nielsen's design heuristics. *International Journal of Medical Informatics*, 117, pp. 19–25.

Nielsen, J. and Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 249–256).

Nielsen, J. (1992). Finding usability problems through heuristic evaluation. In *Proceedings of the SIGCHI conference on Human factors in Computing Systems* (pp. 373–380).

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 152–158).

Nielsen, J. (2020). 10 Usability Heuristics for User Interface Design. Nielsen-Norman Group Website: https://www.nngroup.com/articles/ten-usability-heuristics/

Novick, D. G. and Hollingseed, T. (2007). Usability inspection methods after 15 years of research and practice. In *Proceedings of the 25th Annual ACM International Conference on Design of Communication* (pp. 249–255).

Pritchett, A.R. and Strong, A.C. (2016). Integrating cognitive engineering into industry design teams. *Journal of Cognitive Engineering and Decision Making*, 10(2), pp. 134–137.

Russ, A.L., Baker, D.A., Fahner, W.J., Milligan, B.S., Cox, L., Hagg, H.K. and Saleem, J.J. (2010). A rapid usability evaluation (RUE) method for health information technology. In *AMIA Annual Symposium Proceedings* (Vol. 2010, p. 702). American Medical Informatics Association.

Smith, S.L. and Aucella, A.F. (1983). Design guidelines for the user interface to computer-based information systems. Bedford, MA: MITRE Corporation.

Stanton, N.A., Salmon, P.M., Walker, G.H., Baber, C. and Jenkins, D.P. (2017). *Human Factors Methods: A Practical Guide for Engineering and Design*. CRC Press.

Zhang, J., Johnson, T.R., Patel, V.L., Paige, D.L. and Kubose, T. (2003). Using usability heuristics to evaluate patient safety of medical devices. *Journal of Biomedical Informatics*, 36(1-2), pp. 23–30.