

Benchmarking Accessibility, Usability, and Universal Design

– What Kind of Society Are We Building?

Per-Olof HEDVALL^a, Lilian MÜLLER^b, Larry MCNUTT^c, Gerald CRADDOCK^d

^a *Mid Sweden University, Sweden;*

^b *Lund University, Sweden;*

^c *Technological University Dublin, Ireland*

^d *Centre for Excellence in Universal Design, NDA, Ireland*

ORCID ID: Per-Olof Hedvall <https://orcid.org/0000-0002-0962-2899>

Lilian Müller <https://orcid.org/0000-0003-4897-2629>

Larry McNutt <https://orcid.org/0000-0002-9331-0854>

Gerald Craddock <https://orcid.org/0000-0003-1106-4248>

Abstract. Can and should Universal Design (UD) be evaluated? We can check whether a website meets WCAG criteria, whether a door is wide enough, or whether someone completes a task. But these are accessibility and usability measures that tell us little about what UD, understood as something distinct from both, contributes at a broader scale. Existing UD evaluation tends to focus on the evaluation of two of its neighbours, accessibility and usability, and the field has yet to develop the tools to evaluate UD on its own terms. This paper proposes a Benchmarking Matrix that unburdens UD from compliance and evaluates each concept on its own terms: a 3×3 grid with accessibility, usability, and UD as columns and output, process, and societal impact as rows. Borrowing from the Positive Peace framework, it asks what conditions make UD possible and which indicators let us follow it over time. We tested the matrix in two workshops with about 15 researchers and practitioners, who filled seven cells (two output cells were set aside as already well established). Contrary to our expectation that the least familiar cells would draw the fewest inputs, the UD column, and especially the UD × Societal Impact cell, produced the richest and most novel input, including candidate proxy indicators such as trust in others and democratic participation. The matrix is a thinking tool, not a compliance checklist, and opens a path towards assessing UD as a societal phenomenon.

Keywords. Universal Design, accessibility, evaluation, societal impact, indicators

1. Introduction

How do we know whether UD is actually making a difference? The question sounds straightforward, but the field has yet to establish a shared answer to it. We can check whether a website meets WCAG criteria, whether a door is wide enough, or whether someone completes a task. However, these are accessibility and usability measures, and tell us little about what UD, understood as something distinct from both, contributes at a broader scale. This paper proposes a step toward UD evaluation on its own terms: a Benchmarking Matrix that distinguishes what accessibility, usability, and UD each requires, and at what level of impact each operates.

The paper takes as its foundation a harmonised model (Figure 1) for Accessibility, Usability, and Universal Design (here abbreviated AUUD), developed by Hedvall, Ståhl, and Iwarsson [1], where the three terms are defined as:

- Accessibility: To what extent can you access and use it? It is the measurable baseline, a legal and moral obligation.
- Usability: How effective, efficient and satisfying is the use? It is about the quality of use, seen from the individual's or group's perspective.
- Universal Design: What values is the design process built on? It is the value-based approach to design that should permeate all aspects of society, from the start.

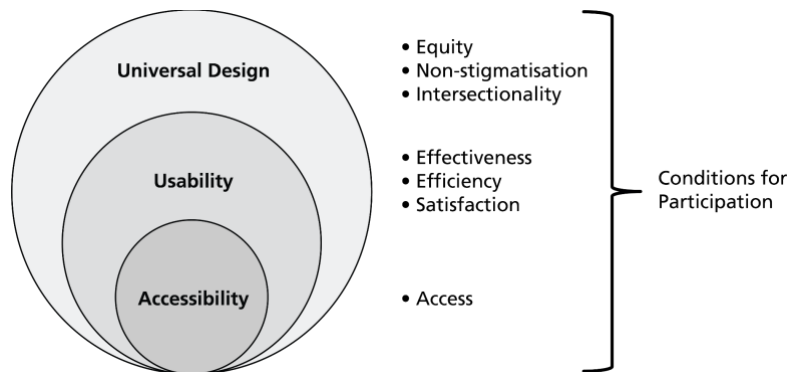


Figure 1. The three concepts accessibility, usability, and universal design, with their respective focus and how they relate to the conditions for participation [1].

Accessibility and usability are concepts with a long history of being assessed and measured, and are integral parts of standardisation and have established instruments, for example, the European Accessibility Act [2] (that incorporates EN 301 549 and WCAG) for accessibility, and ISO 9241-11 for usability. Ireland's Disability Act 2005 [3] defines universal design using the terms “accessed, understood and used”, which, for its time, were important in advancing inclusive environments in Ireland. However, as Steinfeld and Maisel have pointed out, UD “involves continuous improvement [...] including the often neglected outcomes of improved health and social participation” [4,p.29]. Now two decades later, we must therefore ask how UD can evolve to stay relevant and impactful, today and in the years ahead.

The picture regarding evaluation is more mixed when it comes to UD, and existing UD evaluation has a tendency to focus on accessibility (WCAG, ramp gradients, door widths) or usability (task completion, satisfaction scores). That is not UD evaluation. It is the evaluation of two of UD's close relatives, Accessibility and Usability [5].

But can and should UD be evaluated at all? A growing body of scholarship argues that reducing UD to metrics fundamentally misses its original mission. Some of the most pertinent critiques come from scholars who argue that UD's transformative potential has been undermined by its reduction to metrics, technical specifications, and compliance frameworks [6,7]. However, there may be a third path: not rejecting benchmarking, but reorienting it, and this way making it possible to evaluate UD on its own terms. What if we have been trying to evaluate the wrong things? Not because evaluation is impossible, but because we have been looking in the wrong place?

1.1. Unburdening Universal Design from Compliance

When the Center for Universal Design at NC State released the 7 Principles of Universal Design in 1997, the framework was intended to "evaluate existing designs, guide the design process and educate both designers and consumers" [8]. What it explicitly did not include were metrics, standards, or benchmarks. Edward Steinfeld and Jordana Maisel, recognising these (and other) limitations, developed The 8 Goals of Universal Design in their 2012 book *Universal Design: Creating Inclusive Environments*. Their explicit aim was to "define the outcomes of UD practice in ways that can be measured and applied to all design domains" [4].

A range of assessment instruments and standards now exists, though they tend to operationalise UD differently. What these evaluate varies considerably, but few of them systematically assess social integration, dignity, or equity. Hedvall, Ståhl & Iwarsson [1] argue that accessibility, usability, and UD are distinct concepts, where all three are needed and should be evaluated together, but each requires its own evaluation logic. Regarding the evaluation of UD, they write:

"UD was originally conceptualised as being about societal attitudes, democracy, citizenship, equity, flexibility and social inclusion. Consequently, if UD at all should be evaluated, it ought to be done on the premises of how such factors are assessed and not based on operationalizations of accessibility and usability" [1,p.6753].

By evaluating accessibility on its terms, usability on its terms, and UD on its own terms, UD can be unburdened from compliance. The question, then, for UD, is not how well an environment meets accessibility standards or whether people are satisfied when using it, but what kind of society we are building. Measuring ramp gradients is an easy question, and one we already know a great deal about. A lot less has been done regarding how to evaluate UD in relation to attitude changes, democratic participation, and equity. This is, at root, a question about who we design for, and about what it means to belong, in architecture, in products, and throughout our educational and societal structures [9].

Against this background, the present study aims to contribute new insights regarding how UD can be evaluated, if unburdened from accessibility and usability compliance metrics. It reports on the development and early tests of an evaluation matrix bringing together accessibility, usability and UD, and evaluating them in terms of output, process, and societal impact. The research questions that have guided the study are:

1. How can accessibility, usability and UD be evaluated together?
2. How can UD be evaluated on its own terms?

2. Developing and Testing the AUUD Evaluation Matrix – Methods and Results

The development of the matrix took place in a four-year research project on universal design, equity, and equality. In the project, a group of researchers from design science, architecture and linguistics worked in close collaboration with staff from Swedish municipalities and regions. The section below is divided into three parts. We first describe the development of the matrix, then how it was tested, with early results.

2.1. Developing the AUUD Evaluation Matrix

The matrix (Table 1) has three columns (accessibility, usability, UD) and three rows (output evaluation, process evaluation, societal impact evaluation). Current UD evaluation tends to focus on the output row: whether environments meet standards, whether products are usable. Thus, we already know a great deal about the two grey cells, Accessibility $\times$ Output and Usability $\times$ Output. Many of the open questions can be found in the other seven cells.

Table 1. A matrix for capturing Output Evaluation, Process Evaluation and Societal Impact Evaluation of Accessibility, Usability, and Universal Design.

	Accessibility	Usability	Universal Design
Output			
Process			
Societal impact			

The development took Hedvall et al.'s model (Figure 1) as its base, with the concepts as the three columns. UD is understood as ultimately being about societal impact [1]. With such a conceptualisation, the practical challenge became: How does one evaluate, for instance, attitude change, democratic participation, and equity?

Similar factors and societal values are already being evaluated outside of UD, through thoughtful proxy indicators. The Positive Peace Index [10] is particularly apt. Peace may sound impossible to evaluate, yet it has been operationalised through eight interconnected pillars: Well-Functioning Government, Sound Business Environment, Equitable Distribution of Resources, Acceptance of the Rights of Others, Good Relations with Neighbours, Free Flow of Information, High Levels of Human Capital, and Low Levels of Corruption. Importantly, none of those pillars *is* peace, just as no single indicator *is* UD. They are conditions that make peace possible and indicators that peace exists. The same logic could apply to UD: What conditions make Universal Design possible, are those same conditions also signals that we are succeeding, and what direct and indirect indicators make it possible to follow the development of UD over time?

A related logic has emerged in human rights monitoring. The OPERA framework [11], applied in Canada's national monitoring of the right to housing for people with disabilities, triangulates outcomes, policy efforts and resources into an overall assessment [12], yet its primary aim remains compliance with obligations.

The rows "Output Evaluation" and "Societal Impact Evaluation" are grounded in current evaluation theory [13], where output and impact are two central terms. The "Process Evaluation" row was added since it is central to UD [4], to developmental evaluation [14], to utilisation-focused evaluation [15], and, more generally, to everything that concerns access and use. The three rows should be understood as separate and not following a particular sequence or degrees on a scale.

2.2. The Pilot – Testing the AUUD Matrix

The matrix has been tested in two pilot workshops, in Lund and Malmö, with municipal staff and researchers. 15 participants in total worked in groups of two or three for about an hour and filled in seven of the nine cells in the matrix. The matrix was available as printed physical copies and as digital tables. The two cells Accessibility × Output and Usability × Output were intentionally omitted, since there is already a large body of work covering these cells, and because they usually dominate what gets measured.

2.3. Results from the Pilots

The results from the pilots were translated, collated from physical and digital forms and clustered by matrix cell (Table 2). Two patterns emerged: *direct* indicators, e.g. “increased use of space by more people”, and *proxy* indicators, things UD contributes to without being a measure of UD itself, e.g. “democratic participation”. The participants also proposed society-level proxies, such as the experience of the city, and social sustainability accounting.

Familiarity with the cells decreases from top-left to bottom-right. We are most accustomed to evaluating accessibility, then usability, then UD for the columns, and output, then process, then societal impact for the rows. Thus, we expected the least familiar cells to yield the fewest contributions. However, this is not what happened. The Universal Design column, and especially its least familiar cell, UD × Societal Impact, instead produced the most conceptually interesting material.

Several features of the responses in the Universal Design column stood out. In the societal-impact cell, participants asked several questions, such as “what is a city for all?”, “what does UD contribute to?”, and “where do we actually know anything?”. They also named existing data sources that could be repurposed, such as national statistics, social-sustainability accounts, and a national public-health survey, and they named concepts to follow up, such as spatial capital and the three pillars of sustainable development. Another prominent thread was connected to agency, participation and democracy.

Interestingly, the bottom row split between a deficit perspective and a positive one. The accessibility and usability cells listed negative factors, the costs of absence or failure, such as unemployment, isolation, home care and ill-health. In contrast, the UD cell listed *presence and belonging indicators* such as trust and democratic participation. In other words, the same row was read through opposite perspectives. Also, a prominent whole-journey theme ran throughout several cells, where evaluation should follow the chain between destinations, the whole way and not just the lift, rather than staying with separate, isolated features. Throughout, participants seemed to reach for concepts rather than only indicators, drawing on notions such as invisible work, emotional work, spatial capital, and the right to access, with reference to the CRPD.

Table 2. Overview of the results from the pilots.

	Accessibility	Usability	Universal Design
Output	Omitted per workshop instructions	Omitted per workshop instructions	Equity and equality (same solution, distance, time, gradient, conditions) and flexibility (alternatives, multiple ways of use); nonclusive , equitable time cost as invisible/emotional work ; whole-journey/whole-city (time between destinations, connections, not just the lift); social participation and interaction. Stringency from initiation to implementation.
Process	Deviations at building-permit stage; Easily Eliminated Obstacles measurement ; procurement requirements; analysis and systematic follow-up across all stages; accessible material ; dialogues .	Quantitative and qualitative studies ; noise/wind measurement ; procurement ; systematic follow-up/control ; how many user groups were heard ; dialogue and consultation-phase area walks ; target-group selection and involvement ; "the democratic mission "; user participation .	A mandatory UD process with named ownership and user testing ; documented systematic routines ; "relevant scales, the whole way , not just the lift"; Democracy and participation; UD as a principled temperament , organic not mechanical; procurement ; citizen dialogue ; many persons in collaboration ; embedding the UD mindset across the full span.
Societal impact	Measure unemployment, housing-adaptation grants, home care, mobility service; consequence analysis of inaccessibility (ill-health, isolation, compensatory-service costs); experienced participation ; increased mobility for more; access and rights ("access to", CRPD); an Accessibility Index .	Increased ill-health, isolation, need for support services; Societal cost of poor usability ; increased use of spaces by more people ; experience of the city; perceived participation ; a Usability Index .	Full participation ; questions ("what is a city for all?", "what does UD contribute to?" "where do we actually know anything?", "how good is this environment?"); flexibility (alternatives and choices); concepts (spatial capital, the three pillars, the digital democratic twin, "for more people, not more stuff"); instruments (citizen surveys, social sustainability accounting, a national public-health survey); socio-economic analysis ; policy development (public-health goals ignore the physical environment); trust in others, agency ; democratic participation , interaction ; organisational development .

3. Concluding Discussion

The pilot showed promising results, where the matrix seemed to open up for productive conversations. The responses from the participants said more about the field than about the matrix itself, and illustrated how much conceptual work remains before the most ambitious (and interesting) cells can be filled with credible indicators.

The paper's main contributions are twofold. *Conceptual*: the field needs a way to follow UD on its own ontological and epistemological terms over time, not in terms of accessibility and usability. *Practical*: the matrix is a thinking tool and conversation starter, not a compliance checklist. By borrowing from peace research, it opens a path toward assessing Universal Design as a societal phenomenon. Reading the study's results against the research questions, the matrix offers one way to evaluate accessibility, usability and UD together (RQ1), while the pilot indicates that UD can be evaluated on its own terms (RQ2), most promisingly at the level of societal impact.

Why did the UD $\times$ Societal Impact cell, the least familiar of all, draw the richest input? One reading is that universal design's distinctive contribution lies in the affective domain of human experience, such as emotions, values, empathy, and belonging, which output and compliance metrics are not built to capture. Read through the Positive Peace logic, the prevalence of proxy indicators in this column is not a weakness of the result but a confirmation of the approach. To evaluate UD's contribution to values that cannot be measured directly, well-chosen proxies are the method. Indicators such as trust in others, democratic participation, and a sense of belonging let UD, as a value-based approach to design [1], be followed over time.

Evaluation in many fields has historically privileged the cognitive and the measurable over the affective, e.g., in education [16]. A parallel inertia has been described in adjacent fields such as educational technology, where a "techno-centric habitus" lets performance measures crowd out transformative aims [17], even as values-driven conversations reveal a desire to treat education and design as means of human transformation rather than standardised commodities [18].

To evaluate UD on its own terms, then, is also to ask how design cultivates empathy, fosters belonging, and supports democratic participation, a shift from a compliance mindset toward one of co-creation [19]. However, this is a limited, early study, which warrants no grand claims. Thus, more research is needed, and the most ambitious cells (UD $\times$ Societal Impact in particular) cannot yet be filled with credible indicators, even though the pilot tests gave valuable input. In further work, it will also be interesting to continue exploring what findings are about the field and what are about the matrix itself. What the study offers at this point is a structure within which that work can now happen.

In conclusion, the study highlighted the importance of what we commit to benchmarking over time. Whatever a field chooses to benchmark, it commits to tracking over decades. Benchmark ramp gradients and you accumulate ramp data. Reorient to societal impact and you start accumulating data about citizenship and equity. We argue that harnessing UD's transformative potential requires shifting from a narrow focus on the output evaluation of accessibility and usability toward a broader, more meaningful assessment of societal impact. A field that refuses to benchmark what it values ends up valuing what it can benchmark. This paper and the Benchmarking Matrix are an attempt to contribute to the UD evaluation discussion by unburdening UD from compliance and instead following UD on its own terms, focusing on the profound question: "What kind of society are we building?"

Acknowledgement

This work was financially supported by the Swedish Governmental Agency for Innovation Systems (VINNOVA, Grant 2021-02810).

References

- [1] Hedvall P-O, Ståhl A, Iwarsson S. Accessibility, usability and universal design – still confusing? Harmonisation of key concepts describing person-environment interaction to create conditions for participation. *Disability and Rehabilitation*. 2025;47(25):6750–6759; doi:10.1080/09638288.2025.2491831.
- [2] European Union. Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services (Text with EEA relevance) [Internet]. OJ L Apr 17, 2019. Available from: <http://data.europa.eu/eli/dir/2019/882/oj>.
- [3] Definition and Overview of Universal Design (UD) [Internet]. Centre for Excellence in Universal Design. [cited 2026 June 22]. From: <https://universaldesign.ie/about-universal-design/definition-and-overview>.
- [4] Steinfeld E, Maisel JL. *Universal design: creating inclusive environments*. Hoboken, New Jersey: John Wiley & Sons Inc.; 2012.
- [5] Persson H, Åhman H, Yngling AA, et al. Universal design, inclusive design, accessible design, design for all: different concepts—one goal? On the concept of accessibility—historical, methodological and philosophical aspects. *Universal Access in the Information Society*. 2015;14(4):505–526; doi:10.1007/s10209-014-0358-z.
- [6] Imrie R. Universalism, universal design and equitable access to the built environment. *Disability and Rehabilitation*. 2012;34(10):873–882; doi:10.3109/09638288.2011.624250.
- [7] Hamraie A. *Building Access: Universal Design and the Politics of Disability*. University of Minnesota Press; 2017.
- [8] Connell BR, Jones M, Mace R, et al. *The Principles of Universal Design* [Internet]. 1997. Available from: <https://design.ncsu.edu/research/center-for-universal-design/>.
- [9] Imrie R, Luck R. Designing inclusive environments: rehabilitating the body and the relevance of universal design. *Disability and Rehabilitation*. 2014;36(16):1315–1319; doi:10.3109/09638288.2014.936191.
- [10] Institute for Economics & Peace. *Positive Peace Report 2024: Analysing the factors that build, predict and sustain peace* [Internet]. Sydney: Institute for Economics & Peace; 2024. Available from: <https://www.economicsandpeace.org/wp-content/uploads/2024/04/PPR-2024-web.pdf>.
- [11] Corkery A, Way S, Wisniewski Otero V. *The OPERA framework: assessing compliance with the obligation to fulfill economic, social and cultural rights*. Brooklyn, NY: Center for Economic and Social Rights. 2012.
- [12] Canadian Human Rights Commission. *Monitoring the right to housing for people with disabilities* [Internet]. 2024 [cited 2026 June 22]. Available from: <https://www.chrc-ccdp.gc.ca/resources/publications/monitoring-the-right-housing-people-disabilities>.
- [13] OECD. *Glossary of Key Terms in Evaluation and Results-based Management in Sustainable Development (Second Edition)* [Internet]. OECD; 2023; doi:10.1787/632da462-en-fr-es.
- [14] Patton MQ. *Developmental evaluation: applying complexity concepts to enhance innovation and use*. New York, NY London: Guilford Press; 2011.
- [15] Patton MQ, Campbell-Patton CE. *Utilization-focused evaluation*. 5th edition. Los Angeles: SAGE; 2021.
- [16] Scharmann LC, Grauer BL. Critical relationships in managing students' emotional responses to science (and evolution) instruction. *Evo Edu Outreach*. 2020;13(1):13; doi:10.1186/s12052-020-00128-6.
- [17] McNutt L. Educational technologists, universal design and transforming higher education. In: Chew YW, Chan KM, Alphonso A, editors. *Personalised Learning Diverse Goals One Heart ASCILITE 2019 Singapore*. 2019. p. 509–513; doi:10.14742/apubs.2019.321.
- [18] Vance US. *Conclusive Enfranchisement. Beyond Health Capacity*. Routledge; 2025; doi:10.4324/9781003535768-11.
- [19] Craddock G, Doran C, McNutt L, et al., editors. *Transforming our World Through Design, Diversity and Education - Proceedings of Universal Design and Higher Education in Transformation Congress 2018* [Internet]. 2018; ISBN 978-1-61499-923-2.