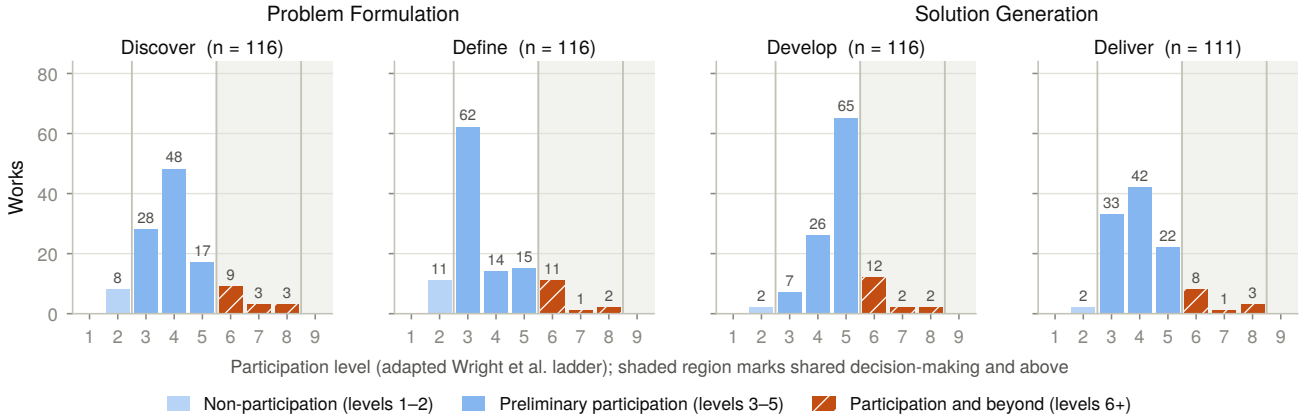
Participation Without Power? A Systematic Review and Call to Center Shared Authority in Accessibility Research

JiWoong (Joon) Jang
jwjang@cmu.edu
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA, USA

Julia Liu
julialiu@cmu.edu
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA, USA

Patrick Carrington
pcarrington@cmu.edu
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA, USA

Andrew Begel
abegel@cmu.edu
Software and Societal Systems Department
Carnegie Mellon University
Pittsburgh, PA, USA



Reported participation of disabled stakeholders in 120 participatory CHI and ASSETS works, 1998–2025

Abstract

Participatory methods in accessibility and assistive technology (AT) research carry a promise: that disabled people will hold real authority over the technologies designed for them. Whether that claim holds in practice at scale in HCI accessibility work has not been examined. To investigate, we systematically reviewed 120 CHI and ASSETS papers (1998–2025) claiming participatory methods of some kind. We coded participation for each work for the four Double Diamond design stages, using Wright, von Unger, and Block's adapted Stufenmodell der Partizipation—a 9-level scale measuring shared authority—to produce a *Participation Profile* for each work. We identify a potentially systemic gap between promise and practice in our corpus: (1) over 80% of works never reach full Participation (level 6-9) at any stage; (2) mean participation does not rise above Consultation (level 4), with works explicitly adopting "participatory design" paradoxically scoring the lowest; and (3) power-sharing peaks during solution generation but is at its lowest during problem framing. Yet, a minority of works sustain consistently higher participation across all stages, showing that substantive power-sharing across the full research process is achievable. Realizing this possibility across the field requires shared commitment: researchers planning participation at each stage of research using the *Participation Profile*, reviewers expecting stage-level consideration of power-sharing from work claiming participatory methods, and the field building shared standards for participatory practice. Collectively, we call on the field to help honor participatory research's promise to disabled participants through these commitments.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
ASSETS '26, Vila Nova de Gaia, Portugal
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2521-0/2026/10
https://doi.org/10.1145/3797867.3829026

ACM Reference Format:
JiWoong (Joon) Jang, Julia Liu, Patrick Carrington, and Andrew Begel. 2026. Participation Without Power? A Systematic Review and Call to Center Shared Authority in Accessibility Research. In *The 28th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '26), October 25–28, 2026, Vila Nova de Gaia, Portugal.* ACM, New York, NY, USA, 21 pages. https://doi.org/10.1145/3797867.3829026

1 Introduction

"*Nothing about us without us*" [11]—names a commitment that accessibility and assistive technology (AT) research has broadly sought to honor through participatory methods [23, 34, 44, 45, 56, 63]. Disabled and allied communities have long held that decisions about disabled lives require the authority of those who live them. This claim centers disabled knowledge and posits that authority over one's own life is not something others should exercise in one's stead. Participatory approaches to designing technologies, represented by a family of methods including participatory design (PD), co-design, and co-creation (and others), offer a vehicle for enacting this principle. Studies across CHI and ASSETS now routinely invoke participatory methods as the expression of this commitment and in doing so, share the promise that disabled people[1] will hold authoritative influence over the technologies designed for them. Yet this promise depends on the practices that enact it, and a growing body of work suggests that these practices are breaking down [18, 67, 68].

Some studies described as using PD involve collaborators as co-authors who shape research questions and hold decision-making authority. Others use the term for a pair of one-hour interviews followed by a researcher-designed prototype and a usability test [55]. Recent work has begun to map this drift. Qi and Yu [55] show that many SIGCHI papers described as participatory design (or methods) involve participants as testers rather than in design decisions. Bødker and Kyng [8] and Huybrechts et al. [35] argue in participatory design venues that contemporary PD practice has drifted from the democratic political commitments that once distinguished it from other forms of user involvement. Parallel observations surface in accessibility-adjacent subdomains at smaller scale: in studies of participatory design with autistic adults [45], gerotechnology [23], and neurodiverse social computing [3]. Together, they suggest that the gap between participatory vocabulary and participatory practice is not a local issue.

The critiques gathered so far have surfaced in pockets, and whether they describe HCI accessibility and AT research as a whole remains an open question. This question matters because a systemic gap between participatory methods' promise and its practice threatens harm to disabled people. Disability research has a documented history of extractive and paternalistic practice in which non-disabled researchers have positioned themselves as translators and interpreters of disabled experience [11, 13, 52], and the accessibility research community has explicitly sought to move past these patterns [27, 44, 69]. In seeking the authority of disabled stakeholders, participatory methods are the methodological commitment the field most directly invokes to enact that move. When the invocation of participatory methods lacks transparent accounting of actual decision-making authority, the participatory vocabulary risks obscuring a reinforcement of the status quo.

Thus we ask: First, *how does the HCI accessibility and assistive technology research community practice participatory methods*, and does the gap between promise and practice hold across that community? Second, if it does, *what should researchers, reviewers, and the field do in response*?

[1]In this work we use identity-first language ("disabled person", "disabled people") rather than person-first constructions ("person with a disability"), reflecting the stated preference of many disabled self-advocates and the convention within much of the Disability Justice tradition [1].

To answer this question, we conducted a systematic review of 120 papers from CHI and ASSETS (1998–2025) that self-describe as using participatory methods — participatory design, co-design, or co-creation — in accessibility or AT research. We characterize each paper along two dimensions to form a *Participation Profile* (Figure 1): a design-stage axis, using the Double Diamond model (Discover, Define, Develop, Deliver) to locate where in the research process participation occurs, and a power-sharing axis, using an adapted nine-level participation ladder (*Stufenmodell der Partizipation*) by participation scholars Wright, von Unger, & Block [34, 70] to measure the degree of recognized decision-making authority disabled participants hold at each stage. A methodological commitment central to this analysis is that we coded only for participation by people with direct lived experience of the conditions the research addresses; proxy stakeholders such as clinicians, caregivers, and family members did not count toward participation. The end result is a set of *Participation Profiles*, which documents the trajectory of power-sharing across the design process for each work in our corpus.

We find the following.

(1) Mean participation across the corpus does not rise above Consultation (level 4 of 9 on Wright et al.'s participation ladder), where participants are asked for input but do not hold decision-making authority. Paradoxically, works self-describing as adopting participatory design (PD), the methodology whose tradition most explicitly commits to redistributing power, score *lower* than those adopting any other participatory label.
(2) Over 80% of papers never reach any shared decision-making (level 6 of 9, where participants first hold decision-making authority) at any stage.
(3) Participation is unevenly distributed across the Double Diamond design stages. Scores are lowest at Define, where researchers frame the problem before consulting participants, and highest at Develop, where co-design workshops and prototyping activities give disabled participants a generative role.
(4) A minority of works, identified through k-means clustering, sustain a consistently higher level of participation across all four stages. Participation in these works does not decline at the Define stage of research as it does across the rest of the corpus, indicating that reduced participation during problem framing likely reflects prevailing practice rather than a structural limitation of participatory methodology.
(5) Finally, we report that declared disabled authorship is strongly associated with both higher and more uniform participation across stages.

Together, these findings strongly suggest that the gap between participatory methods' promise and how they are reported is a systemic issue within our corpus, pointing to a need for clearer methodological accountability. Closing it would let the accessibility research community hold participatory design as a methodological standard that picks out a consistent and transparent set of practices. This has implications for researchers, reviewers, and the field.

(1) For researchers, we propose the *Participation Profile* as a framework to plan inclusion by disabled stakeholders stage

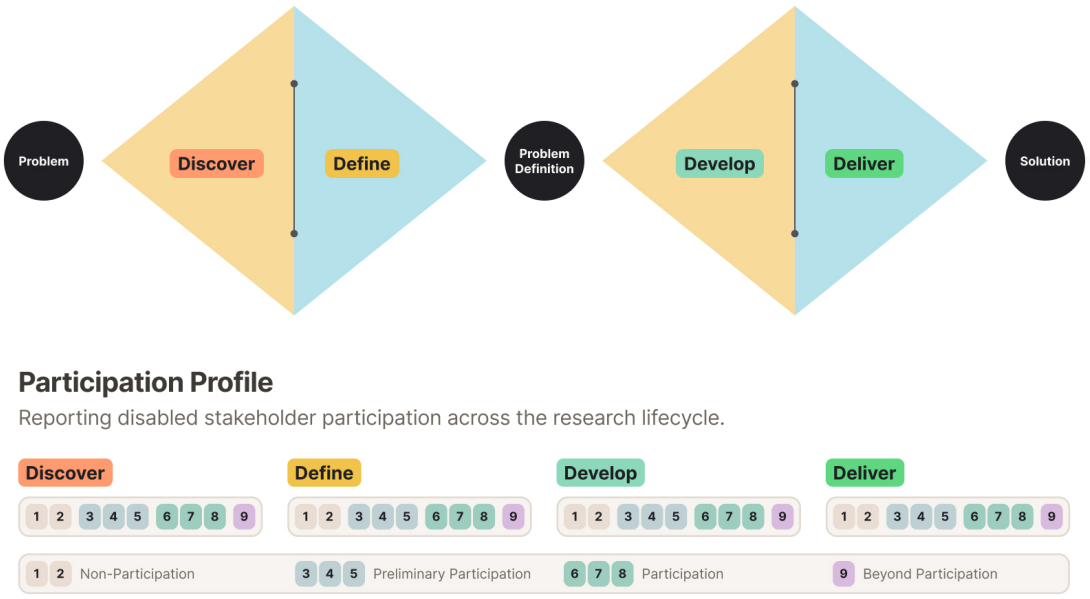


Figure 1: The Participation Profile combines the Double Diamond model of design with the adapted Wright, von Unger, and Block's adapted *Stufenmodell der Partizipation* ladder. Each paper in our corpus receives a rating from 1 (lowest participation) to 9 (highest participation) for each of the four Double Diamond stages, capturing how principals with lived experience participated at that stage.

by stage, at specific levels of decision-making authority, rather than as a general commitment. In the discussion, we pair the framework with exemplars that show what tractable incremental moves between levels look like in practice — what raising participation at a given stage costs and returns.

(2) For reviewers and program committees at AT-related venues, the finding that the gap is field-wide implicates the review process itself. We propose that stage-level participation reporting should be an expected feature of any participatory method claim in accessibility venues. Current reporting conventions do not require authors to specify how participation was distributed across design stages, which leaves reviewers without the information needed to evaluate PD claims against the practices they describe.

(3) For the accessibility research community, we call on the field for shared reflection on participatory methods' commitments to authority-sharing with disabled people. We propose that the *Participation Profile* can serve as the basis for a standard reporting artifact — a data sheet or model card for participation — that authors declare as part of any paper claiming participatory methods. Any such standard must navigate the tension between participatory methods' historical commitment to direct representation and the reality that many disabled people exercise agency through interdependence with others. We explore this tension and its implications for participatory practice in the discussion, and invite the community to shape this work with us.

The remainder of the paper situates these contributions against prior work on participatory design and participation frameworks (§2), describes our corpus and coding procedure (§4), presents the *Participation Profiles* and k-cluster analysis (§5), and develops the implications introduced for researchers, reviewers, and the field (§6).

As researchers, we turn to participatory methods in accessibility to affirm that disabled people should hold real authority over the technologies built around their lived experiences, and to reify the redistributive commitments participatory methods were designed to carry. This work argues that those commitments become real only when authority is visible and accountable at the level of specific decisions, where participants, reviewers, and the field can see it, examine it, and hold it to account.

2 Background and Related Works

This section situates our study along the four lines of prior work it draws on and departs from: the political origins and contemporary definitional drift of participatory design[2] (§2.1); the participation-framework and power-ladder tradition we borrow our measurement instrument from (§2.2); prior reviews of PD within accessibility and

[2]We use PD as the banner term throughout this paper while keeping the adjacent methods like co-design and co-creation in scope where authors invoke them. The three labels co-occur frequently in the accessibility literature, and PD's Scandinavian labor-movement origins give it the most explicit commitment to redistributing authority among the three[8] — a commitment we take as the clearest benchmark against which to assess the family as a whole.

adjacent domains (§2.3); and the design-process frameworks from which we draw our stage scaffold (§2.4).

2.1 A Primer On Participatory Design's Political Origins and Current-Day Drift

Participatory design emerged from the Scandinavian cooperative-design and labor movements of the 1970s and 1980s as an explicitly political methodology. Legislative changes in Denmark, Norway, and Sweden granted workers new rights to influence technology decisions in their workplaces, and researchers collaborated directly with trade unions to develop the theoretical and methodological groundwork for worker participation in sociotechnical systems design [7, 16]. Early projects such as DEMOS and DUE paired researchers with labor unions to build workers' capacity to negotiate with management over technology deployments, and the later UTOPIA project paired researchers with the Nordic Graphic Workers' Union to develop alternative newspaper production technologies that would preserve and extend typographers' craft skills rather than deskill them [16, 21]. The distinguishing feature of this "first generation" of PD was its grounding in what the tradition called the Collective Resources Approach: the explicit premise that design is political work, that technology deployments redistribute power between workers and management, and that workers as a collective have both the right and the capacity to co-determine those deployments [16, 22].

As PD traveled outside Scandinavia through the late 1980s and 1990s — most visibly through the first Participatory Design Conference in Seattle in 1990 and through Muller and colleagues' work on adapting PD for North American contexts [48, 49] — its political core was selectively retained. U.S. and broader HCI adaptations emphasized the cooperative and pragmatic dimensions of PD (better products, higher user satisfaction, more effective knowledge acquisition) while softening its democratic-empowerment commitments, partly because North American workplaces lacked the union infrastructure that had underwritten PD's Scandinavian origins [22, 39, 49]. Participation on North American soil also carried a distinct domestic genealogy: as Kensing and Greenbaum recount in the *Routledge International Handbook of Participatory Design*, it was through grass-roots action on "burning social issues such as civil rights and urban problems" that participatory action research — rather than union-grounded PD — first took serious hold in US research communities [40]. The two genealogies share democratic commitments but institutionalized them differently, an entanglement that resurfaces in the PD/AR boundary questions we address below and in our search-scoping decisions (§4).

By the 2000s and 2010s, PD had spread into healthcare, education, civic technology, and accessibility [26, 55], and in doing so had expanded its stakeholder base well beyond workers to include any group affected by a sociotechnical design choice. This expansion has sharpened a contemporary debate across different research fields over what PD requires on its own terms and where its boundaries sit against adjacent methodologies [8, 18, 35, 55, 67, 68]. Two strands of that debate bear directly on the *Participation Profile* we develop.

2.1.1 What Counts as Participatory? The question is live because PD's use has arguably outpaced its self-accountability: Frauenberger et al. [18] posit that as PD has spread across domains, its methodological commitments and evaluative criteria have not kept pace, leaving the field without shared standards for assessing when a participatory claim is warranted. Wacnik et al.'s [68] recent systematic review documents the resulting variance: papers self-identifying as participatory diverge sharply in who is included, at which points in the design process, through what methods, and with what degree of authority, and these dimensions are reported too inconsistently to permit comparison across studies.

Several authors have attempted to supply that missing standard, each foregrounding PD's political commitments. Bødker & Kyng [8] argue that contemporary PD often focuses on "here-and-now" issues with low technological ambition, neglecting the structural political commitments that distinguished PD from other forms of user involvement, and they call for a re-politicization of PD around issues that matter at scale. Huybrechts et al. [35] articulate "visions that change" as the missing political dimension of current PD practice. Udoewa [67] makes the political lineage most explicit: they distinguish PD that is participatory "to the root or core — full inclusion as equal members of the research and design team" from what they term *colonial participatory design*, and proposes a typology of participation along dimensions of who initiates, who participates and how fully, and who leads. Qi and Yu [55], in a review of 185 PD cases from SIGCHI venues, offer a complementary characterization organized around four constitutive features — cooperation, political commitment (democracy and empowerment), mutual learning, and creativity — and name three recurring misconceptions: overlooking the political commitment to democracy and empowerment, narrowly equating PD with design ideation or prototyping, and introducing unnecessary PD terminologies.

2.1.2 What Differentiates Participatory Design from Other Methodologies? Distinguishing PD from adjacent methodologies in practice is harder than it appears, and this difficulty motivates what our *Participation Profile* is designed to capture. PD sits within a crowded space of similar-sounding approaches — action research, co-design, co-creation, user-centered design — that share vocabulary and activities but carry different underlying commitments.

On action research (AR), Hayes [32] defines AR as a class of collaborative research approaches characterized by longitudinal engagement, an explicit commitment to producing change in the participating community, and shared problem-setting authority. AR projects are oftentimes led by practitioner-researchers with a direct concern regarding the target community, with some AR projects involving "empathetic outsiders" and others being entirely community-owned [62]. However, not all community-owned initiatives are AR [6]. Because PD and AR both involve direct stakeholder engagement and democratic commitments, they frequently exhibit surface-level similarities that invite conflation. The difference is that AR focuses on sustainable processes, organizational interventions, and knowledge generation through action [32, 33], while PD is explicitly oriented towards design. Hayes [32, 33] cautions that the increasing tendency to collapse the two, particularly when AR is applied to HCI, erodes both traditions' distinctive contributions. Grigorovich and colleagues [23] surface the same confusion in gerotechnology, where projects invoking PD are sometimes evaluated against participatory action research's distinct empowerment

commitments and found to fall short of commitments the PD label was taken to imply.

On co-design/creation, Sanders and Stappers [59] describe co-creation and co-design as broad umbrella terms that encompass a range of collective creative practices, with co-design describing "the creativity of designers and people not trained in design working together in the design development process." The authors name PD as an example of co-design, but emphasize that co-design also encompasses a wider range of participatory practices, including methods from consumer product development such as lead-user innovation. Co-design does not directly trace its lineage to labor movements or similar political engagements, nor does the umbrella term carry the same political weight as PD. Nevertheless, the authors emphasize that power-sharing is structural to co-design: some level of control must be "relinquished" by trained designers and "given to potential customers, consumers, or end-users." In HCI research, the PD and co-design/creation labels are frequently used interchangeably in published work, which leaves readers unable to tell whether a given paper is claiming the explicit political content PD carries or a different set of commitments that the umbrella also covers.

The consequence of this porousness is subtle on its face but damaging. A paper's chosen label — "participatory design," "co-design," etc — is read as a claim about what authority participants held, but the labels have drifted far enough in use [55, 68] that readers cannot reliably recover that from the methodological label alone. We express concern that this condition also creates the potential that this confusion could extend to authors, researchers, and practitioners, where a label choice is driven by subfield convention or rhetorical fit rather than a principled distinction. We do not attempt to adjudicate the full definitional debate. The critiques surveyed above converge on one feature as PD's distinctive content: the commitment to democratic decision-making. Two lines of thinking in accessibility research sharpen why that commitment, in particular, is the feature whose distribution across our field most needs to be characterized.

The first is an epistemic argument advanced by disability studies and critical-disability HCI scholarship. Disabled people hold situated knowledge about their own lives that non-disabled researchers do not and cannot fully access [11, 44], and design decisions made without that knowledge reproduce the ableist defaults that accessibility research is meant to disrupt [27, 69]. Recognizing this epistemic position requires more than inviting disabled people into research as informants; it requires that they hold legitimate authority over decisions their knowledge bears on [5, 13]. In a field with a documented history of extractive and paternalistic practice, democratic decision-making is the feature of PD that operationalizes this epistemic position in research practice [52].

The second is the observation that design authority is not a constant within the lifecycle of a project. Framing the problem, scoping the inquiry, generating candidate solutions, and selecting what proceeds to evaluation are distinct decision-sites, and authority at one does not imply authority at another. Qi and Yu's diagnosis that contemporary PD tends to narrow toward ideation and prototyping [55] suggests precisely this asymmetry: authority at the generative stages can coexist with its absence at the stages where a project is framed and where its outputs are selected. Characterizing democratic decision-making in the accessibility literature therefore requires an instrument that resolves authority at the level of individual design stages rather than treating it as a single per-paper property.

2.2 Participation Frameworks and Power Ladders

The question of *how to measure participation* has a long history outside HCI. Arnstein's foundational "Ladder of Citizen Participation" [2] distinguished eight rungs of community involvement in urban planning, ranging from manipulation (citizens are informed of decisions already made) through tokenism (citizens are consulted without binding influence) to degrees of citizen power (partnership, delegated power, and citizen control). Arnstein's Ladder shows that participation is not binary but ordinal, and that the difference between adjacent rungs can be politically decisive. It has seeded decades of adaptations across public health, development studies, environmental planning, and education [31, 66, 70].

Within public-health participatory research, Wright, von Unger, and Block's nine-level model [70] refines Arnstein's scheme for community-based health research, distinguishing *non-participation* (instrumentalization, educating and treating), *preliminary participation* (information, consultation, inclusion), *participation* proper (co-determination, partial decision-making authority, decision-making power), and *beyond-participation* (autonomous organization). The central distinction between preliminary participation and participation proper is whether the target group holds binding decision-making authority — the same line Qi and Yu [55] identify as the definitional boundary between PD and adjacent methodologies. Heitplatz et al. [34] have recently demonstrated that Wright et al.'s model can be productively applied to participatory technology development with disabled people, building practice guidelines from six in-depth AT projects.

What remains unaddressed is a *field-wide, stage-sensitive* application of this kind of instrument to the accessibility and AT literature: one that characterizes participation not as a single per-paper label, but as a profile across the discrete stages of the design process each paper engages. Qi and Yu's review [55] provides a venue-wide feature-level characterization but does not resolve participation at the stage level; Heitplatz et al. [34] resolve participation at the project level but across only six cases. Halskov and Hansen [24] document heterogeneity across a decade of PD conference proceedings but without a power axis. The gap this paper addresses sits at the intersection of these prior efforts: a corpus-scale application of an ordinal power-sharing scale, resolved to the level of the design-process stage.

Adjacent critical frames in HCI foreground power in complementary ways that our instrument is in dialogue with but does not subsume. Design Justice [12, 13], crip technoscience [27], interdependence-as-frame [5], and counterventionist HCI [69] each articulate normative commitments about how design ought to distribute authority; disability-studies and intersectional critiques of accessibility research [14, 28, 29, 44] extend the concern to *whose* participation is being solicited in the first place. These frames inform the political stakes of our analysis and several of our coding decisions (see §4), but they do not themselves supply a stage-level measurement

Table 1: The adapted Wright, von Unger, and Block Stufenmodell der Partizipation [70] used in this review: nine ordinal levels of participation grouped into four meta-categories. The point between levels 5 and 6 marks the threshold at which decision-making authority begins to be shared.

Meta-category	Level	Label	Description
Non-participation	1	**Instrumentalization**	Participants are used to legitimize a predetermined agenda; their presence confers validity without influencing any decision.
	2	**Instruction**	One-way directive communication: participants are told what to do, think, or accept, with no channel for response.
Preliminary stages of participation	3	**Information**	Participants are informed about decisions, procedures, or findings but have no mechanism to shape them.
	4	**Consultation**	Participants' views are solicited (e.g., via interviews or surveys), but researchers retain full discretion over whether and how input is used.
	5	**Inclusion**	Participants are systematically involved in the process and their perspectives are weighed, though final authority still rests with researchers.
Participation	6	**Shared decision-making**	Participants and researchers negotiate decisions jointly, with equivalent standing in the outcome.
	7	**Partial delegation of decision-making authority**	Participants hold decisive authority within a bounded scope (e.g., feature selection, study-design elements), and those decisions stand.
	8	**Decision-making authority**	Participants hold authority across the full scope of the activity; researchers act in a supporting or facilitative role.
Beyond participation	9	**Community-owned initiatives**	The activity is initiated, led, and governed by the community itself; researchers, where present, participate on the community's terms.

instrument, which is what the Wright et al. × Double Diamond pairing developed in this paper contributes.

2.3 Prior Reviews of Participatory Methods in Accessibility and AT Research

Prior reviews of PD in AT-adjacent domains each document a version of the definitional-drift problem within their subdomain. Quintero's review of PD in accessible technology [56] identifies the tendency to apply the PD label to processes more accurately described as user testing. Maye and colleagues [45] document this pattern specifically in PD with autistic adults, showing that many studies involving autistic participants never grant them authority beyond feedback provision. Grigorovich and colleagues [23] document the same pattern in gerotechnology and explicitly frame it as a PD–participatory-action-research tension, arguing that PD's empowerment commitments are frequently weaker than the older adults involved would warrant. Baillargeon and colleagues [3] take up the same pattern in neurodiverse social computing research, with attention to how the "participatory" label obscures which neurodivergent voices are included and which are not. Each of these reviews establishes the definitional-drift pattern within a bounded subdomain — whether through conceptual synthesis [23, 56], small-*n* scoping of a single population [45], or decade-scale coding of a single application area [3].

The preponderance of evidence raises the question of whether the field at large encounters similar issues, and to what depth and scale. The closest example on depth is from Heitplatz et al. [34], who apply the Stufenmodell der Partizipation ladder to six AT projects and build practice guidelines from their findings. The closest precedent on scale is Qi and Yu's review [55], whose SIGCHI-wide review characterizes 185 papers across 26 venues over ten years at the feature level (cooperation, political commitment, mutual learning, creativity), of which healthcare is the largest single application area but accessibility is specifically addressed.

The contribution of this paper is to combine the two. Where Heitplatz and colleagues ask *what does high-quality participatory AT development look like in practice?*, and Qi and Yu ask *what constitutive features of PD are present in SIGCHI at large?*, we ask: *what is the distribution of participatory engagement across the accessibility and AT literature, and where do the gaps concentrate?* Operationalizing the democracy prong of Qi and Yu's political-commitment dimension at stage resolution lets us identify systemic patterns and propose field-wide solutions.

2.4 Design-Process Frameworks as Analytical Scaffolds

Characterizing participation at the level of the design-process stage requires a framework for segmenting the design process. Several candidates exist in HCI and design scholarship, each foregrounding different properties. Linear waterfall models separate requirements, design, implementation, and testing as discrete phases [57], but are poorly suited to PD analysis because they assume a one-directional flow that most PD projects (and most AT research) do not follow. Iterative and spiral models [9] capture the cyclical character of design but at the cost of clearly identifiable stages, making cross-paper comparison difficult. Design-thinking frameworks such as the Stanford d.school's five-stage model (empathize, define, ideate, prototype, test) [54] offer clearly delineated stages but were developed for design education rather than for analytic use, and their stage definitions map awkwardly onto the dominant patterns of AT research

reporting. Sanders and Stappers' [59] fuzzy-front-end / development / production distinction captures the depth-of-participation question but is too coarse-grained for stage-level coding.

The Double Diamond framework, articulated by the UK Design Council [15], divides the design process into two successive diamonds — a *problem diamond* (Discover, Define) and a *solution diamond* (Develop, Deliver) — each consisting of a divergent and a convergent stage. Discover involves open exploration of the problem space (formative interviews, ethnography, literature review); Define synthesizes that exploration into a bounded problem framing (themes, requirements, design goals); Develop generates and iterates candidate solutions (brainstorming, prototyping, co-design workshops); and Deliver refines and evaluates a converged solution (usability testing, deployment, evaluation). Three properties recommend the Double Diamond for our analysis. First, its stages are clearly delineated and internally consistent in the sense that divergent and convergent stages alternate, which makes it possible to segment reported research activities systematically. Second, its stages correspond well to the activities AT research papers typically report — formative interviews, requirement synthesis, co-design workshops, and evaluation studies — even when papers do not explicitly invoke the framework. Third, the problem–solution distinction it encodes is substantively relevant to our research question: the problem diamond is where consequential framing decisions are made, and the solution diamond is where implementation and evaluation decisions are made. One of the central empirical claims of this paper is that participation is distributed unevenly across these two diamonds.

We treat the Double Diamond as an analytical scaffold rather than a normative commitment. We do not claim that AT research should follow Double Diamond structure, nor do we contribute to design-process-framework scholarship. We use it because it segments the design process into stages that map onto reported activities across the corpus and that correspond to substantively different kinds of design decisions. Several boundary cases require explicit handling: works that engage only one diamond (e.g., publications describing early stage exploratory work that discovers and defines without developing or delivering), papers with nested or iterated diamonds (e.g., iterative prototyping cycles that each contain their own discover–define–develop–deliver loop), and papers that report activities in an order that does not match the framework's canonical sequence. We describe our treatment of these cases in (§4). We note one structural boundary of the framework — it ends at Deliver, leaving post-project sustaining labor out of frame — and discuss its consequences for our analysis in §7.

3 On Proxies and Interdependent Representation

Before describing our methodology, we take up a question that shaped it and that we believe the field must confront on its own terms: how should participation by *proxies* — caregivers, family members, educators, clinicians, or others who participate in interdependence with a disabled principal — be counted when work claims a participatory method? For this review, we make a deliberate decision to code only for direct participation by people with lived experience of disability, and correspondingly did not rate works which involved only proxy participants. As we detail in §4, this choice was made partly out of practicality: proxy treatment was the largest single source of disagreement in our initial coding. But the deeper reason we treat proxies as out of scope is that the political status of proxy participation in disability research is a conundrum we do not believe can be resolved inside a coding protocol.

Excluding proxies from our coding is not a claim that proxy involvement is unimportant or that it should never occur in participatory work. We note that some PD-based traditions privilege direct participation as a baseline commitment: Arnstein's Ladder itself grants higher ratings only to direct representation by members of the target group [2], and Hamidi et al. echo this within PD and assistive technology, cautioning that proxy involvement "*should only be used when working directly is impossible*" [26]. But this position raises two thorny questions we flag to the reader.

3.1 Who Decides When Direct Participation Is "Impossible" in Participatory Design?

The "only when direct participation is impossible" principle depends on *who* is making the determination, and that someone is, in practice, almost always the researcher. This position is not benign. It runs against the documented tendency of people in power to systematically underestimate disabled people's capabilities [11, 44], and it generally means an easier recruitment path for researchers who have institutional incentive find such impossibility. Without explicit accountability for who made the call, on what evidence, and in consultation with whom, "impossibility" becomes a discretionary category that can license whatever substitution the researcher finds convenient.

3.2 What About Chosen Interdependence in Participatory Research?

A stronger challenge to direct participation as telos comes from disability studies. Scholars there argue that interdependence — reliance on others as an irreducible part of how many disabled people exercise self-determination — is not a diminished form of agency, and proxy involvement by a trusted ally or caregiver does not automatically constitute diminished participation [5, 36]. On this view, insistence on direct participation as the singular, ultimate goal imports an ableist assumption about independence that we should treat with caution [27].

But this rebuttal is itself complicated. Recent reviews by Yildiz & Gerling [71] and by Jang et al. [37] find that interdependence in disability research is frequently *not* chosen by the disabled person — it is configured by researchers or by surrounding structures. Jang et al. sharpen the distinction: the politically significant question is not whether a disabled person is directly present, but whether the disabled party sets the terms of the arrangement. The same proxy configuration — a caregiver speaking on behalf of a non-speaking person — can constitute "recognized interdependence," chosen and bounded by the principal, or *conditional interdependence*, imposed without the principal's input. A proxy arrangement the principal did not choose is straightforwardly an inferior condition to direct participation on any participatory account. The hard case is *recognized interdependence*: a proxy arrangement the disabled principal actively chose, on terms they set, does not obviously rank below

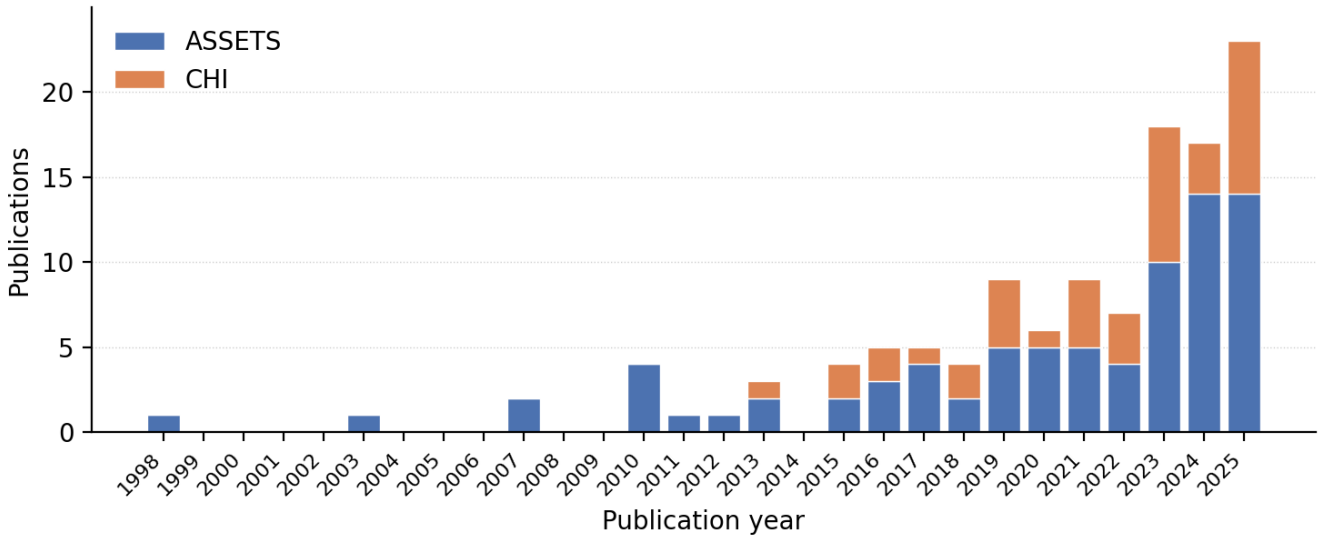


Figure 2: Counts of the 120 included papers across publication years from 1998 to 2025, with five papers published before 2010 and the remainder distributed across the subsequent fifteen years. The last three years together account for 59 papers, or approximately 45% of the corpus.

direct participation — it may in fact be the form of participation that best reflects their agency. Jang et al.'s concept of *relational sovereignty* supplies a framework for treating recognized interdependence and direct participation as normatively comparable on these grounds: what matters is whether the principal sets the terms, not whether they are directly present. But the framework depends on the principal's agency to set and voice those terms, and this capacity is exactly what is contested in populations that feature prominently in accessibility research — participants with cognitive disabilities, or minors whose agency is legally mediated by guardians [25, 45]. In these cases, relational sovereignty reframes the question rather than answering it.

A further complication is that the proxy and principal categories themselves can blur. Proxies may hold overlapping lived experience with the principle: a parent or teacher answering for a child with a visual disability may themselves have grown up with one, and disability's heterogeneity means a proxy can simultaneously be a member of an adjacent — or the same — disability community. In such cases, their contributions are not reducible to merely secondhand testimony; they carry direct experiential knowledge alongside the proxy role. Published reports rarely make this intersectional understanding visible, and a coding protocol reliant on the published record, such as ours, therefore cannot credit it. The inverse configuration is just as opaque: disabled people themselves frequently serve as proxies — for family members, students, or clients — within and beyond accessibility research, and the field currently has little accounting of how often this occurs or how it should figure in a participatory claim.

We do not attempt to resolve these questions. The narrower claim our data does support is that reporting conventions in accessibility and AT research routinely collapse the distinction between principals and proxies, and between conditional and recognized interdependence, treating all of them as interchangeable substrates for a participatory design claim. Whatever the accepted normative position turns out to be, the field cannot currently adjudicate it from its own published record, because that record does not consistently distinguish the cases.

This has a concrete implication for §6.3. Any prescriptive framework the field adopts for participatory methods — including the *Participation Profile* we develop in this paper — must not treat proxies as a category to be abstracted away or scored out of the picture. The field's current difficulty deciding what proxy involvement means for a participatory claim is a reason to demand more transparent reporting of proxy arrangements, not less. We therefore do not recommend that other researchers adopt our principals-only rule as a methodological standard for their own work; that rule was fit for our analytical purpose (characterizing direct participation by disabled principals at corpus scale) and should not be generalized to prescriptive use. What we do recommend is that any participatory-methods claim in accessibility venues declare, at minimum, which stakeholder populations participated directly at which stages, which were represented by proxies at which stages, and on what basis those proxy arrangements were configured. Reporting at this level of detail does not resolve the normative debate; it gives the field the record it would need to have that debate in the first place.

Positionality and Reflexivity

This work is conducted by multiply disabled, neurodivergent, and allied researchers. Many of us hold dual identities as both disabled individuals and academic researchers, which influence our perspectives on issues of power-sharing in accessibility research and research with disabled communities.

Table 2: Search terms for participatory methodologies and accessibility/disability topics, and full-text search strings used for retrieval of publications from the ACM Digital Library. For ASSETS, a publication was included if any participatory term matched the title or abstract. For CHI, a publication was included if any participatory term matched the title or abstract *and* any accessibility term matched the title or abstract.

Concept	Search Terms
Participatory methodology	`"participatory"` OR `"codesign"` OR `"co-design"` OR `"co-creation"`
Accessibility / Disability	`"accessibility"` OR `"disability"` OR `"disabled"` OR `"assistive technology"`

Executed Venue Query Strings

ASSETS Query:

```
     (Title:("participatory" or "codesign" or "co-design" or "co-creation")
  or Abstract:("participatory" or "codesign" or "co-design" or "co-creation"))
```

CHI Query:

```
     (Title:("participatory" or "codesign" or "co-design" or "co-creation")
  or Abstract:("participatory" or "codesign" or "co-design" or "co-creation"))
  and (Title:("accessibility" or "disability" or "disabled" or "assistive technology")
  or Abstract:("accessibility" or "disability" or "disabled" or "assistive technology")))
```

Our interest in this topic grew out of anecdotal observations across our work and personal experiences in prior projects. In some cases, we have observed gaps between the way participatory methodologies are described in theory and how those methodologies are implemented in practice. While we have also encountered many examples of strong participatory work that lives up to stated principles, observing this gap motivated us to pursue this research for two major reasons. First, many of us have personal stakes in the equitable treatment of disabled people. While the misalignment is unideal at an anecdotal level, we were concerned that our observations may be reflective of a larger field-wide pattern that consistently undercuts disabled authority. Secondly, as academics working in accessibility and HCI, we have a professional stake in the ideas behind participatory methods being reflected in how the term "participatory" is actually applied. A persistent gap between label and practice makes it harder for the field to communicate, collaborate, and build cumulative knowledge.

Our positionality is not neutral: we recognize that as academic researchers, we hold institutional power and privilege that shape how we understand and evaluate dynamics between researchers and disabled stakeholders. At the same time, we believe our diverse team composition strengthens our analysis. When coding papers for their participatory practices, team members with relevant lived experience were often more carefully attuned to how disabled participants' contributions were characterized in a paper's language, helping to surface nuances in how contributions were framed that informed our coding decisions. Our backgrounds also shaped specific design decisions. We deliberated over the status of proxies as long as we did partly because some of us live the interdependence at issue (§3), and we chose to credit declared disabled authorship without penalizing silence (§6.2) knowing first-hand what disclosure can cost.

4 Methodology

In our methodological process, we attempt to answer how much decision-making authority disabled participants are reported to hold during different stages of the design process, and do so by evaluating every included work against a defined rubric [20, 50]. Our systematic review occurs in five steps, summarized in Table 4. We first constructed a corpus of candidate publications of HCI accessibility and assistive technology research by searching titles and abstracts in the ACM Digital Library for participatory-methodology terms, paired with accessibility terms. We then manually screened the retrieved records and applied inclusion and exclusion criteria, arriving at 120 publications reporting participatory methods. We then read and coded the 120 works remaining in the analytical corpus on Wright et al.'s nine-level Stufenmodell der Partizipation ladder (Table 1) at each stage of the Double Diamond in the work, leading to its *Participation Profile*; works whose participation ran exclusively through proxy stakeholders were flagged but not scored. Finally, we validated the coding schema through an interrater-reliability process on a 30-paper sample.

Table 3: Corpus counts across retrieval, screening, and coding. Counts reflect the number of publications retained after each filtering step.

Stage	ASSETS	CHI	Total
Retrieved via keyword search	103	56	159
Retained after in/exclusion screening	80	40	120
Coded on all Double Diamond stages	74	34	108

4.1 Constructing the Corpus

To prepare for analysis, we first constructed a corpus of publications from two ACM venues: the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI) and the ACM SIGACCESS Conference on Computers and Accessibility (ASSETS). Both are premier venues for accessibility research within HCI. We selected these venues because HCI as a field routinely engages adjacent scholarly traditions — including design research, disability studies,

Table 4: Our five-step corpus review process, including initial retrieval, manual screening, inclusion/exclusion decisions, analytical coding, and proxy filtering. For initial retrieval, manual screening, and inclusion/exclusion decisions, we reviewed paper titles and abstracts. For analytical coding and proxy filtering, we utilized paper full-texts.

Step	Material Reviewed	Description
1. Initial Retrieval	Title & Abstract	Retrieve works matching keywords for methodology and disability/accessibility using search strings shown in Table 3.
2. Manual Screening	Title & Abstract	Exclude semantic mismatches (e.g., papers using "accessibility" to mean public file access).
3. Inclusion / Exclusion	Title & Abstract (Full-Text Fallback)	Apply inclusion/exclusion criteria (include only works conducting an active study with disabled primary end-users).
4. Analytical Coding	Full-Text	Apply the 9-level participation scale across all 4 Double Diamond stages.
5. Proxy Filtering	Full-Text	Remove scores from publications involving exclusively proxy stakeholders.

and disability rights — with which participatory methodology is closely associated. Authors publishing in CHI and ASSETS are therefore likely to encounter participatory commitments as part of a shared methodological vocabulary, making these venues a well-suited site in which to examine how participatory design is actually practiced in accessibility research.

Publications from these venues were initially retrieved on February 3rd, 2026, and re-retrieved on April 21st, 2026, from the ACM Digital Library. No date filter was applied to the initial retrieval; the search resulted in from 1998 through the end of 2025. The April re-retrieval was restricted to publications dated 2025 or earlier for consistency.

We developed two sets of search terms: one targeting participatory methodologies, and one targeting disability, accessibility, and assistive technology. We bounded our methodological search terms around the vocabulary of the PD and co-design lineages in HCI, as we were primarily interested in participatory and democratically-oriented traditions pertaining to *design practice*. As such, we adopted search terms such as "participatory", "co-design", and "co-creation" to capture the terminology used within these traditions. While adjacent non-HCI specific paradigms discussed previously such as action research (AR) also imply significant democratic commitments, they are not inherently connected to the practice of design, and were determined to be out of scope[3]. Nevertheless, works self-describing as both an adjacent paradigm and as PD/co-design were still captured by our search criteria; specifically, one study that self-described as both "participatory action research" and "co-design" was captured.

Because ASSETS publications are necessarily accessibility-focused by venue convention, we included any ASSETS publication matching our participatory methodology terms. For CHI, we required matches to both sets of terms (using the AND operator) to ensure relevance given the venue's broader scope. We searched for keyword matches in titles and abstracts, which retrieved 159 publications (103 ASSETS, 56 CHI). A complete summary of corpus attrition across retrieval, screening, and coding is given in Table 3, and our full search strings and search terms are given in Table 2. We report our study selection following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines for systematic reviews [53], summarized in the flow diagram of Figure 3.

Not all retrieved publications proved relevant. We therefore manually screened titles and abstracts to exclude papers where search terms appeared in irrelevant ways: for example, some papers used "accessibility" to refer to public availability or general access, rather than disability access.

4.2 Inclusion and Exclusion Criteria

Following manual screening, we developed and applied a set of inclusion and exclusion criteria to further refine the corpus for relevance to our research questions. We evaluated these criteria primarily using publication titles and abstracts, reviewing the full text only when the title and abstract did not include enough information to make a determination. We included a publication if all three of the following were met:

(1) The work reports a participatory methodology that was actually conducted, not just intended or referenced. This excludes publications describing planned or intended future work, as well as meta-analyses or literature reviews of participatory studies.
(2) The work primarily concerns disability, accessibility, or assistive technology. We adopt an expansive definition of disability, encompassing physical, sensory, cognitive, and psychiatric conditions, as well as chronic illness, neurodivergence, and topics related to aging. This reflects an understanding of disability as a broad and heterogeneous category [38].
(3) The primary end users of design artifacts, or primary subjects of qualitative study, are disabled people.

[3]To verify that this scoping decision did not exclude a substantial body of action research (AR)-lineage work, we re-ran our exact retrieval pipeline with "action research" substituted for the participatory-methodology terms, keeping the accessibility terms and screening procedure unchanged. The search surfaced few additional work: none at CHI, and 3 additional at ASSETS — these additional works were not added to our corpus or analysis. This scarcity suggests that within HCI accessibility research, democratically oriented design work conventionally describes itself in the participatory family's vocabulary — "participatory," "co-design," "co-creation" — rather than AR's, and our search terms follow that convention.

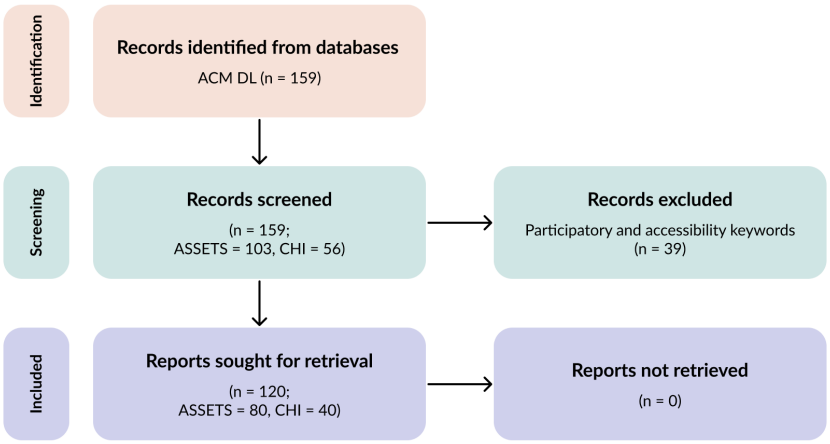


Figure 3: PRISMA flow diagram summarizing the study selection process.

This third criterion was particularly consequential for comparability. As an example, consider a tool designed to support speech-language pathologists (SLPs) who work with disabled clients: if disabled people did not participate in its design, that absence constitutes something different compared to excluding disabled people from the design of a tool used by disabled people themselves. By restricting our corpus to only include work where disabled people are the intended primary beneficiaries, we avoided conflating cases where involvement or non-involvement of disabled stakeholders carried different implications.

We excluded publications where disability was not the primary focus, where a participatory methodology was not reported, or where design artifacts were primarily intended for non-disabled end users such as caregivers. Workshop proposals and panel descriptions were also excluded, as these describe planned events rather than conducted research. Applying these criteria reduced the corpus to 120 publications (80 ASSETS, 40 CHI).

4.3 Coding Framework

We required a coding framework that could meaningfully distinguish between different modes of participation. Rather than constructing our own scale from scratch, we adapted established participatory models from adjacent disciplines. Specifically, we adapted the nine-level Wright et al. model of participation [70] as operationalized for HCI and participatory technology development by Heitplatz [34], drawing additionally on Arnstein's foundational Ladder of Citizen Participation [2]. Our goal was to assess reported levels of involvement and power-sharing within a participatory process.

We read and coded the 120 works remaining in the analytical corpus. Our corpus review process can be observed in Table 4. For each, we assessed which parts of the documented research process mapped to the four Double Diamond stages, giving each stage a score on the Stufenmodell der Partizipation. In doing so, we treat each work as providing up to 4 coded data points following similar traditions in mixed-methods research [58, 60] and meta-research practice at CHI and ASSETS [43]. Because the Stufenmodell der Partizipation is ordinal, this allows us to perform further statistical tests (Mann-Whitney, Wilcoxon signed-rank, Friedman, Kruskal-Wallis, Spearman) [61, 64] and report means as descriptive summaries with medians alongside [51]; Table 7 showcases all tests reported in this paper.

4.3.1 Assessing for Double-Diamond Stages. A holistic participation rating could have obscured variation within a project: whether the distribution of power was consistent across stages or whether it shifted at particular points. We therefore coded participation separately across the four stages of the Double Diamond design framework — Discover, Define, Develop, and Deliver — treating each stage as a distinct site of analysis. Stages did not need to occur in order, and in cases where the boundaries between adjacent stages were ambiguous, we assigned activities based on their general purpose within each diamond: conceptually divergent activities oriented toward generating possibilities were coded as Discover or Develop, while convergent activities framed as synthesis were coded as Define or Deliver. For projects which documented iterative loops, we avoided coding at the level of individual cycles, instead assigning stages based on what the iteration was collectively oriented toward achieving for the project.

For publications not explicitly using the Double Diamond framework, we mapped reported activities onto stages by function:

Discover: Early exploratory work such as interviews, ethnography, diary studies, and surveys.

Define: Synthesis activities such as thematic analysis and the formalization of design goals or requirements.

Develop: Solution-generative and iterative activities such as brainstorming, prototyping, and co-design sessions.

Deliver: Evaluative and convergent activities such as usability studies, deployment studies, and the synthesis of design recommendations.

For qualitative publications, we mapped stages onto the research process itself along a diverge–converge logic:

Discover: Exploratory background work motivating the study.

Define: Determination of research questions and study design.
Develop: Data collection.
Deliver: Analysis and synthesis of findings.

If a given publication did not engage with a particular stage, we chose to leave that stage unrated rather than rate it a score of 0. Of the 120 publications in the analytical corpus, we coded 108 as engaging with all four Double Diamond stages; the remaining 12 engaged with a subset of stages and were coded on those stages alone. Where multiple disabled stakeholders participated at different levels within a single stage, we retained the rating corresponding to the highest level of involvement at that stage.

4.3.2 Assessing with the Nine-Level Participation Scale. We adapted the Wright et al. and Arnstein frameworks for the specific context of participatory design with disabled participants. We clarified levels 2 and 3 from existing models. In Heitplatz's operationalization of the Wright et al. model, level 2 characterizes research that frames its target group as deficient and in need of correction by external experts, while level 3 characterizes research that shares outcomes with the target group but does not seek their input. In our coding framework, we defined level 2 to describe work that excludes disabled stakeholders and conceptualizes disability through a deficit lens, and level 3 to describe work without a deficit framing that nonetheless does not involve disabled stakeholder input. In Wright et al.'s model, the boundary between levels 3 and 4 marks the fundamental threshold between non-involvement and involvement. We preserved this boundary in our operationalization, consistent with both source frameworks.

For levels 4 through 9, we found existing frameworks describe the character of each level, but do not specify boundaries precisely enough to resolve ambiguous cases. We addressed this by reducing each boundary to a single explicit criterion: the distribution of structural power between disabled stakeholders and researchers. The resulting levels are:

Level 4: Disabled stakeholders are consulted but hold no decision-making power.
Level 5: Disabled stakeholders hold some decision-making power, but less than researchers.
Level 6: Decision-making power is equally shared between disabled stakeholders and researchers.
Level 7: Disabled stakeholders hold majority power within bounded domains.
Level 8: Disabled stakeholders hold majority power across the entire project.
Level 9: The project is fully and autonomously organized by disabled people at every level.

In practice, this reduced coding of levels 4–9 to a three-step process: first, determine whether disabled stakeholders held any decision-making power at all; second, determine how much power they held relative to researchers; and third, if they held more, determine whether that extended across the entire project or only within bounded domains.

In cases where disabled members were present on the research team, we applied this operationalization to assess power distribution within the team itself, comparing power held by disabled team members against power held by non-disabled researchers. In practice, the presence of disabled researchers rarely resulted in a rating below level 5, as these researchers generally held at least some decision-making authority at at least some stage(s) of the process. To determine the presence of disabled authors or researchers, we referred exclusively to claims made directly within the paper's text, and did not refer to external sources or our personal knowledge of certain researchers' positionalities.

Table 5: Self-described methodologies across the 120 scored publications. Works invoking more than one methodological label are counted under each label, so counts sum to more than 120; the percentage column reports the share of the 120-work corpus invoking each label.

Self-Described Methodology	Works	% of corpus
Co-design	69	57.5
Participatory design (PD)	46	38.3
General participation[a]	13	10.8
Co-creation	5	4.2
Inclusive design	4	3.3
Research through design (RtD)	3	2.5
Participatory action research (PAR)	1	0.8
Autoethnography	1	0.8
Human-centered design (HCD)	1	0.8
User-centered design (UCD)	1	0.8
Action research (AR)	1	0.8
Value-sensitive design (VSD)	1	0.8

[a]Applied to publications that describe their work as "participatory" without committing to a named methodological tradition.

4.4 (Not) Handling Proxy Stakeholders

While we initially intended to score all publications on the above criteria, we revised our approach after discovering systematic inconsistencies in how the first two authors coded publications involving proxy stakeholders representing disabled interests. We ultimately only scored publications involving some level of direct participation by people with lived experience of the relevant disability or condition; publications involving exclusively the participation of proxy stakeholders (such as clinicians, caregivers, or family members) were flagged for analysis but not scored. We made this decision following discussions within the research team where we identified complexities surrounding proxy involvement detailed in §3. Proxy involvement occupies a grey area between participation and non-participation that cannot be cleanly categorized within our coding scale, and this ambiguity produced inconsistencies between coders. Not assigning scores to publications without direct involvement from disabled stakeholders was our single most effective decision in resolving interrater disagreement.

4.5 Interrater Reliability

The first two authors independently coded a randomly selected reliability sample of 30 papers to validate the coding framework. We report Krippendorff's α as our primary reliability measure. Following an initial coding round of 15 papers, we identified systematic

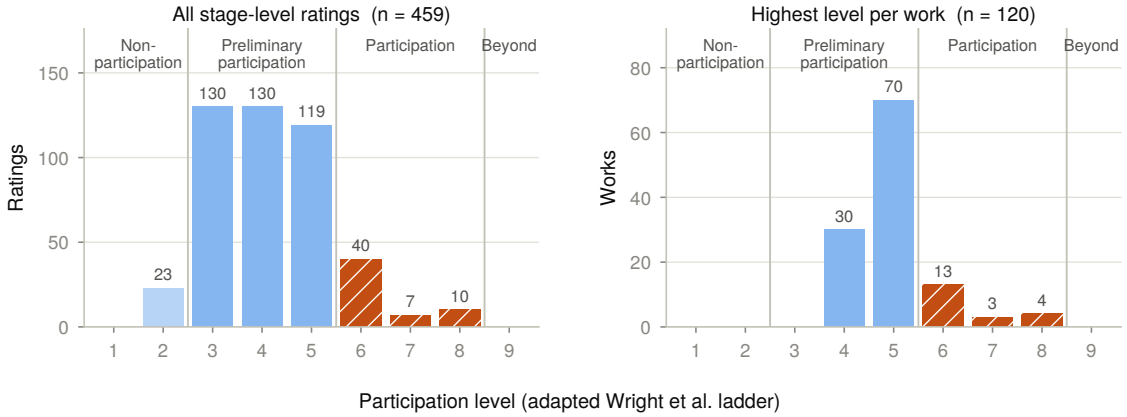


Figure 4: Participation ratings on the adapted Wright et al. [70] ladder: all 459 stage-level ratings (left) and each work's highest level at any stage (right). Hatched orange marks levels at which participants hold decision-making authority (6+). Only 20 of 120 works (16.7%) ever reach this threshold.

Table 6: Per-stage inter-rater reliability on a 30-paper reliability sample after codebook reconciliation, reported as Krippendorff's α

Stage	Krippendorff's α
Discover	.825
Define	.921
Develop	.824
Deliver	.820
Overall	.903

sources of disagreement and revised our approach, holding discussion meetings to resolve ambiguities and updating our codebook accordingly. A major revision involved modifying coding instructions to not score publications involving exclusively proxy participation as described above, which resolved the primary source of disagreement. After a reconciliation round applying these revisions, overall Krippendorff's α was .903; per-stage values are reported in Table 6. When we analyzed residual disagreements, they were found to be concentrated in two areas: the boundary between participation levels 2 and 3, and the boundary between participation levels 4 and 5.

5 Findings

We present our findings in five parts. First, we show that participation depth across the corpus falls predominantly below the threshold of shared decision-making, with the majority of papers operating within the preliminary-involvement band at all stages. Second, we highlight that the participatory design (PD) label paradoxically correlates with lower participation than co-design. Third, we demonstrate that participation is unevenly distributed across the design process — peaking during solution generation (Develop) and dropping during problem framing (Define) — a pattern consistent across both venues and topic areas. Fourth, through a clustering analysis, we identify two distinct paper archetypes: a large majority that mirrors this uneven aggregate pattern, and a smaller minority that maintains a uniformly higher participation profile. This indicates that the drop at Define is not an inherent structural limitation of the methodology, but rather a reflection of how participatory efforts are currently distributed in practice. Finally, we report that declared disabled authorship is strongly associated with both higher and more uniform participation across stages.

Our corpus of 120 included papers yields 459 stage-level ratings. Unless otherwise noted, paper-level analyses use $N = 120$ and stage-level analyses use $N = 108$. The corpus grand mean participation level across all ratings is $M = 4.18$ ($SD = 1.24$; $Mdn = 4$), situating the average level squarely within the "Consultation" rung of the adapted Wright et al. ladder — the band in which researchers actively seek participant input but retain decision-making authority.

5.1 A Ceiling on Participation

Participation depth across the corpus falls predominantly below the threshold of shared decision-making. Of the 459 stage-level ratings in our corpus, 82.6% fall within the "Preliminary Participation" band (levels 3–5), where disabled stakeholders may be involved, but researchers retain most decision-making authority. Only 12.4% reach the "Participation" band (levels 6–9), where disabled stakeholders hold partial or full decision-making power. The remaining 5.0% represent stages where disabled people were not involved at all.

A parallel pattern emerges at the publication level. Despite consisting entirely of self-described participatory work, only 16.7% of the papers reached the Participation band at any stage of the design process. The remaining 83.3% never exceeded level 5 at any point. For these works, disabled stakeholders were present, but structural authority remained with researchers at every stage.

Taken together, these results show a systemic discrepancy between the participatory framing of the included studies and the

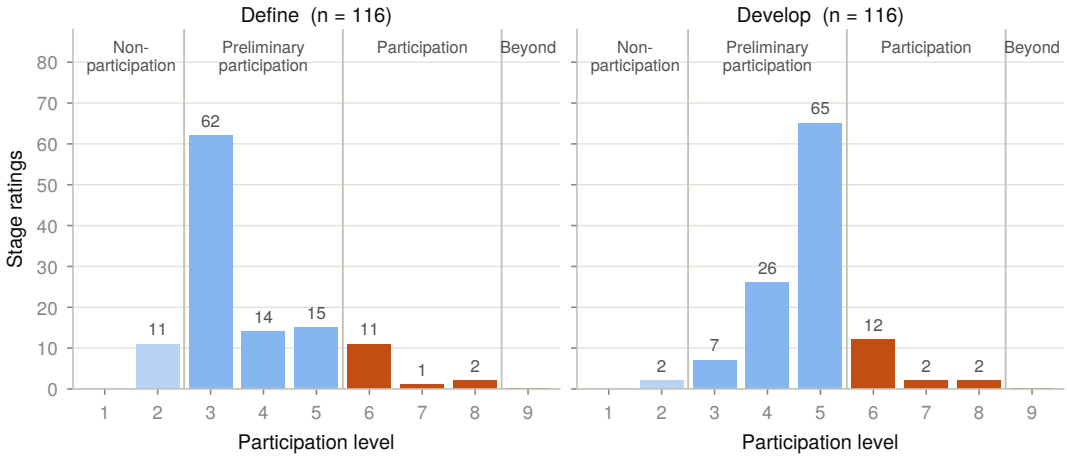


Figure 5: Distribution of participation scores at the Define (left) and Develop (right) stages across the nine-level ladder.

reality of their reported practice: the majority of papers never reach the threshold of genuine participation as defined by Wright, von Unger, and Block [70]. The participatory label predominantly describes a practice of involvement where decision-making authority remains with the research team.

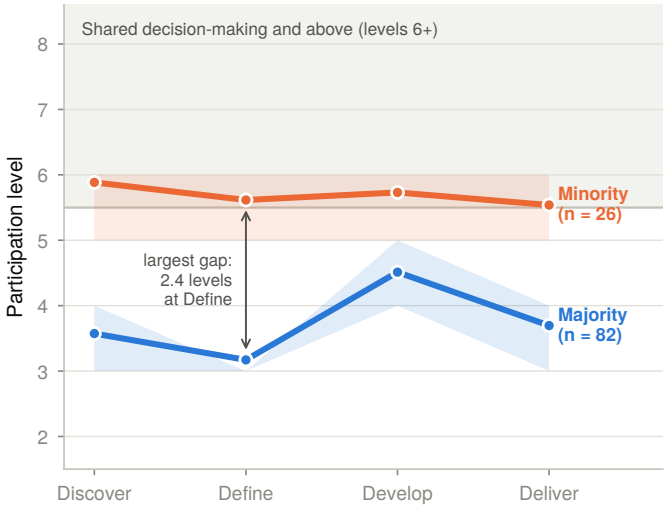


Figure 6: Participation trajectories across Double Diamond stages for the two k-means clusters ($k = 2$, $n = 108$).

5.2 The PD Label is Linked with Lower Participation

This ceiling is not evenly distributed across methodological labels. As detailed in Table 7, papers exclusively using the "participatory design" (PD) label reported significantly lower participation scores ($M = 3.75$) than those exclusively using "co-design" ($M = 4.42$).

This directional effect persists across every stage of the Double Diamond, with the widest gap occurring during Define. Papers claiming both labels patterned closely with the PD-only subset, indicating that the PD vocabulary itself anchors the lower participation levels. This is a counterintuitive finding, given that PD's Scandinavian labor-movement origins are explicitly committed to redistributing decision-making authority [8, 67].

5.3 Many Works Emphasize Develop, While Define Remains Underaddressed

Participation is not distributed evenly across the Double Diamond framework. Statistical testing confirms a significant effect of the design stage on participation levels (see Table 7 for full Friedman and Wilcoxon results). Develop is rated significantly higher than every other stage, while Define is rated significantly lower. Discover and Deliver fall between these two extremes. (Median at each stage: 4 at Discover, 3 at Define, 5 at Develop, and 4 at Deliver.)

The pattern becomes starker when analyzing per-paper maximums. Among the 57 papers with a single highest stage, 75.4% peaked at Develop; zero papers peaked at Define. For the majority of included studies, Develop is the primary — and often only — site of deeper participation, while Define is consistently deprioritized.

This inequality reflects two structural dimensions. First, participation is systematically higher in the solution phase (Develop and Deliver) than in the problem phase (Discover and Define). Second, divergent stages (Discover, Develop) sustain higher participation than convergent stages (Define, Deliver). Develop sits at the intersection of both higher-participation dimensions, while Define sits at the exact opposite.

5.4 Two Paper Archetypes

The aggregate "Develop-peak" describes the average shape of self-labeled PD in accessibility research, but it leaves an important question unanswered: is this peak a uniform practice across all papers, or is it an average of fundamentally different approaches? If it is uniform, the drop at Define marks the ceiling of what the field currently knows how to do. If it is a mixture, the aggregate shape reflects current community norms and concentrations of effort rather than the structural limits of the methodology.

ID	Analysis	Test	n	Test statistic	Result
§5.2 Participation and methodological label					
1	Across four label subgroups (PD only, co-design only, both, neither)	Kruskal–Wallis	108	$H(3) = 11.55$, $p = .009^{**}$	Subgroups differ
2	PD (any) vs. non-PD (any), overall	Mann–Whitney	108	$U = 795.5$, $p < .001^{***}$, $r = 0.39$	PD lower
3	PD (any) vs. non-PD (any) at Discover	Mann–Whitney	108	$U = 861.0$, $p = .002^{**}$	PD lower
4	PD (any) vs. non-PD (any) at Define	Mann–Whitney	108	$U = 821.5$, $p < .001^{***}$	PD lower
5	PD (any) vs. non-PD (any) at Develop	Mann–Whitney	108	$U = 932.0$, $p = .007^{**}$	PD lower
6	PD (any) vs. non-PD (any) at Deliver	Mann–Whitney	108	$U = 968.5$, $p = .020^{*}$	PD lower
7	PD-only vs. co-design-only, overall	Mann–Whitney	83	$U = 449.5$, $p = .003^{**}$, $r = 0.41$	PD-only lower
8	PD-only vs. co-design-only at Discover	Mann–Whitney	83	$U = 499.5$, $p = .009^{**}$	PD-only lower
9	PD-only vs. co-design-only at Define	Mann–Whitney	83	$U = 501.0$, $p = .008^{**}$	PD-only lower
10	PD-only vs. co-design-only at Develop	Mann–Whitney	83	$U = 511.0$, $p = .008^{**}$	PD-only lower
11	PD-only vs. co-design-only at Deliver	Mann–Whitney	83	$U = 550.5$, $p = .035^{*}$	PD-only lower
§5.3 Stage-level inequality					
12	Overall stage effect	Friedman	108	$\chi^2(3) = 80.87$, $p < .001^{***}$, $W = 0.25$	Stages differ
13	Discover vs. Define	Wilcoxon	108	$W = 238.0$, $p_{\text{adj}} < .001^{***}$	Discover > Define
14	Discover vs. Develop	Wilcoxon	108	$W = 590.0$, $p_{\text{adj}} < .001^{***}$	Develop > Discover
15	Discover vs. Deliver	Wilcoxon	108	$W = 1100.0$, $p_{\text{adj}} = 1.00$	No difference
16	Define vs. Develop	Wilcoxon	108	$W = 138.5$, $p_{\text{adj}} < .001^{***}$	Develop > Define
17	Define vs. Deliver	Wilcoxon	108	$W = 421.5$, $p_{\text{adj}} = .0015^{**}$	Deliver > Define
18	Develop vs. Deliver	Wilcoxon	108	$W = 357.0$, $p_{\text{adj}} < .001^{***}$	Develop > Deliver
§5.4 Cluster analysis					
19	Cluster-count selection	Silhouette score	108	$k = 2$: 0.50; 3: 0.38; 4: 0.30; 5: 0.29	$k = 2$ selected
20	Majority vs. minority cluster (overall)	Mann–Whitney	108	$U = 0.0$, $p < .001^{***}$	Clusters differ
21	Stage effect within majority cluster	Friedman	82	$\chi^2(3) = 94.27$, $p < .001^{***}$, $W = 0.38$	Pattern present
22	Stage effect within minority cluster	Friedman	26	$\chi^2(3) = 4.93$, $p = .18$, $W = 0.06$	Pattern absent
23	Develop vs. Define, minority cluster	Wilcoxon	26	$W = 12.0$, $p = .39$	No difference
24	Paper mean × within-paper variance	Spearman	108	$\rho = -0.35$, $p < .001^{***}$	Higher mean, flatter profile
§5.5 Disabled authorship					
25	Disabled-author vs. other works, per stage	Mann–Whitney	111–116	U = 2259.5/2293.0/2135.0/2041.0, all $p < .001^{***}$	Disabled-author higher
26	Within-work variation across stages	Mann–Whitney	108	$U = 840.5$, $p = .005^{**}$	More uniform profiles
27	Cluster membership × declared authorship	Fisher's exact	108	23/26 vs. 11/82, $p < .001^{***}$	Concentrated in minority

Table 7: Summary of statistical tests reported in the Findings section. All tests use the reconciled corpus ($N = 120$ included works; $N = 108$ with complete four-stage ratings). All Wilcoxon tests are paired signed-rank tests; all Mann–Whitney tests are two-sided. Post-hoc comparisons after Friedman's test use Bonferroni correction for six pairwise comparisons (p_{adj}). Effect sizes are reported where applicable: Kendall's W for Friedman tests, rank-biserial r for Mann–Whitney tests. Per-stage disabled-authorship comparisons use the works engaging each stage ($n = 116$ for Discover, Define, and Develop; $n = 111$ for Deliver), with U values listed in stage order. Significance: $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

To distinguish these possibilities, we clustered the subset of works which we were able to assign scores for all four Double Diamond stages ($N = 108$) using k-means ($k = 2$). The two resulting clusters differ substantially in both level and shape:

- **The Majority Cluster (76% of papers):** This group reproduces the "higher Develop scores / lower Define scores" shape. With an overall mean of 3.74, this cluster operates within the preliminary-participation band, touching level 5 only during Develop.
- **The Minority Cluster (24% of papers):** This group is completely flat across all four stages, with an overall mean of 5.69. These works demonstrate consistent and substantive participant involvement across the entire research lifecycle.

The uneven distribution is concentrated in the majority cluster. If the minority archetype genuinely shares authority more evenly across the research lifecycle, the unequal distribution reported earlier should disappear when this cluster is isolated. As shown in Table 7, this is exactly what happens. Within the majority cluster, the effect of the design stage on participation remains large and significant. Within the minority cluster, the effect essentially vanishes.

This addresses the hypothesis that Define stage of research is inherently difficult to incorporate participation because synthesis

and problem-framing require a type of analytical expertise that is difficult to share. If that were true, the lower scores at Define would persist even among papers that invest heavily in participatory methods. Instead, we observe the opposite: in the 26 papers that sustain higher participation overall, Define is not significantly lower than in any other stage.

Interpretation. These results locate the uneven distribution of participation within a specific subset of the corpus: the ∼76% of papers operating at the lower rungs of the preliminary-participation stages. This confirms, at scale, the misconception that narrowly equates participatory design with design ideation or prototyping [55]. However, the minority cluster demonstrates that this pattern is not an inevitability of the research process. The gap between the participatory label and reported practice is concentrated at specific points in the process — widest when consequential framing decisions are made.

5.5 Disabled Authorship Tracks With Higher, More Uniform Participation

Of the 120 included works, 34 (28%) declare a disabled researcher or co-author in some form, whether through a positionality statement, an author note, or in-text identification. These works rate higher at every Double Diamond stage than works without declared disabled authorship (Mann–Whitney; Discover $U = 2259.5$, Define $U = 2293.0$, Develop $U = 2135.0$, Deliver $U = 2041.0$; all $p < .001$; medians 5 vs. 4, 5 vs. 3, 5 vs. 5, and 5 vs. 4, respectively), and their Participation Profiles are more uniform, with lower within-work variation across stages (Mann–Whitney on within-work standard deviations, $U = 840.5$, $p = .005$; median SD 0.54 vs. 0.96).

Declared disabled authorship also concentrates where participation is highest. In the cluster analysis of §5.4, 23 of the 26 minority-cluster works (88%) declare a disabled author or co-author, against 11 of 82 (13%) in the majority cluster (Fisher's exact, $p < .001$). Of the 20 works that reach the Participation band (level 6 and above) at any stage, 18 declare disabled authorship. Declared disabled authorship is, by these measures, the strongest single correlate of sustained and substantive participation in our corpus. These figures describe declaration in the published record; works with undeclared disabled authors are counted among the 86 without declared disabled authorship (§7).

6 Discussion

Our findings (§5) gave us strong indications about how participatory methods are currently enacted in accessibility and AT research at CHI and ASSETS, and our coding process surfaced quandaries about proxy representation that we took up in §3. In this discussion, we situate these patterns in the broader context of accessibility research's historical commitments to disabled people, the epistemic stakes of where decision-making authority sits, and the implications for the disabled people whose lives are affected by our work.

6.1 One Vocabulary, Divergent Practices

Our k-means clustering ($k = 2$) demonstrated that our corpus consisted of two groups with markedly different Participatory Profiles. The majority archetype (76% of the corpus) concentrates participation at Develop and thins it sharply at Define and Deliver. In contrast, the minority archetype (24%) sustains participation evenly across all four stages, with the Define trough absent entirely. These configurations imply vastly different things about where disabled stakeholders hold authority, yet both groups use the exact same labels to describe their work — e.g., "participatory design," "co-design," "co-creation." This disconnect between vocabulary and reported practice is sharpest at participatory design: the label whose political lineage most explicitly commits to sharing decision-making authority. Ultimately, this indicates that the participatory label has become decoupled from the commitment its history carried.

We recognize that our data only grants us access to what authors report, which may differ from what actually occurred. A paper with substantive Define-stage participation might score poorly if space constraints prevented documentation. Therefore, we categorically reject a shallow reading of our data that assumes researchers in the majority cluster are actively choosing to hoard authority, or that deliberate label misuse is rampant. Rather, our data highlights that the uneven low-Define, higher-Develop profile is not inherent or inevitable. The minority cluster proves that constant, uniformly substantive participation is achievable — which, in our corpus, typically occurred when projects featured a disabled researcher or co-author with lived experience (§5.5).

6.1.1 Define and Deliver is Where Authority Manifests. The corpus overall demonstrates notably lower participation during the convergent phases of a project (Define and Deliver). We argue these stages matter disproportionately because they represent the points of commitment to key decisions. While divergent stages (Discover and Develop) generate possibilities for what the problem or solution *could* be, convergent stages dictate which framing the project commits to and which solution actually proceeds.

When researchers retain authority at Define and Deliver, participant contributions arrive merely as inputs to decisions researchers ultimately make; these inputs can be incorporated, reframed, or set aside. Consequently, in the majority archetype, disabled stakeholders hold the least authority precisely when the most consequential decisions occur.

This concentration of authority carries dual consequences for the ethics and substance of our work. Ethically, "nothing about us without us" [11] is a claim of epistemic justice: disabled people hold authoritative knowledge about their own lives [44] and deserve power over decision-making as they ultimately bear the consequences. Consulting disabled participants during ideation while researchers retain deciding authority falls short of that commitment. It risks reducing participatory design to what Qi and Yu [55] describe as "design ideation or prototyping," hosting participatory activities only where it won't significantly redistribute project control. Substantively, the technologies emerging from the majority cluster inherently reflect problem framings and solution selections prioritized by researchers rather than decided by disabled stakeholders. This is not to say these technologies are harmful or produced in bad faith, but is instead a narrower observation about design authority: where disabled stakeholders hold true authority at Define and Deliver — as they do in the minority archetype — the

final work reflects their framings of their own priorities, not just researchers' interpretations of them.

6.2 The Politics of Declared Disabled Authorship

The association in §5.5 is an association with *declared* authorship, and declaration is a consequential act. Disability disclosure in academia carries professional risk, and disabled and chronically ill scholars weigh that risk against institutional cultures that were not built to include them [10]; many disabilities are invisible, and this precarity means that disabled authors with lived experience choose not to disclose. The 29% of works in our corpus with declared disabled authorship should therefore be seen as the floor on disabled participation in the corpus.

We surmise the following: disabled researchers may steer projects toward deeper power-sharing, bringing lived experience to bear on which decisions are shared and with whom. And projects already committed to power-sharing may be the environments in which disclosure feels safe, relevant, and valued enough to make. We note that there is no way to separate these possibilities from the written text alone, and the association we report of higher participation for works with declared disabled authorship should be read as evidence of their entanglement rather than of either alone.

This entanglement shapes what accountability should considered of researchers. A reporting regime that rewarded declared disabled authorship would incentivize disclosure, pressing scholars to out themselves for their work to be credited — a dynamic that disability scholars have warned against [10]. In this vein, we ask the field to consider the *Participation Profile Card* (§6.3) as a potential instrument organized around arrangements rather than identities: it merely asks who held which decisions at which stages, and leaves disclosure of any participant's or author's disability status where it belongs: up to one's individual discretion.

6.3 Toward Shared Accountability: Researchers, Reviewers, and the Field

The gap between participatory vocabulary and practice requires structural intervention. We propose three prescriptive moves aimed at researchers planning studies, reviewers evaluating claims, and the field collectively building infrastructure for accountability. These moves rely on the *Participation Profile* — the Double Diamond paired with the Stufenmodell der Partizipation ladder from Wright et al. [70] — as a shared diagnostic. We note that the Profile framework does not resolve the philosophically hard questions raised by this analysis; its purpose is to make those questions legible and discussable where they have largely been implicit.

6.3.1 For researchers: plan stage-by-stage with tractable exemplars. We propose the *Participation Profile* as a framework for planning engagement before a project begins, not just for reporting it afterwards. Rather than treating participation as a monolithic target to maximize, researchers can use the *Profile* to make stage-by-stage commitments visible and to surface tradeoffs honestly. The majority archetype's participation is thinnest at Define and does not carry its Develop peak through Deliver, which suggests that the most effective intervention is not an immediate leap from Consultation (Level 4) to Decision-Making Authority (Level 8), but a deliberate move from Level 4 to Level 5 or 6 at particularly consequential stages. Elevating participation at any given stage carries real costs: time before scoping is settled, compensation for early involvement, and the willingness on the part of researchers to have initial framings altered. An honest and realistic plan identifies where deeper participation is prioritized and where it is not, rather than optimizing those tradeoffs away. The minority archetype demonstrates that sustained participation across stages is achievable; the corpus supplies concrete exemplars of incremental moves between levels.

At Discover, Lee et al. [42] illustrates a shift from researcher-led needs-finding toward joint authorship of the problem space: the work of identifying current practices and challenges in accessible collaborative editing was conducted through co-design sessions with a blind coauthor on the research team, rather than through interviews synthesized after the fact. The cost is the added complexity of structuring early-stage work around a collaborator's availability and authority rather than a researcher's timeline; the return is a problem frame that enters Define already jointly held. At Develop, Mouallem et al. [47] grant BLV collaborators partial authority within a bounded scope during a sustained 2.5-year co-design process for an accessible circuit simulator, at the cost of researchers accepting design choices they might not personally prefer and with the return that solution candidates entering Deliver are not pre-filtered by researcher bias. At Deliver, Barbareschi et al. [4] document an even more striking move: after an initial workshop, four older women chose independently to run a second session for fifteen local children, with the research team in a supporting rather than leading role. The Delivery stage in that project goes beyond co-evaluated to self-determination by participants. Suchanek et al. [65] sit at a further extreme still, documenting a project in which deaf collaborators drove the agenda across both solution stages, with hearing collaborators in an explicitly supporting capacity.

6.3.2 For reviewers: concrete expectations, concrete mechanisms. The field-wide prevalence of the majority archetype directly implicates peer review. A paper can currently claim "co-design" or "participatory design" without specifying how power was distributed across the design lifecycle, leaving reviewers without a basis to evaluate the claim. We propose three concrete changes to how participatory work is reviewed at accessibility venues.

First, program committees could add an explicit prompt to review forms for any paper claiming participatory methods, asking reviewers to assess whether the authors have reported stage-level engagement: at which stages disabled principals engaged, with what level of decision-making authority, through what practices, and — if proxies were involved — on what basis.

Second, reviewers should treat undifferentiated claims of "participatory design" or "co-design" with the same skepticism they would apply to undifferentiated claims of "qualitative analysis," requiring authors to specify the practices that enacted participation at each stage rather than accepting the label as self-evidencing.

Third, and critically, reviewers should treat this as a reporting standard rather than a scoring standard. A paper transparently reporting Level 4 engagement at Define and offering principled reasons for that choice serves the field better than a paper implying higher participation through vague language. Reviewers

should evaluate the coherence of a paper's *Participation Profile* — whether its stated commitments align with its reported practices and its claimed contributions — rather than demand uniform maxima across all stages.

6.3.3 For the field: shared infrastructure, and two open questions. To combat the definitional drift documented in this review and others, the field needs shared infrastructure: a common vocabulary for contested categories and concrete artifacts through which authors declare, and reviewers evaluate, participation claims. We propose as a starting point a *Participation Profile Card*, in the lineage of Model Cards for model reporting [46] and Datasheets for Datasets [19], to accompany any paper claiming participatory methods.

The Card has four parts.

(1) It describes the participants and collaborators: the disabled principals engaged, how they were recruited and compensated, and — if proxies such as caregivers, clinicians, or teachers were involved — the basis for that involvement and the principals' relationship to those proxies.
(2) It specifies stage-level engagement: for each Double Diamond stage present in the project, the participation level of principal participation, the specific practices that enacted that level, and the concrete decisions made at that stage and by whom.
(3) It documents the boundaries of ceded authority: what researchers explicitly retained (publication, funding, technical constraints, scope-setting), and why those boundaries were drawn where they were.
(4) It offers a reflection: what structural constraints limited participation, and what the authors would do differently.

The *Participation Profile Card* is not meant to be something used to "keep score" but an artifact of disclosure that makes the participation claim legible and accountable.

Like Model Cards or Datasheets, the *Participation Profile Card* is designed as a diagnostic and accountability instrument. Similarly, Participation Profile Cards do not resolve the philosophically hard questions raised by this analysis; they make those questions legible where they have largely been implicit. We close by naming two such questions that the accessibility and AT communities must collectively confront.

The first is the question of ends. Does (ever) higher participation on the participation ladder represent a destination the field should be climbing toward, or is it itself a symptom of a more fundamental commitment the field has not yet made? Suchanek et al. [65] argue explicitly that participatory methods "do not require researchers to actively reflect on power relationships as would be required when aiming for transformative impact," and propose solidarity-driven research as a methodological move beyond participation rather than at its apex. Under that framing, Wright et al.'s ladder is a diagnostic of current practice, not a destination. Whether the field's aspiration should be ever-higher rungs, or an alternative orientation altogether — solidarity, disability-led research, something yet to be articulated — is not a question this paper can settle. Even full-stage participation leaves open *who*, within a disability community, is participating: access to recruitment networks, information, and resources is unevenly distributed among disabled people themselves [28, 30], and a project may satisfy every stage of the Profile while drawing only on its community's most-resourced members. Participation, on this view, remains a spectrum even at the ladder's upper rungs. It is a question the field has yet to address in the open, and part of the Participation Profile's goal is to encourage such discussion.

The second is the question of proxies. The principals-only rule we adopt in Section 4 is perhaps a defensible methodological choice, but it is not a normative verdict. It treats proxy engagement — with clinicians, caregivers, family members, teachers — as categorically different from engagement with disabled principals, and rightly so: the two raise distinct questions of authority and representation. But the rule leaves unsettled how the field ought to reason about projects in which principal participation is structurally limited (young children, people with severe cognitive impairments, communities that are themselves proxy-mediated) and proxies are carrying much of the representational weight, for better or worse. Flattening these cases to a single low score discards information; treating proxies as equivalent to principals erases a distinction that matters politically. We again do not attempt to resolve this, and turn to the field for its engagement.

Neither of these questions can be answered by an accountability artifact. Both can be asked more productively once the field shares a vocabulary for what is actually happening in the design processes it produces. That is what the Participation Profile is for, and what the Participation Profile Card would institutionalize.

7 Limitations

Our analysis represents one situated reading of participatory design in accessibility research, shaped by deliberate methodological choices. We examined English-language publications from CHI and ASSETS between 1998–2025, enabling systematic comparison within two venues central to HCI's accessibility discourse. We therefore make no claim about participatory commitments as they are practiced in adjacent sites, such as disability studies, rehabilitation research, and regional or non-English outlets — where these commitments may take substantively different forms.

We also note that the *Participation Profile* — the Double Diamond paired with an adapted von Unger ladder — is one analytical path among several. Alternative lenses, such as Sanders and Stappers' convivial toolbox [59] or Arnstein's original ladder [2], would foreground different aspects of participation, and we view these as complementary directions rather than competing ones. We also note that the Double Diamond framing of a design process carries a blind spot: it ends at Deliver. The long labor of sustaining an intervention in a community after a project formally concludes — maintenance, repair, handoff, and the relational work of remaining accountable to participants — has no representation in the Double Diamond framework. Harrington, Erete, and colleagues have argued that dominant design engagements systematically neglect this relational work: the trust-building, administrative, and community-care labor that surrounds and outlasts the formal design process is rendered invisible by methods and institutions that privilege discrete, reportable design activity [17, 28]. Empirical accounts of community-based computing initiatives likewise identify sustainment as among the most difficult and least-supported dimensions of such work [41]. Process frameworks that render sustaining labor

illegible make it easier for projects to treat departure as completion. Our stage-resolved instrument inherits this boundary: participation in post-project sustainment is invisible to the *Participation Profile* as we operationalize it.

Our principals-only coding rule, while central to the reproducibility of our analysis, does not credit the frequently essential contributions of caregivers, clinicians, educators, and other supporters whose involvement is often inseparable from working alongside disabled participants. This is a deliberate scoping choice, not a claim that such contributions are unimportant.

Finally, our ratings reflect what works report in writing, which is not necessarily the same as what participation-sharing actually occurred. A study that underdocuments a substantive Define-stage engagement receives the same rating as one that never held it. Relatedly, many self-declared participatory papers locate their problem formulation almost entirely in related work and then engage disabled participants from Develop onward. Our stage-level ratings treat each paper as the unit of analysis, though we attempted to ascertain whether prior work was conducted in a participatory fashion and reflect this in our ratings (this was rare). Our working principle was that authority inherited from the literature is not authority exercised in the reported study. In the rare cases where we did credit antecedent participatory work, our ratings may in fact be inflated: this review's own findings suggest that a substantial share of self-declared PD work is itself thin on participation, meaning that taking prior participatory claims at face value likely led to a more generous interpretation than a closer audit may warrant.

As noted in §(6.2), we've chosen to give some leeway in the case of disclosed disabled authorship. When a paper's text indicated that an author was themselves a member of the disabled community under study, we rated as if that author held decision-making authority across the stages in which they were involved — an interpretive choice that extends benefit of the doubt and, as a result, may cause our analysis to *undercount* tokenization where disabled authorship was declared but not substantively enacted. Conversely, where we personally knew an author to be disabled but this was not disclosed in the paper, we made no such assumption. The patterns we identify are therefore patterns in the published record of participatory accessibility research — a record that itself warrants continued scrutiny, and one we hope this review helps to examine more systematically.

8 Conclusion

Our review asked whether participatory methods in accessibility and AT research fulfill their core promise: that disabled people hold authoritative influence over the technologies designed around their lives. Across 120 papers from CHI and ASSETS, we found they predominantly do not. Mean participation stalls at Consultation (level 4), where disabled participants provide input but lack decision-making power. More than 80% of works never reach shared authority at any stage. This gap opens widest at problem framing (Define) and, paradoxically, is most severe among papers explicitly invoking the PD label.

Yet, this systemic gap is not a structural inevitability. A minority of works in our corpus successfully sustain high participation across all four design stages, demonstrating that substantive power-sharing is practically achievable. Their existence demonstrates that the common pattern of focusing on participation at solution generation (Develop) while overlooking participant authority at Define reflects selective researcher investment rather than an inherent limitation of the methodology.

We call on the field to shift from claiming participation as a broad label to accounting for it as a specific, distributed practice. We offer the Participation Profile as a shared mechanism for this accountability. For researchers, it provides a framework to intentionally plan power-sharing stage by stage. For reviewers, it offers a standard to evaluate participatory claims against reported practice. For the accessibility community at large, it serves as a foundation for establishing the shared methodological standards we currently lack. By making authority visible and accountable, we can move beyond including disabled communities and begin honoring the redistributive promise that participatory methods were designed to carry.

The authors used generative AI tools (Claude and ChatGPT) on a limited basis during the preparation of this manuscript. Specific uses included copy-editing assistance for grammar and phrasing, generating LaTeX scaffolding for tables, and drafting Python code for statistical analyses and figure generation. All conceptual contributions, coding decisions, statistical interpretations, and substantive writing are the authors' own, and any AI-assisted output was reviewed, verified, and edited by the authors before inclusion.

Acknowledgments

We would like to thank Patricia Piedade, Beatriz Severes, and Ningjing Tang for their support and encouragement in this work. In addition, we would like to express our gratitude to our anonymous reviewers for their engaged feedback through the review process.

References

[1] Erin E Andrews and Anjali J Forber-Pratt. 2022. Disability culture, identity, and language. In *The positive psychology of personal factors: Implications for understanding disability*, Dana S. Dunn and Michael Wehmeyer (Eds.). Lexington Books, London, UK, 27–40. https://books.google.com/books?hl=en&lr=&id=MQdSEAAAQBAJ&oi=fnd&pg=PA27&dq=disability+justice+identity+first+language&ots=l_FHNVaMWi&sig=614pFwcNBEdRb5-zBWf88XESlTA

[2] Sherry R. Arnstein. 1969. A Ladder Of Citizen Participation. *Journal of the American Institute of Planners* 35, 4 (July 1969), 216–224. doi:10.1080/01944366908977225

[3] Philip Baillargeon, Jina Yoon, and Amy Zhang. 2025. Who Puts the "Social" in "Social Computing"?: Using A Neurodiversity Framing to Review Social Computing Research. *Proceedings of the ACM on Human-Computer Interaction* 9, 2 (May 2025), 1–44. doi:10.1145/3711106

[4] Giulia Barbareschi, Chihiro Sato, Seray Senyer, Michael Pan Junpeng, Jianrui Zhao, Dunya Chen, Kirsten Ellis, and Kai Kunze. 2025. Sticking With Electronics for Crafting Practices: An Inclusive Approach to Promote Making Literacy Among Older Adults. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 668, 18 pages. doi:10.1145/3706598.3714234

[5] Cynthia L. Bennett, Erin Brady, and Stacy M. Branham. 2018. Interdependence as a Frame for Assistive Technology Research and Design. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Galway Ireland, 161–173. doi:10.1145/3234695.3236348

[6] Jarg Bergold and Stefan Thomas. 2012. Participatory Research Methods: A Methodological Approach in Motion. *Historical Social Research / Historische Sozialforschung* 37, 4 (142) (2012), 191–222.

[7] Gro Bjerknes, Pelle Ehn, and Morten Kyng (Eds.). 1987. *Computers and Democracy: A Scandinavian Challenge*. Avebury, Aldershot, UK.

[8] Susanne Bødker and Morten Kyng. 2018. Participatory Design that Matters—Facing the Big Issues. *ACM Transactions on Computer-Human Interaction* 25, 1 (Feb. 2018), 1–31. doi:10.1145/3152421

[9] Barry W. Boehm. 1988. A Spiral Model of Software Development and Enhancement. *Computer* 21, 5 (1988), 61–72. doi:10.1109/2.59

[10] Nicole Brown and Jennifer Leigh. 2018. Ableism in Academia: Where Are the Disabled and Ill Academics? *Disability & Society* 33, 6 (2018), 985–989. doi:10.1080/09687599.2018.1455627

[11] James I. Charlton. 2000. *Nothing About Us Without Us: Disability Oppression and Empowerment.* University of California Press, Berkeley, CA.

[12] Sasha Costanza-Chock. 2018. Design Justice: towards an intersectional feminist framework for design theory and practice. doi:10.21606/drs.2018.679

[13] Sasha Costanza-Chock. 2020. Design Practices: "Nothing about Us without Us.". In *Design Justice.* https://designjustice.mitpress.mit.edu/pub/cfohnud7

[14] Aashaka Desai, Aaleyah Lewis, Sanika Moharana, Anne Spencer Ross, Jennifer Mankoff, and Christina N. Harrington. 2025. Race, Disability, and Technology: A Call to Action for Accessibility Researchers. *ACM Transactions on Accessible Computing* (Nov. 2025), 3772276. doi:10.1145/3772276

[15] Design Council. 2005. The Double Diamond: A Universally Accepted Depiction of the Design Process. https://www.designcouncil.org.uk/our-resources/the-double-diamond/. Framework launched by the UK Design Council in 2005; subsequently codified in the Design Council's 2007 report *Eleven Lessons: Managing Design in Eleven Global Brands* and updated in the 2019 *Framework for Innovation.*

[16] Pelle Ehn. 1988. *Work-Oriented Design of Computer Artifacts.* Arbetslivscentrum, Stockholm, Sweden. Ph.D. dissertation, originally published by Almqvist & Wiksell International.

[17] Sheena Erete, Yolanda A. Rankin, and Jakita O. Thomas. 2023. A Method to the Madness: Applying an Intersectional Analysis of Structural Oppression and Power in HCI and Design. *ACM Transactions on Computer-Human Interaction* 30, 2, Article 24 (April 2023). doi:10.1145/3507695

[18] Christopher Frauenberger, Judith Good, Geraldine Fitzpatrick, and Ole Sejer Iversen. 2015. In pursuit of rigour and accountability in participatory design. *International Journal of Human-Computer Studies* 74 (Feb. 2015), 93–106. doi:10.1016/j.ijhcs.2014.09.004

[19] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. Datasheets for Datasets. *Commun. ACM* 64, 12 (2021), 86–92. doi:10.1145/3458723

[20] Maria J. Grant and Andrew Booth. 2009. A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies. *Health Information & Libraries Journal* 26, 2 (2009), 91–108. doi:10.1111/j.1471-1842.2009.00848.x

[21] Joan Greenbaum and Morten Kyng (Eds.). 1991. *Design at Work: Cooperative Design of Computer Systems.* Lawrence Erlbaum Associates, Hillsdale, NJ.

[22] Judith Gregory. 2003. Scandinavian Approaches to Participatory Design. *International Journal of Engineering Education* 19, 1 (2003), 62–74.

[23] Alisa Grigorovich, Pia Kontos, Amanda Jenkins, and Susan Kirkland. 2022. Moving Toward the Promise of Participatory Engagement of Older Adults in Gerotechnology. *The Gerontologist* 62, 3 (March 2022), 324–331. doi:10.1093/geront/gnab026

[24] Kim Halskov and Nicolai Brodersen Hansen. 2015. The Diversity of Participatory Design Research Practice at PDC 2002–2012. *International Journal of Human-Computer Studies* 74 (2015), 81–92. doi:10.1016/j.ijhcs.2014.09.003

[25] Foad Hamidi, Melanie Baljko, and Isabel Gómez. 2017. Using Participatory Design with Proxies with Children with Limited Communication. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility.* ACM, Baltimore Maryland USA, 250–259. doi:10.1145/3132525.3132527

[26] Foad Hamidi, Patrick Mbullo, Deurence Onyango, Michaela Hynie, Susan McGrath, and Melanie Baljko. 2018. Participatory design of DIY digital assistive technology in Western Kenya. In *Proceedings of the Second African Conference for Human Computer Interaction: Thriving Communities.* ACM, Windhoek Namibia, 1–11. doi:10.1145/3283458.3283478

[27] Aimi Hamraie and Kelly Fritsch. 2019. Crip Technoscience Manifesto. *Catalyst: Feminism, Theory, Technoscience* 5, 1 (April 2019), 1–33. doi:10.28968/cftt.v5i1.29607

[28] Christina Harrington, Sheena Erete, and Anne Marie Piper. 2019. Deconstructing Community-Based Collaborative Design: Towards More Equitable Participatory Design Engagements. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–25. doi:10.1145/3359318

[29] Christina N. Harrington, Aashaka Desai, Aaleyah Lewis, Sanika Moharana, Anne Spencer Ross, and Jennifer Mankoff. 2023. Working at the Intersection of Race, Disability and Accessibility. In *The 25th International ACM SIGACCESS Conference on Computers and Accessibility.* ACM, New York NY USA, 1–18. doi:10.1145/3597638.3608389

[30] Christina N Harrington, Aashaka Desai, Aaleyah Lewis, Sanika Moharana, Anne Spencer Ross, and Jennifer Mankoff. 2023. Working at the Intersection of Race, Disability and Accessibility. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '23).* Association for Computing Machinery, New York, NY, USA, 18 pages. doi:10.1145/3597638.3608389

[31] Roger A. Hart. 1992. *Children's Participation: From Tokenism to Citizenship.* Technical Report 4. UNICEF International Child Development Centre, Florence, Italy. https://www.unicef-irc.org/publications/pdf/childrens_participation.pdf

[32] Gillian R. Hayes. 2011. The Relationship of Action Research to Human-Computer Interaction. *ACM Transactions on Computer-Human Interaction (TOCHI)* 18, 3 (2011), 15:1–15:20. doi:10.1145/1993060.1993065

[33] Gillian R. Hayes. 2014. Knowing by Doing: Action Research as an Approach to HCI. In *Ways of Knowing in HCI*, Judith S. Olson and Wendy A. Kellogg (Eds.). Springer, New York, NY, 49–68. doi:10.1007/978-1-4939-0378-8_3

[34] Vanessa Heitplatz, Susanne Dirks, Bastian Pelka, Lukas Baumann, Ann Christin Schulz, Leevke Wilkens, and Jana York. 2025. Developing Technologies in an Ethical and Participatory Approach. Guidelines for Participatory Technology Design with People with Disabilities. *Interacting with Computers* (April 2025), iwaf012. doi:10.1093/iwc/iwaf012

[35] Liesbeth Huybrechts, Maurizio Teli, Mela Zuljevic, and Mela Bettega. 2020. Visions that change. Articulating the politics of participatory design. *CoDesign* 16, 1 (Jan. 2020), 3–16. doi:10.1080/15710882.2020.1728907

[36] Sins Invalid. 2017. Skin, tooth, and bone - the basis of movement is our people: A disability justice primer. *Reproductive health matters* 25, 50 (May 2017), 149–150. doi:10.1080/09688080.2017.1335999

[37] Jiwoong (joon) Jang, Patrick Carrington, and Andrew Begel. 2026. From autonomy to sovereignty - A new telos for socially assistive technology. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA, 1–19. doi:10.1145/3772318.3791585

[38] Alison Kafer. 2013. *Feminist, Queer, Crip.* Indiana University Press, Bloomington, MN.

[39] Finn Kensing and Jeanette Blomberg. 1998. Participatory Design: Issues and Concerns. *Computer Supported Cooperative Work (CSCW)* 7, 3–4 (1998), 167–185. doi:10.1023/A:1008689307411

[40] Finn Kensing and Joan Greenbaum. 2013. Heritage: Having a Say. In *Routledge International Handbook of Participatory Design*, Jesper Simonsen and Toni Robertson (Eds.). Routledge, New York, NY, USA, 21–36.

[41] Yasmine Kotturi, Julie Hui, TJ Johnson, Lutalo Sanifu, and Tawanna R. Dillahunt. 2024. Sustaining Community-Based Research in Computing: Lessons from Two Tech Capacity Building Initiatives for Local Businesses. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1, Article 178 (April 2024). doi:10.1145/3641017

[42] Cheuk Yin Phipson Lee, Zhuohao Zhang, Jaylin Herskovitz, JooYoung Seo, and Anhong Guo. 2022. CollabAlly: Accessible Collaboration Awareness in Document Editing. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22).* ACM, New York, NY, USA, Article 596, 17 pages. doi:10.1145/3491102.3517635

[43] Kelly Mack, Emma McDonnell, Dhruv Jain, Lucy Lu Wang, Jon E. Froehlich, and Leah Findlater. 2021. What Do We Mean by "Accessibility Research"? A Literature Survey of Accessibility Papers in CHI and ASSETS from 1994 to 2019. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21, Article 371).* Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3411764.3445412

[44] Jennifer Mankoff, Gillian R. Hayes, and Devva Kasnitz. 2010. Disability studies as a source of critical inquiry for the field of assistive technology. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility.* ACM, Orlando Florida USA, 3–10. doi:10.1145/1878803.1878807

[45] Laura Maye and Nicolai Brodersen Hansen. 2025. Involvement of Autistic Adults in the Participatory Design of Technology: A Scoping Review. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems.* ACM, Yokohama Japan, 1–17. doi:10.1145/3706598.3713961

[46] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model Cards for Model Reporting. *Proceedings of FAccT* (2019). doi:10.1145/3287560.3287596

[47] Aya Mouallem, Mirelys Mendez Pons, Ali Malik, Trini Rogando, Gene S-H Kim, Trisha Kulkarni, Charlene Chong, Danyang Fan, Shloke Nirav Patel, Lauren Aquino Shluzas, Helen L. Chen, and Sheri D. Sheppard. 2025. IncluSim: An Accessible Educational Electronic Circuit Simulator for Blind and Low-Vision Learners. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25).* Association for Computing Machinery, New York, NY, USA, Article 338, 18 pages. doi:10.1145/3706598.3713437

[48] Michael J. Muller. 1991. PICTIVE—An Exploration in Participatory Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91).* ACM, New Orleans, LA, USA, 225–231. doi:10.1145/108844.108896

[49] Michael J. Muller. 2002. Participatory Design: The Third Space in HCI. In *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, Julie A. Jacko and Andrew Sears (Eds.). Lawrence Erlbaum Associates, Mahwah, NJ, 1051–1068.

[50] Zachary Munn, Micah D. J. Peters, Cindy Stern, Catalin Tufanaru, Alexa McArthur, and Edoardo Aromataris. 2018. Systematic Review or Scoping Review? Guidance for Authors When Choosing Between a Systematic or Scoping Review Approach. *BMC Medical Research Methodology* 18, Article 143 (2018). doi:10.1186/s12874-018-0611-x

[51] Geoff Norman. 2010. Likert Scales, Levels of Measurement and the "Laws" of Statistics. *Advances in Health Sciences Education* 15, 5 (2010), 625–632. doi:10.1007/s10459-010-9222-y

[52] Mike Oliver. 2013. The social model of disability: thirty years on. *Disability & Society* 28, 7 (Oct. 2013), 1024–1026. doi:10.1080/09687599.2013.818773

[53] Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, et al. 2021. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* 372, Article n71 (2021). doi:10.1136/bmj.n71

[54] Hasso Plattner, Christoph Meinel, and Larry Leifer (Eds.). 2010. *Design Thinking: Understand – Improve – Apply.* Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-13757-0

[55] Xiang Qi and Junnan Yu. 2025. Participatory Design in Human-Computer Interaction: Cases, Characteristics, and Lessons. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama, Japan, 1–26. doi:10.1145/3706598.3713436

[56] Christian Quintero. 2022. A review: accessible technology through participatory design. *Disability and Rehabilitation: Assistive Technology* 17, 4 (May 2022), 369–375. doi:10.1080/17483107.2020.1785564

[57] Winston W. Royce. 1970. Managing the Development of Large Software Systems. In *Proceedings of IEEE WESCON*. IEEE, Los Angeles, CA, 1–9.

[58] Margarete Sandelowski, Corrine I. Voils, and George Knafl. 2009. On Quantitizing. *Journal of Mixed Methods Research* 3, 3 (2009), 208–222. doi:10.1177/1558689809334210

[59] Elizabeth B.-N. Sanders and Pieter Jan Stappers. 2008. Co-Creation and the New Landscapes of Design. *CoDesign* 4, 1 (2008), 5–18. doi:10.1080/15710880701875068

[60] David Williamson Shaffer. 2017. *Quantitative Ethnography.* Cathcart Press, Madison, WI, USA.

[61] Sidney Siegel and N. John Castellan. 1988. *Nonparametric Statistics for the Behavioral Sciences* (2nd ed.). McGraw-Hill, New York, NY, USA.

[62] Bridget Somekh. 1995. The Contribution of Action Research to Development in Social Endeavours: a position paper on action research methodology. *British Educational Research Journal* 21, 3 (June 1995), 339–355. doi:10.1080/0141192950210307

[63] Katta Spiel, Kathrin Gerling, Cynthia L Bennett, Emeline Brulé, Rua M Williams, Jennifer Rode, and Jennifer Mankoff. 2020. Nothing About Us Without Us: Investigating the Role of Critical Disability Studies in HCI. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3334480.3375150

[64] S. S. Stevens. 1946. On the Theory of Scales of Measurement. *Science* 103, 2684 (1946), 677–680. doi:10.1126/science.103.2684.677

[65] Oliver Suchanek, Janis Lena Meissner, Robin Angelini, and Katta Spiel. 2025. From Participation to Solidarity: A Case Study on Access of Maker Spaces from Deaf and Hearing Perspectives. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. ACM, New York, NY, USA, Article 292, 15 pages. doi:10.1145/3706598.3713202

[66] Jonathan Quetzal Tritter and Alison McCallum. 2006. The Snakes and Ladders of User Involvement: Moving Beyond Arnstein. *Health Policy* 76, 2 (2006), 156–168. doi:10.1016/j.healthpol.2005.05.008

[67] Victor Udoewa. 2022. An introduction to radical participatory design: decolonising participatory design processes. *Design Science* 8 (2022), e31. doi:10.1017/dsj.2022.24

[68] Peter Wacnik, Shanna R. Daly, and Aditi Verma. 2025. Participatory design: a systematic review and insights for future practice. *Design Science* 11 (2025), e21. doi:10.1017/dsj.2025.10009

[69] Rua Mae Williams, Louanne Boyd, and Juan E. Gilbert. 2023. Counterventions: a reparative reflection on interventionist HCI. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–11. doi:10.1145/3544548.3581480

[70] Michael T Wright, Hella von Unger, and Martina Block. 2010. Partizipation der Zielgruppe in der Gesundheitsförderung und Prävention. In *Partizipative Qualitätsentwicklung in der Gesundheitsförderung und Prävention*, Michael T Wright (Ed.). Hans Huber, Bern, 35–52.

[71] Zeynep Yildiz and Kathrin Gerling. 2026. Rethinking interdependence in HCI: A systematic literature review for understanding its use in accessibility studies. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–21. doi:10.1145/3772318.3790600