Transforming our World through Universal Design for Human Development
I. Garofolo et al. (Eds.)

doi:10.3233/SHTI220823

Making Research More Inclusive: Is Universal Design of Research the Answer?

Gerd BERGET[a,1] and Birgit KVIKNE[a]
[a] *Dep. of Archivistics, Library and Information Science, Oslo Metropolitan University*

Abstract. Researchers typically declare in the methods section of scientific papers that the study included a representative sample. A closer look at the composition of participant groups, however, reveal that these samples are typically based on criteria such as age, educational background, and a binary gender division. Nothing is mentioned about other characteristics e.g., functional, or neurocognitive variations. Consequently, many "representative samples" do not really represent the whole population, but rather the majority. In this perspective paper, we argue that there is a need for more inclusive research considering the broad diversity of people. We discuss whether Universal Design of Research (UDR) is a purposeful approach. We go through the proposed definition of UDR and apply three topics as examples, namely participant characteristics, communication, and study design. The overall conclusion is that UDR might be purposeful for many studies but is not ethical or purposeful in all types of research. There is, however, a need for a more precise and comprehensive definition of UDR to comply with ethical requirements and to be purposeful for researchers. We therefore conclude by suggesting a revised definition.

Keywords. Universal design of research, user diversity, experimental design

1. Introduction

Empirical research must comply with rigorous requirements related to recruitment, data collection, and data analysis [1]. In many scientific papers, researchers claim to have studied a representative sample of participants. This statement is usually based upon the distribution of characteristics such as age, educational background, and a binary gender division. Characteristics such as cultural background, gender identity, functional diversity, or neurocognitive variation, are generally not mentioned. Neither is the need for adjusting settings in the experimental setup, whether participants rely on assistive technology or whether an interview guide must be revised so all participants comprehend the questions and can communicate with the researchers. Consequently, the so-called "representative sample" is often just representative of the majority population, and a significant portion of people is left out. There might be many reasons for lacking diversity in samples. Some examples are recruitment challenges or bias, experimental design or interview guides that do not consider people with functional or cognitive variations, or lacking awareness in the scientific communities.

In this paper, we will use research fields such as human computer interaction (HCI) and interactive information retrieval (IIR) as a starting point. A typical experimental

---

[1] Gerd Berget, Department of Archivistics, Library and Information Science, Oslo Metropolitan University, Postboks 4 St. Olavs plass, N-0130 Oslo, Norway; gerd.berget@oslomet.no:

design within HCI or IIR measures variables such as performance or preferences, often by comparing different modalities or system components presented to a selection of users. In other cases, qualitative studies are conducted to explore e.g., how people search for and evaluate information. Sometimes, different cohorts are compared to evaluate the impact of certain characteristics, such as domain knowledge or system experience. In other cases, researchers are studying specific parts of the population or exploring research questions that do not address the whole population, such as the design of an airplane cockpit intended for pilots only. In such cases, limited diversity in the samples may be required to answer the research questions and to reduce the number of variables. Other studies compare people with different abilities [2], for example people who are blind versus people with no visual impairments. In the latter case, the purpose is to examine the impact of impaired vision and investigate how to develop more inclusive systems. Such studies are important to ensure accessible systems for all types of users. This paper will, however, mainly address the "mainstream studies".

This paper is structured as follows: The background introduces "the average user", the capabilities approach, and Universal Design of Research (UDR). This is followed by an examination of the UDR definition through the lens of three key areas that needs attention in context of inclusive research, namely participant characteristics, communication, and study design. This section is followed by a discussion of using a capabilities approach within research and the applicability of UDR. Finally, a revised definition of UDR is suggested.

2. Background

2.1. Is there such a thing as an average user?

As stated in the introduction, most mainstream research does not include a broad diversity of users in their samples. We refer to diversity as "*the existence of variations of different characteristics in a group of people. These characteristics could be everything that makes us unique, such as our cognitive skills and personality traits, along with the things that shape our identity (e.g. ethnicity, age, gender, religion, sexual orientation, cultural background)*" [3]. The question is whether a representative sample is possible to attain. A representative sample is based upon the assumption that there are certain "average users", which can be identified by researchers.

Many research projects study user performance or preferences, often comparing groups or cohorts. A consequence of such comparisons is that each participant is turned into an "average person" who can be compared with other "average persons", where certain specific qualities are regarded as essential. According to Rose and Ogas [4] there is no such thing as an average person. By focusing on the sameness of people, an important perspective is lost. Consequently, Rose and Ogas [4] argue the need for embracing and acknowledging user diversity rather than focusing on stereotypes. This is a key issue within the "*science of the individual*" perspective [4], and has also been addressed by critics of universal design. If we design for everyone, we end up designing for no one [5].

2.2. The capabilities approach

The capabilities approach has been discussed in various fields over the years. For instance, Sen [6] emphasised that traditional economic models have omitted debating which activities people are able to do and what kind of people they can be. Nussbaum [7] discussed how women have been given unequal human capabilities due to political and social circumstances. Nussbaum argues that the capabilities approach represents universalism, and can be related to human rights, political justification and implementation [7].

Capabilities are also relevant in context of impairments. In many settings, “disabled people”, “people with impairments” or similar terms are used when referring to people who deviate from certain norms regarding physical, sensory, or mental abilities. Consequently, a division arises between “people with and without disabilities”. An alternative approach is a capabilities approach, which acknowledges that all people have varying capabilities due to many factors. For example, a person who has dyslexia may read slower and with more errors than others. However, people who are very tired, ill, or multitasking may experience the same reduction of reading speed and decoding errors. It might therefore be purposeful to pay more attention towards different contexts and diverse abilities. A related concept is situated abilities, suggesting that an aim should be to design for situated abilities over a narrow focus on people with impairments [8].

2.3. Universal Design of Research (UDR)

Many definitions of universal design exist. Common for all definitions is the underlying principle that one main solution should be applicable to as many people as possible without adjustments for specific groups. The type of characteristics and the terminology applied vary, but many definitions refer to age, gender identity, sexual orientation, cultural background, functional diversity, and neurocognitive variation [9]. In addition to the definition, a set of principles are also put forward to provide more tangible guidelines on how to make things universally designed [10]. Universal design has also been introduced and applied in specific contexts, such as the Universal Design for Learning (UDL) and Universal Design of Research (UDR).

Williams and Moore [11] discussed the need for a broader representation of users in biomedical research and proposed the concept of UDR, defined as “*the design of research so that all people can be included as potential participants, to the greatest extent possible, without the need for adaptation or specialized design*” [11]. This definition builds upon the original definition of universal design by Mace [12]. Williams and Moore [11] also suggest a set of four “rules” for UDR, such as i) plan research to include all types of people despite disabilities, ii) only apply inclusion criteria when necessary, iii) provide multisensory options in the research design, and iv) consult people if you do not have the required knowledge about a disability.

3. Applying Universal Design of Research (UDR)

UDR may be a useful approach to achieve more inclusive research. There are, however, some potential challenges related to this approach. Rigorous experimental design and data analysis methods do not always make it possible to design research that consider all types of abilities. In other cases, just a few small measures are needed to comply with

UDR. In the following section we address the different parts of the definition of UDR through examples from participants characteristics, communication, and study design.

3.1. All people as potential participants

Participants can be described in many ways, and a broad diversity of attributes are associated with universal design. Examples of characteristics are age, gender identity, sexual orientation, cultural or socio-economic background, education, digital literacy, in addition to functional and neurocognitive variation. Important questions, however, are i) which characteristics comprise relevant variables in a study?, ii) which characteristics should be documented?, and iii) how should such data be collected? In most cases, many of these characteristics are not relevant, and might be uncomfortable for participants to provide. Others may be vital, depending on the research questions and study design. In this section we will apply gender identity as example on how studies can exclude potential participants due to a lacking awareness among researchers.

According to the UDR definition, all people are potential participants. Nevertheless, most commonly, samples are presented in a binary gender division, typically in percentage distribution of females and males. This might be an erroneous way of referring to the participants. Moreover, such an approach may be perceived as highly exclusive, and offensive for certain participants. In all populations, there are people with gender identities that do not fit into these binary categories, for example people who are gender-fluid or non-binary. Another question is which category researchers apply to trans people, who do not identify as the sex they were assigned at birth. Are they included with the gender assigned at birth or the gender they identify as when participating in the study? Documenting gender identity can be stigmatizing and result in erroneous data. It might therefore be purposeful to consider whether data on gender is even relevant for the data analysis. If not, do we need to ask participants about that?

Gender identity might be challenging for researchers to address. For instance, participants may not want to state their gender identity, and the categories provided by the researchers may not fit. Moreover, terminology changes quickly. In anonymous surveys, it might be better to get participant data on gender identity by providing several alternatives rather than using a binary division. A commonly reported problem in surveys is that people who do not identify with a binary division often terminates answering surveys when met with stereotypical categories. In interviews or experiments, researchers might erroneously assume a gender identity based on appearance or name.

How can we then achieve a more inclusive research design in context of gender identity? In surveys, several categories can be provided, including the options to self-describe or "prefer not to answer". It may be challenging to decide how many and which categories that should be included. There is probably no right solution to comply with all identities and an ever-changing terminology. It might therefore be helpful to contact user organizations to get advice on the proper, updated, and respectful terminology. Another issue is how questions of gender identity are formulated. Instead of the traditional "*Gender:*", the question could rather be "*To which gender do you most identify as?*". In personal meetings, the researcher can ask the participant what gender the person identify as and which personal pronoun the researcher should use, rather than making assumptions based on appearance or name. To conclude, it is vital that researchers have all types of identities in mind when recruiting to ensure that all people are regarded as potential participants. The first part of the definition is there both purposeful and necessary.

3.2. To the greatest extent possible

In the field of HCI and IIR, certain user groups, such as people with aphasia, have been heavily overlooked [2]. A characteristic related to aphasia is impaired language skills. This group is therefore often excluded from research, since most study designs depend on the ability to produce and comprehend language [2]. Moderations of rigorous methodological requirements may be necessary, for example by offering response alternatives and suggestions during interviews. Without such adjustments the researcher can miss out on essential information. For example, blank sheets of paper for drawing, communication support in handouts and sufficient time, can accommodate communication with people with aphasia. (See [13] for a more detailed discussion of this methodology.)

It would not be ethical to include people with aphasia in a general study, knowing that there would be challenges in understanding and communicating with the researcher. The same applies for people with dementia and some people with severe cognitive impairments. Some of these users may not have the ability to consent or there might be other reasons why the mainstream experimental design is not sufficient. In such cases, UDR will neither be purposeful nor ethical, because special considerations are required to protect the participant. Consequently, it is important that this is acknowledged within the UDR approach, which is confirmed by the inclusion of "*to the greatest extent possible*" in the definition. This does not mean such participants should not be included in any types of research, but UDR may not be the answer for such users.

3.3. Without the need for adaptation or specialized design

Functional diversity comprises a wide range of senses and abilities, such as vision, hearing, and motor skills. All these characteristics can vary without being related to an illness or impairment and occur on a spectrum. For example, one person can have reduced vision and rely on visual aids such as glasses. In contrast, another person can be legally blind relying on more advanced assistive technology. Nevertheless, functional diversity is rarely mentioned in research papers, unless the study comprises people with functional variations regarded as impairments, e.g., being blind, deaf, or paralysed. An exception might be studies referring to participants wearing glasses, which is sometimes mentioned in for instance eye tracking studies because it may impact the data quality.

According to the definition of UDR, there should be no need for adaptation or specialized design. An important component of study design related to functional variation is assistive technology. Through assistive technology, many people with impairments can perform equally well as others. An important question, however, is where the border is set between common adjustments in the setup and "*adaptations or specialized design*" in the UDR definition. For example, there are many settings included in operating systems, such as font size, contrasts, brightness, sound options, sticky keys, and colour settings. In other cases, the participants may depend on using assistive technology such as speech recognition, speech synthesis or a refreshable braille display. Which of these adjustments or assistive technologies would be regarded as adaptations or specialized design according to the definition, and which, if any, would not?

If one decides that all settings adjustments are regarded as adaptations, a large group of people will be excluded from the UDR definition. One could argue that for a study to comply with UDR, all types of assistive technology should be available in the original design, thus regarding assistive technology as a part of the setup. This goal, however, is not realistic, among others due to funding and expertise. As a minimum requirement, however, the experimental design should allow for some level of user control of basic settings, so the study resembles the participant's regular setup as much as possible. In some cases, the research questions may not allow for such modifications. Researchers should as a minimum ask the participants about their usual setup and the one applied at the test. By including some simple accessibility adjustments as a part of the research design, less people will fall under the category "*adaptations or specialized design*".

Neurocognitive variation is seldomly addressed when reporting on the sample. The exception is when the study includes people who based on a formal diagnosis are not regarded as neurotypical, e.g., people with intellectual impairment, autism spectrum disorder (ASD), dementia, or aphasia. Nevertheless, when including all types of people in a study, neurocognitive variation will naturally occur because abilities are also affected by e.g., situation or context. A person who is nervous or tired may for instance produce more spelling errors than a rested and confident participant, while a person with a headache might perform tasks more slowly than others. This is a good argument for incorporating more cognitive tests in experimental design, such as tests for short term memory, concentration, and language skills where these variables are regarded as relevant, to get an overview of the person's situated abilities. At the same time, researchers should not include more tests than necessary, again keeping in mind the research questions, study design, and considerations of the participants.

Neurodiverse abilities tend to occur on a spectrum. For example, speech, and reading or writing skills may vary in degree rather than being present or absent. Consequently, the questions arise regarding which types of users that can be included in a mainstream study design, and when there is a need for special considerations in the study. A vital point in research ethics is informed consent. The researcher must ensure that all information provided about the study and consent form is understood by all participants. For example, if people with dyslexia are included, it might be necessary to ensure informed consent by confirming that the content in the written recruitment letter is understood.

Informed consent can be ensured by providing information through different channels and formats, such as pre-recorded videos, voice recording of texts and easy-to-read information letters. Accessible and easy to read information can be produced through plain language, highlighting important points, and visual support. Such measures would benefit all participants, which entails that such considerations make the research more accessible for all types of people. There is, however, again the question of where to draw the line for defining such measures as special considerations.

4. Discussion

By limiting research to "representative samples" and "typical users" a valuable perspective on user diversity is lost. Moreover, a faulty impression is given that the persons included in the study are representative for all types of people. The performance of all types of users may vary over time, due to for instance illness, tiredness, fatigue, or temporary injuries. Including all types and levels of abilities in research will therefore benefit all types of users and result in more usable systems. This is not meant as an argument against all studies of specific cohorts, e.g., accessibility studies. In many settings, it is purposeful to apply research on narrower user groups. What we are referring to here, are the more general studies, which may benefit from a broader perspective on user diversity, for instance when exploring search interfaces or other types of products and services applied by a broad diversity of users.

By acknowledging that people are neurodiverse and supporting individuals as they are, researchers can gain important knowledge about user behaviour. This expertise can in turn contribute to the development of more inclusive and user-friendly information retrieval systems. Successful and usable interface design builds upon research and user testing. If such knowledge acquirement is limited to stereotype thinking about users and search behaviour, there will be several shortcomings in the system. This simplified view benefits neither users nor system developers. By establishing data collecting methods that include a wide range of user behaviour, one can design for a broader group of end users. Moreover, information systems that support the information behaviour of neurodiverse people, can benefit all users. Research that builds upon simplifications and stereotypes can be purposeful when creating information systems that works for most people. However, when universal design is the goal, one must aim to design for the edges rather than the average. In this context, an understanding of human complexity is essential. Consequently, this needs to be reflected in all aspects of the research process.

One problem with categorizing people is that we end up defining a person only based on one characteristic [4], for instance that a person is blind, has a learning challenge or a gender-fluid identity. People are complex, and there are many different characteristics that affect user behaviour. A person with autism spectrum disorder can for instance be a teenager, gender-fluid, extremely talented in computer gaming, and speak three different languages. By including this person in the sample of "neurodiverse people", all the other aspects of this person are lost, such as high digital literacy. Participant categories such as gender, sexual orientation, socio-economic status, religion, age, nationality, and ability relate to both sameness and difference. Identity categories consist of various characteristics that overlap and interconnect and do not necessarily function as mutually exclusive entities. The UDR approach can facilitate for intersectionality which can mirror a diversity of users and user behaviour. This intersectionality can in turn provide a truer picture of complex human information behaviour.

The original definition of UDR put forward by Williams and Moore [11] was "*the design of research so that all people can be included as potential participants, to the greatest extent possible, without the need for adaptation or specialized design*". In this paper we have regarded this definition as consisting of three parts, as discussed above. The first part of the definition addresses who should be regarded as participants. This requirement holds as a premise that researchers must be aware of the various characteristics and identities people hold and not make assumptions about stereotypes. Moreover, some characteristics may not be necessary to address if it might result in stigmatization or exclusion of certain users.

The second part of the definition includes an important limitation, emphasising that that it will not be possible to design research studies which include everyone. Including everyone is not possible without violating ethical guidelines. Examples are certain people who have challenges with communication or cannot provide informed consent. In some cases, a few considerations will accommodate user needs, in other situations, people will not understand what they are giving consent to or are participating in, making the research highly unethical. We therefore suggest adding an ethical aspect to the definition.

The third part of the definition is a demand for no adaptations or specialized design. We argue that several small adaptations can be purposeful without affecting the research design or results. Examples are allowing some user control over certain parts of the setup if using a computer or by providing information about the study or asking questions in different modalities to make sure all participants comprehend. We therefore suggest modifying this part of the definition by adding a modification to the demand of no adjustments or specialized design.

We suggest the following revised version of the definition applied by Williams and Moore [11]: *"The design of research so that* ***as many people as ethically justifiable*** *can be included as potential participants, to the greatest extent possible.* ***The research design shall enable and empower people to participate with equal opportunities*** *without the need for* ***extensive*** *adaptations or specialized design"*

5. Conclusion

The Universal Design of Research approach may be useful as a framework to increase the awareness among researchers that a representative sample must represent the whole population. To be applicable as a framework or methodology, however, there is a need for a clear definition of the concept. Our suggestion should not be regarded as a final definition, but rather an attempt to develop the concept further. Moreover, it seems necessary with additional principles, for example inspired by the principles of universal design [10], but revised to fit within a research context. Williams and Moore [11] suggested four rules, which seem purposeful. They are, however, limited to people with disabilities. We therefore suggest as future work to both look into the definition of UDR, but also at a set of principles to guide researchers in making their research more inclusive.

References

[1] Berget G, MacFarlane A. Experimental methods in IIR: The tension between rigour and ethics in studies involving users with dyslexia. In Azzopardi L, Halvey M, editors. Proceedings of the ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR); 2019 March 10-14; Glasgow, Scotland: ACM. p. 93-101.

[2] Berget G, MacFarlane, A. What is known about the impact of impairments on information seeking and searching? Journal of the Association for Information Science and Technology 2020; 71(5): 596-611.

[3] Steinmetz CHD. Criticism of the concepts of diversity and inclusion in Western countries. Advances in Social Sciences Research Journal 2021; 8(9): 116-132.

[4] Rose T, Ogas O. The end of average: How to succeed in a world that values sameness; 2015. New York: HarperOne.

[5] Harper S. Is there design-for-all? Universal Access in the Information Society 2007; 6(1): 111-113.

[6] Sen A. Equality of what?: The Tanner Lecture om human values delivered at Stanford University May 22; 1979. Stanford: Stanford University.

[7] Nussbaum MC. Women and human development: The capabilities approach; 2000. Cambridge, UK: Cambridge University Press.

[8] Saplacan D, Herstad J, Schulz TW. Situated abilities within universal design: A theoretical exploration. International Journal on Advances in Intelligent Systems 2020; 13(34): 278-291.

[9] Steinfeld E, Maisel J.L. Universal design: Creating inclusive environments; 2012. New Jersey: Wiley.

[10] Story, MF, Mueller JL, Mace RL. The universal design file: Designing for people of all ages and abilities; 1998. Raleigh: Center for Universal Design.

[11] Williams AS, Moore SM. Universal design of research: Inclusion of persons with disabilities in mainstream biomedical studies. Science translational medicine 2011; 3(82): 1-10.

[12] Center for Universal Design. About UD. 2008; Available from: https://www.ncsu.edu/ncsu/design/cud/about_ud/about_ud.htm.

[13] Kvikne B, Berget G. "My words were completely gone": A qualitative study of the information seeking behaviour of people with aphasia. Information Research 2022; 27(1): paper 916.