ABSTRACT

POLANCO JR., ALEXIS ANDRE (AJ). Exploring the Deaf and Hard of Hearing Inclusion Gap in Middle School Assessment Design, and a Multi-Disciplinary Framework Proposal to Address It. (Under the direction of Professor Tsai Lu Liu and Dr. M. Elen Deming).

The design of digital assessments that measure the student performance (e.g. math and languages) of Deaf or Hard-of-Hearing (DHH) middle-school aged students in the United States is supported by three constructs: accessibility, usability, and child-computer interaction (CCI). Each of these constructs represents relatively new fields of research. As a result, experts have claimed that research into DHH students is seemingly non-existent. This poses challenges for designers who are tasked with creating work in this interdisciplinary space, because they lack the resources necessary to ensure that they can best serve the needs of DHH students.

First, this study seeks to investigate why designers' resources are lacking for DHH assessment and, in particular, provides evidence of assessments that are inequitable to DHH students. Additionally, this study seeks to aid in the creation of Deaf and Hard-of-Hearing Middle School Assessment Design guidelines (DHH-MSAD) to empower designers to better serve DHH students. To that end, two rounds of interviews were conducted with subject-matter experts and design practitioners. The first interview asked participants to describe how they perform their work and to share any design guidelines they use. The second interview used a scale of semantic probes to ensure that all research participants understood the difference between guidelines in general and DHH-MSAD. The probes were as follows:

- **Q1.** When you hear the word "guidelines", what type of information would you expect to see?
- **Q2.** If you knew that "guidelines" were created for designers, would you expect to see different information? If so, what type of information?

- **Q3.** During the [previous] interview, we talked about empowering designers who create products for Deaf children. If you were reviewing “Deaf Design Guidelines”, would you expect to see different information? If so, what type of information?
- **Q4.** If you were responsible for creating guidelines for designing Deaf middle school assessments, what type of process would you follow?

Through their responses, the research participants confirmed the need for multi-disciplinary collaboration, revealed previously unspoken preconceptions (e.g. a designer taking for granted that they would be included in the DHH-MSAD formation process), and provided a research framework proposal for creating the next-generation of design guidelines for Deaf and Hard-of-Hearing students, as well as design guidelines for other under-represented populations.

Exploring the Deaf and Hard of Hearing Inclusion Gap in Middle School Assessment Design, and a Multi-Disciplinary Framework Proposal to Address It

by
Alexis Andre Polanco, Jr. (AJ)

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Design

Raleigh, North Carolina
2024

APPROVED BY:

Augustus Wendell	Dr. Nilda Cosco
Professor Tsai Lu Liu Co-chair of Advisory Committee	Dr. M. Elen Deming Co-chair of Advisory Committee

DEDICATION

I dedicate this work to my research participants, who believed in my research and volunteered their time and talent toward the advancement of digital accessibility for all. The accessibility, disability, and design communities owe you a debt of gratitude. We appreciate you.

BIOGRAPHY

Alexis Polanco Jr (AJ) is a design professional with over 15 years of experience in user experience design and product management. Professionally, he has executed strategic visions across education, pharmaceuticals, hospitality, financial technology, e-commerce, human capital management, sales, and marketing. He self-identifies as a person with disabilities and strives to advocate for the under-represented.

In the field of EdTech, his accomplishments include co-authoring and securing ETS's first NAEP Platform Development contract with the U.S government—valued at $101.7 million USD—and spearheading the release of the first open-source style guide for digital assessments (NAEP Style Guide). In the e-commerce space, his product portfolio was greater than $320 million in sales & call-in reduction, and he is a named co-inventor on a Verizon Wireless patent. He's currently a Director of UX Design at ADP, a global provider of human capital management.

AJ’s degrees include a Master of Business and Science (MBS) from Rutgers University and a Dual Bachelor of Arts and Science Digital Design & Architecture from NJIT, respectively. Soon after graduation from Rutgers, he was asked to lead the UX Design concentration for Rutgers MBS, where he discovered his passion for teaching design students. With a foot in both Academia and Professional practice, he came to learn that digital accessibility design training was in high demand and people with disabilities were continually being underserved.

To further his passions, AJ sought to achieve a life-long goal of a doctoral degree and chose to specialize in digital accessibility. Throughout his career, AJ has attended university while working, and NCSU’s Doctor of Design program allowed him to continue that trend. After completing his doctoral studies, AJ looks forward to finding new opportunities to advance the design profession and to uplift the next generation of design professionals.

ACKNOWLEDGMENTS

I would like to thank my co-chairs Professor Tsai Lu Liu and Dr. M. Elen Deming for their mentorship and guidance throughout this Doctor of Design journey. I also would like to extend gratitude to Augustus Wendell, Nilda Cosco, and Percy Hooper for believing in my research and allowing me to see my studies to completion. Throughout this journey, I have experienced tremendous highs and lows and throughout it all, my doctoral committee was always supportive and understanding.

I would also like to give a shout-out to my fellow Cohort 1 Doctor of Design students! It was scary and amazing to be the pioneer class for a new academic program, and every one of you are awesome human beings. I hope this dissertation motivates more of you all to complete your DDes journey! As always, I'm here to help in whatever way I can!

I want to thank my parents, Alex and Linder Polanco, who always strove to provide the best opportunities for me and made it possible for me to receive my dual bachelor degree from NJIT; my brother and best man, Andrew Polanco, the Pipe Gang (Leandro, Edwin, Erin, Dennis, Matt, and Mateusz), and the Homies who supported me while I was juggling getting married while trying to complete my Doctor of Design degree.

Lastly, I want to express immense gratitude and boundless love to my wife Jenifer L. Elizondo-Polanco, who has been lockstep with me throughout this entire doctoral journey and without whom this dissertation would not be possible. We experienced incredible highs and endured immense sorrows over the past few years (rest in peace, Dusty), but we did it all together and there is no-one else I would have taken this journey with. I love you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCING THE RESEARCH STUDY

1. Research Background

Learners who are Deaf or Hard-of-Hearing (DHH) are among the least understood student groups, despite their prevalence in the United States. For context on the scope of hearing loss in America, it is estimated that 14.3% of all Americans aged 12 and older have some form of hearing loss (Goman 2016). Unfortunately, research on how to measure the academic performance of these learners is seemingly non-existent, even though the number of DHH students in American schools is increasing each year (Pizzo and Chilvers 2016). This poses a challenge for designers of digital assessments who must create fair and equitable products for DHH students. This challenge is further compounded by the fact that the two of the most popular design heuristics and principles, Jacob Nielsen's 10 Usability Heuristics for User Interface Design and North Carolina State University's Principles of Universal Design, provide little to no DHH-specific recommendations (Nielsen 1994) (NC State University 1997).

Supporting the idea that the measurement of DHH performance is misunderstood, digitally based assessment and educational support for students with disabilities are both recent developments. For reference, the National Assessment of Educational Progress (NAEP) is the largest nationally representative assessment of American students' proficiency in various subjects. First administered in 1969, it took until 1998 for special needs accommodations to be authorized, and until 2017 for Main NAEP to fully transition to a digitally based assessment (NCES 2019). As a result, much of the literature surrounding the performance of Deaf and Hard-of-Hearing students comes back to one or two underlying research citations, with the most prominent being the Stanford Achievement Test (Traxler 2000). From this citation comes the oft-

quoted statistic that, at age 14, all but the top 20% performing DHH students place below a grade-4 reading level.

While there may be some validity to this claim, the author believes the disparity in scores between hearing students and DHH students is more indicative of failures with the testing instruments than the failures of the students. For example, assessments are often created in English, which may not align with the language DHH students are used to, i.e. a) English may not be the DHH student's primary language; b) DHH students may be accustomed to receiving classroom instruction via American Sign Language, not English; and c) reading comprehension is not directly observed in assessments; scores are generated from the result of students' interaction with an assessment product, not the process of reading itself (Qi and Mitchell 2012; Lollis and LaSasso 2008).

Regarding the last point, recent investigations have linked e-learning with heightened fatigue and lower test sources amongst DHH students due to these digital interfaces requiring greater levels of multi-focal attention (Rodrigues, et al. 2022). This issue of multi-focal attention is most readily exemplified by the NAEP Tutorial.

Case Study of the NAEP Tutorial

The NAEP tutorial exists as a student's first entry point into the NAEP. It is a semi-autonomous sequence of instructional content, interspersed with interactive elements that guide the student test-taker into understanding how the assessment's user interface (UI) works. As the UX Lead for the NAEP Platform Development (NPD) contract from 2019-2022, the author was tasked with making the existing NAEP tutorials accessible on the web. During this timeframe,

the 2019 NAEP assessment was administered nationally across the United States, and the author was kept apprised of real-time updates on how students were faring in the field.

Unfortunately, it was quickly determined in the field that the NAEP tutorial was inaccessible to Deaf and Hard of Hearing students who were permitted the American Sign Language (ASL) accommodation. For reference, this accommodation is an approved deviation from the assessment experience wherein an ASL interpreter is provided to a student to assist them as they take the NAEP assessment. In instances when an assessment's goal is to measure knowledge, it is reasonable to expect this accommodation (National Deaf Center 2019).

The reason why the NAEP tutorial was deemed inaccessible is that portions of the assessment consisted of auto-playing content that could not be paused. To illustrate the issue, consider the following scenario:

You are a Deaf student in the fourth grade. You are given a Surface Pro laptop, which may be unlike any other laptop you have used before. To better acclimate you to the computer hardware, its many functions, and to better understand the user interface of the NAEP assessment, a tutorial is provided. Upon the start of the tutorial, images rapidly move across the screen; although captions at the bottom of the screen describe what is being verbally communicated to hearing students, you cannot hear the audio yourself.

You look to your ASL interpreter who translates what is being said by the computer. Your interpreter translates the audio to an approximation of, "Sometimes you will be asked to do something in the tutorial. The arrow will show you where to tap." You then look back to the Surface Pro monitor, only to see that there is no arrow on the screen. You then attempt to communicate to the ASL interpreter that you don't see an arrow, but it is too late. At this point, an illustrative hand taps the "Next" button in the system toolbar, and the tutorial advances to the

next instructional module. Now you're confused, frustrated, and you haven't even gotten to the graded portion of the NAEP assessment yet!

This is a very real situation that affected real students. At this point, you should be asking yourself, why did this happen?

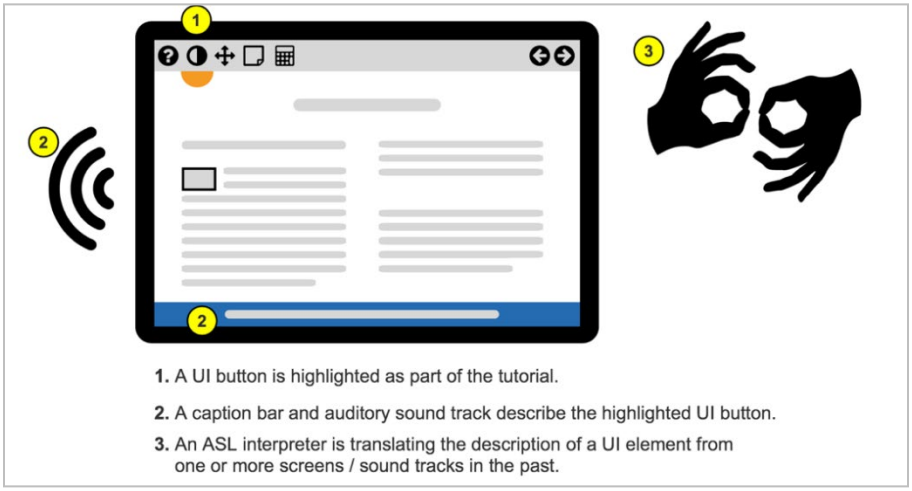

Figure 1.1 - An illustration of the multi-focal challenge posed by the NAEP Tutorial[1], as experienced by Deaf and Hard of Hearing students with the ASL interpreter accommodation.

Unfortunately, because the NAEP assessment is timed, students are only afforded a short amount of time to complete the tutorial modules. To keep the assessment orderly and regimented, the computer will automatically progress the student through portions of the assessment once a time limit has been exceeded.

This scenario raises the question of why this issue had never been observed before, and/or why no-one had considered this possibility before inflicting this frustration on DHH child students. The short answer is that the tutorial was not tested with Deaf or Hard of Hearing

[1] The NAEP Tutorial on the Web can be accessed at https://enaep.cotw-ng.naep.ed.gov/totw/2024/English.html

students before 2019 NAEP was administered to the field. Additionally, Deaf or Hard of Hearing professionals were not engaged in the tutorial creation process.

To remedy the situation, the author and other NAEP Platform Development contractors recruited teachers of DHH students to provide their expertise as part of user experience (UX) interviews via Zoom video conferences. Very quickly into the first user interview, a teacher asked the UX researcher who was facilitating the user test to mute the audio so that the meeting participants could experience the tutorial as a Deaf student. (Note that during this first user interview, the author served as a note-taker to assist another UX researcher.) In hindsight, the NAEP Development contractors felt as though the suggestion of muting the audio to simulate the Deaf experience was obvious—but surprisingly, no-one had considered that possibility before. Had a teacher of DHH students or DHH students been consulted earlier in the design process, the NAEP tutorial inaccessibility issue in the field could have been avoided altogether.

This experience motivates this research study to explore ways to prevent designers from facing similar issues when designing assessments for DHH students in the future. In other words, the general research problem is as follows: How can UX designers have earlier involvement and greater input in the design of assessments for DHH students and other underrepresented or underserved learners with disabilities? To lay the groundwork for this study and shape its more precise research questions, the author sought to review the theoretical landscape surrounding DHH digital assessment design.

2. Theoretical Context of Study

A shift needs to occur wherein digital assessments need to be increasingly recontextualized as user interface problems. This recontextualization will lead to the

development of more equitable design strategies for DHH students via the consolidation and integration of related research disciplines. Furthermore, this author acknowledges the problem of limited research on the academic performance of DHH learners that was first identified by Pizzo & Chilvers in 2016. The current dissertation seeks to reframe that problem as a lack of much-needed design resources for designers. To aid in the forming of a research question, the author sought to identify literature that relates to the design of digital assessments and compile this literature as a resource for future designers of DHH digital assessments. The topics that emerged are as follows: Measuring student performance; Test Score Validity; Accessibility & Usability; Child-computer Interaction.

2.1 Measuring Student Performance

What is Being Measured? The score received on the computerized test is predicated on the constructs being measured. The term "construct" is defined as "an abstract image, idea, or theory, especially a complex one, formed from a number of simpler observable elements" (NCES 2020). For example, one theoretical reading assessment may seek to measure a command of English vocabulary whereas another assessment may seek to measure proficiency with grammar and punctuation.

To understand what the measurement goals of an assessment are, individuals should seek out the underlying "frameworks" of the test they are using. Let us use NAEP again as an example. The Reading Assessment Framework for the 2022 and 2024 NAEP breaks down reading into the following three characteristic parts:

> **Understanding written text:** Readers attend to ideas and content in a text by locating and recalling information and by making inferences needed for literal comprehension of the text. In doing so, readers draw on their fundamental skills for decoding printed words and accessing their vocabulary knowledge.

> **Developing and interpreting meaning:** Readers integrate the sense they have made of the text with their knowledge of other texts and with their outside experience. They use increasingly complex inferencing skills to comprehend information implied by a text. As appropriate, readers revise their sense of the text as they encounter additional information or ideas.

Using meaning: Readers draw on the ideas and information they have acquired from text to meet a particular purpose or situational need. The use of text may be as straightforward as knowing the time when a train will leave a particular station, or it may involve more complex behaviors such as analyzing how an author developed a character's motivation or evaluating the quality of evidence presented in an argument. (National Assessment Governing Board 2022)

Returning to the topic of DHH student proficiency, there is no explicit mention in the 2022 and 2024 NAEP Framework for how these constructs apply to DHH students. The framework does, however, note that readers' background and experience influence reading performance. To better understand student performance, it is important to understand the psychometric principles that inform assessment scores.

How are scores measured? There are two primary methods of measuring student performance in an assessment: Classical Test Theory (CTT) and Item Response Theory (IRT).

Classical Test Theory. The first method, Classical Test Theory, was developed to determine measurement error values to correct test scores. It seeks to explain why the same construct could be measured multiple times with different results (Steyer 2001, 1955). The primary concepts are "test observed" (X), "true score" (T), and "error score." It is best explained by the following formula, assuming that "true score and error scores are independent" (Magno 2009, 1).

$$X = T + E$$

In CTT, a true score represents the average score a student would receive if they completed a test an infinite number of times. As a result, this approach is very dependent on the size of pretesting test-taker samples. Additionally, to properly calculate an assessment's true score in the CTT model, the same question must be administered each time to appropriately determine the error score (Zanon et al. 2016).

Item Response Theory. The second method, Item Response Theory, was formed in response to the CTT limitation of needing to deliver the same test items for analysis. Its key concepts are a person's "latent ability" and an "item characteristic curve (ICC)" (Yang and Solon 2014, "Basic measurement properties for IRT"). Latent ability refers to a construct being measured, and the item characteristic curve represents the probability of receiving a correct score based on a test-taker's latent ability. Item response theory is founded upon a few key assumptions— assumptions which have caused it to be the more prevalent model in educational assessment.

1) "Monotonicity: as the latent trait increases, the probability of a correct answer will always increase.
2) Unidimensionality: there is one dominant trait that is being measured by an item, and it is the primary factor in all item scores.
3) Local independence: scores received on different items have no statistical bearing on one another; i.e. they are mutually independent.
4) Invariance: Item characteristics can be estimated from any point on the ICC." (Yang and Solon 2014, "Basic assumptions in IRT")

While there are many possible Item Characteristic Curves, the one-, two-, and three-parameter IRT models are among the most referenced and utilized. Their formulas are represented below, starting with the one-parameter IRT model.

$$P_i(\theta) = \frac{1}{1 + e^{-(\theta - b_i)}}$$

$$P_i(\theta) = \frac{1}{1 + e^{-D_{ai}(\theta - b_i)}}$$

$$P_i(\theta) = c_i + (1 - c_i)\frac{1}{1 + e^{-D_{ai}(\theta - b_i)}}$$

The IRT variables are defined as follows:

- $P_i(\theta)$ is the probability of a current response for the i[th] item;
- b_i is the difficulty parameter for the i[th] item;
- a_i is the discrimination parameter for the i[th] item;
- c_i is the guessing parameter for the i[th] item;
- θ is the ability level (i.e. latent ability);
- D represents a scaling factor. (Awopeju et al 2017, p.199)

To further clarify the variables, the following semantic definitions shall be provided, and representative/interactive graphs will be linked as well.

- The difficulty parameter (b_i) represents the median probability of receiving a correct answer. It causes the center of the graph to shift left or right on the x-axis. (**Figure 1.1**)

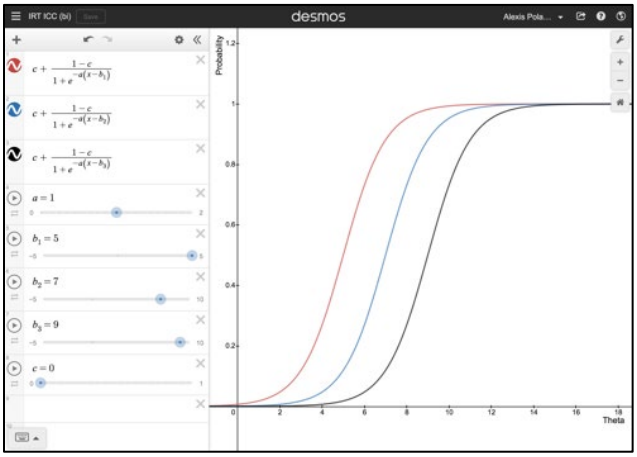

Figure 1.2 - ICC where only the difficulty variable is variable.

Interactive Graph: https://www.desmos.com/calculator/wm3cfpnlgd

- The discrimination factor (a_i) refers to the rate of change between ability and score. An ICC with a high discriminating factor will only see correct answers at very high ability levels. This factor must never be negative, or else it would mean that scores will decrease with rising ability level (represented in black below (**Figure 1.2**).

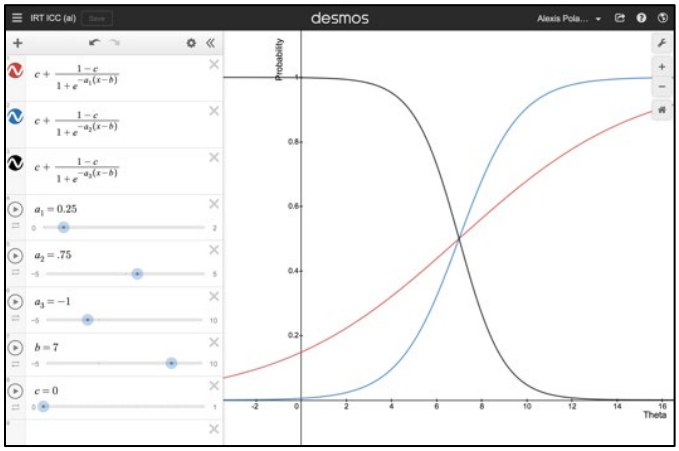

Figure 1.3 - ICC where only the discrimination factor is variable.

https://www.desmos.com/calculator/iqu9nb9l2m

- The guessing parameter is a number between 0 and 1, which represents the statistical probability of receiving a correct score, independent of latent ability; i.e. 1.0 for a question with a single answer choice, 0.5 for a question with two answer choices, $0.3\overline{3}$ for two answer choices, etc (**Figure 1.3**).

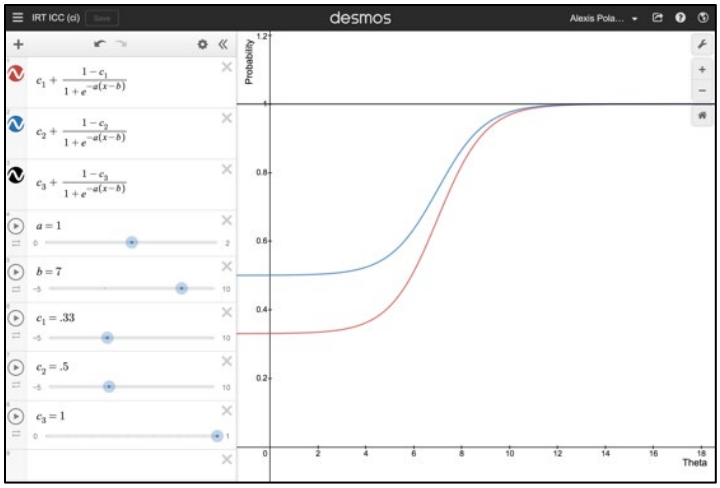

Figure 1.4 - ICC where only the guessing parameter is variable

https://www.desmos.com/calculator/oe66lf9jcf

- The scaling factor represents a numerical constant, typically a value of 1.7, which is used for convenience to make item analysis easier (Camilli 2017).

Unfortunately, even in the three-parameter IRT model, the author was unable to locate research that linked the efficacy of IRT to the performance of DHH students, nor students with other disabilities. Based on the NAEP case study described previously, the author argues that these measurement theories were developed without consideration for students with disabilities. To explore this argument further, it is necessary to explore the topic of test score validity.

2.2 Test Score Validity

In addition to the CTT and IRT theories, there are concepts of construct validity. Construct validity is defined as "the degree to which inferences can legitimately be made from the operationalizations in your study to the theoretical constructs on which those operationalizations were based" (Trochim n.d., para. 1). In other words, there is an understanding

that there may be discrepancies between actual cognitive processes that occur in a test-taker's head and observed outcomes, independent of test item factors like an item's difficulty parameter, etc. a la IRT.

To ensure validity, construct-irrelevant factors must be minimized, to as great an extent as possible. The most common example of a construct-irrelevant factor is English Language proficiency when other constructs are being measured; e.g. algebra proficiency on a mathematics assessment. Other examples of known construct irrelevant factors are the following:

- "Unfamiliar vocabulary that is not related to the construct
- Cultural references or idiomatic expressions (such as "being on the ball") that are not equally familiar to all students
- Syntax that may be confusing or ambiguous (such as negatives or double negatives)
- Low-frequency, long, or complex words and long sentences
- Sentence structure that may be confusing or difficult to follow (such as passive voice or sentences with multiple clauses)
- Syntax or vocabulary that is above the test's target grade level" (Young 2009)

Therefore, the literature suggests that the problem of designing fair and equitable digital assessments for DHH students is less a problem with psychometric assessment theories (like IRT) and are, instead, matters of test score validity. Using the above list of construct irrelevant factors, assessment designers can be conditioned to be on the lookout for things like user-interface labels and system tooltips that use language that is unfamiliar to DHH students and/or might exceed their reading grade level. If this occurs, a designer may unintentionally negatively impact the validity of DHH students' assessment measurement outcomes. Building upon this new insight, the author explores the impacts of assessment design and delivery on assessment measurement outcomes.

Effect of Digitally Based Assessment on Test Scores

Studies into the outcomes of computerized testing can focus on two aspects of student performance: motivation and achievement (Gbollie and Keamu 2017).

Motivation. In Illinois, research was conducted on grade-4 test takers to determine if one-to-one technology (i.e. one laptop to one student) had an impact on student motivation to learn. Motivation was correlated with student attendance. The result was that technology was not a factor (Harris et al. 2016). Unfortunately, the researchers admitted that fourth graders have less autonomy than middle- and high-school students; e.g. grade-4 students are more likely to have their actions decided by their parents, and so the implied motivation attributable to attendance may be due to parental factors vs solely attributable to the child students. Thus, the general applicability of this outcome is questionable.

Another study, conducted by the University of New England, provided another perspective; students self-identified as being either very motivated or motivated by technology vs not motivated/unmotivated (Francis 2017). This study consisted of surveying 348 students of the New England Charter School, with a response rate of 27.3%; the results showed a ratio of 73:21 (technologically motivated vs. unmotivated). Caveats to this study include the low response rates, as well as the fact that this study may not be representative of student motivations in non-charter schools (e.g. public schools).

Achievement. When attempting to determine whether computerized assessment had a construct-irrelevant result, there have been studies and administrations wherein different samples of students were given both paper-based assessments (PBAs) and item-equivalent digitally based assessments (DBAs). The most representative example of this is the 2017 Digital Transition of the NAEP assessment.

This study represented a re-administration of items from the 2015 PBA via a 2017 PBA to 40,000 students, and administration of the same items as DBA to 150,000 students as a 2017 DBA. The results of this study are shown below and demonstrate close congruence between PBA and DBA with respect to mean item scores (**Figure 1.4**).

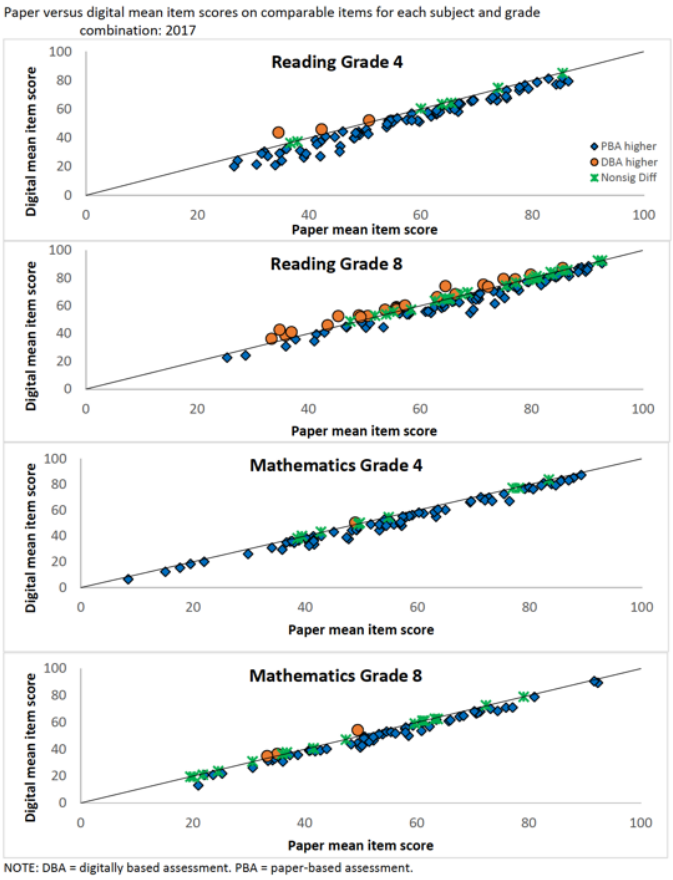

Figure 1.5 - Results from 2017 NAEP Digital Transition. Reprinted from "2017 NAEP Transition to Digitally Based Assessments" by Paul Jewsbury, Robert Finnegan, Nuo Xi, Yue Jia, Keith Rust, Samantha Burg.

Prior to finding this study from Jewsbury et. al, the author assumed that there would be a significant difference in measurement outcomes between paper-based assessment and digital-based assessment. Given that this was not the case, the author was then able to shift his

investigation into matters of digital accessibility and usability without fear that PBA studies would need to be cross-referenced throughout this study.

2.3 Accessibility & Usability

When it comes to the design of digital assessments, the relationship between accessibility and usability can be ambiguous. It is understood that both disciplines emerge from a desire to create a product that emphasizes ease of use via human-centered design (HCD). Usability is defined by ISO standard 9241-11:2018 as "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO 2018). Among the most cited papers on usability is Jakob Nielsen's "10 Usability Heuristics for User Interface Design" (henceforth, Nielsen-10). As of June 2024, Google Scholar reports that the Nielsen-10 have been academically cited over 3200 times across two listings. As drawn from Nielsen (1994), the ten heuristics are as follows:

> [H1] Visibility of system status
> The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
>
> [H2] Match between system and the real world
> The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
>
> [H3] User control and freedom
> Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
>
> [H4] Consistency and standards
> Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

> [H5] Error prevention
> Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.
>
> [H6] Recognition rather than recall
> Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
>
> [H7] Flexibility and efficiency of use
> Accelerators — unseen by the novice user — may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
>
> [H8] Aesthetic and minimalist design
> Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
>
> [H9] Help users recognize, diagnose, and recover from errors
> Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
>
> [H10] Help and documentation
> Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

Unlike usability, which could derive its definition from decades-old sources like Nielsen-10 and ISO 9241-11:2018, accessibility had no source for a standardized definition until very recently; as such, it was often defined by its problem set. For example, problems that affected disabled people exclusively were “pure accessibility” problems. (Petrie and Kheir 2007, p. 398) Conversely, problems that affected non-disabled people were seen as pure usability problems. For these pure accessibility problems, designers and software developers would look to the World Wide Web Consortium (W3C) Web Content Accessibility Guidelines (WCAG) to ensure

their web products were accessible to people with disabilities (W3C 2018). The overarching principles of WCAG are as follows:

> [P1] Perceivable - Information and user interface components must be presentable to users in ways they can perceive.
> - This means that users must be able to perceive the information being presented (it can't be invisible to all of their senses)
>
> [P2] Operable - User interface components and navigation must be operable.
> - This means that users must be able to operate the interface (the interface cannot require interaction that a user cannot perform)
>
> [P3] Understandable - Information and the operation of user interface must be understandable.
> - This means that users must be able to understand the information as well as the operation of the user interface (the content or operation cannot be beyond their understanding)
>
> [P4] Robust - Content must be robust enough that it can be interpreted reliably by a wide variety of user agents, including assistive technologies.
> - This means that users must be able to access the content as technologies advance (as technologies and user agents evolve, the content should remain accessible)

Somewhere between these pure usability and pure accessibility problems sits Universal Design, or “the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design” (NC State University 1997, "Universal Design" section, para. 1). The follow-up question is whether these two disciplines overlap, or if they are completely distinct. To answer this question, researchers have sought to form a correlation between Nielsen-10 and WCAG. The last work on this demonstrated a strong relationship between Nielsen’s H1, H2, H3, and H6 and WCAG’s P1, P2, P3, P4. This in turn sets up a relationship between accessibility and usability as complements of one another and is represented below (**Figure 1.5**) (Casare et al. 2016).

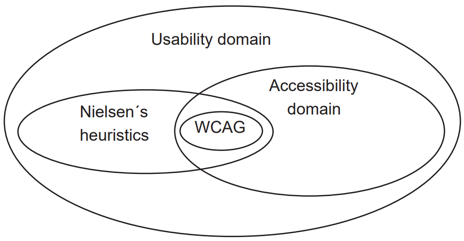

Figure 1.6 - Association between usability and accessibility. Reprinted from "Usability heuristics and accessibility guidelines: a comparison of heuristic evaluation and WCAG", by Casare et. al, 2016, ACM, p. 213-215

One potential criticism of the analysis done by Casare et al. is their failure to look beyond Petrie & Kheir's statement that there was no standard definition for accessibility. Given that the research aim of Cesare et al was to link guidelines across the usability and accessibility disciplines, those authors may have felt that studying the definition of accessibility was outside their research scope. Fortunately, the International Electrotechnical Commission (IEC), the International Organization for Standards (ISO), and the International Telecommunication Union (ITU) jointly published the *IEC/ISO/ITU Policy on Standardization and accessibility* in 2014.

This policy stated that the core objective of accessible design was to "[ensure] that products, systems, services, environments and facilities can be used by persons from a population with the widest range of characteristics and abilities" (IEC, ISO, ITU 2014, p. 2). This definition for accessibility is very similar to that described by the Center for Universal Design at North Carolina State University. In this way, the HCD field has converged to make the terms accessibility and universal design synonymous, and their objectives equal. This shift appears to have been largely pushed by the U.S. government, as is evidenced by Section508.gov (the site that is tasked with assisting federal agencies with accessibility compliance) choosing to

lump these two terms together (Section508.gov 2023). As such, throughout this current research study, when speaking on issues affecting DHH students, the terms of accessibility and universal design will be used interchangeably.

2.4 Child-computer Interaction

Child-computer interaction (CCI) is a sub-discipline that emerged from the human-computer interaction field. Human-computer interaction is defined by the Association for Computing Machinery (ACM) as "a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them" (Hewett et al. 1992, p. 5). At its core, CCI's definition seeks to replace the word "human" with "children" to give it a clearer and more specific focus. Unfortunately, the definition of child is still being defined. During a CCI Special Interest Group (SIG), ACM members sought to expand the scope of CCI literature to include children who are five years or younger (Hourcade et al. 2016). In *The Nature of Child Computer Interaction*, Read & Bekker demonstrates that most work in child-computer interaction focuses on children between the ages of five and eleven, with an understanding that ages three to sixteen are acceptable ranges for CCI research (Read and Bekker 2011, p. 3). In this same conference preceding, it is acknowledged that the CCI community is actively defining itself by reassessing the scope of their concerns.

While it is acknowledged that teenagers represent a valid subset of CCI, it has been documented that research into this population group has been on the decline. (Giannakos et al. 2020). One of the reasons posed for this is that authors may have failed to add sufficient metadata and keywords to link their studies with the CCI discipline. That said, teenagers

represent a demographic that will lend itself to better test results when conducting usability studies, due to their greater ability to verbalize concerns. Literature backing this assertion includes, "*On the assessment of usability testing methods for children."* In Section 4, titled "A framework for comparing usability testing methods of children," the following three dimensions are provided (Markopoulos and Bekker, 2003):

- **Assessment Criteria**: What is being measured (e.g. how many UX problems exist)
- **UTM Characteristics:** User testing operational methods (e.g. number of users, tasks)
- **Children Characteristics:** How children influence the usability test (e.g. a child's inability to verbalize concerns to adults)

The last characteristic of Markopoulos & Bekker's framework for user testing offered a new critical lens for analyzing digital assessments. In much the same way that children's abilities may influence the results of a usability test, a DHH child's inability to discern the system requirements of a DBA would likely have detrimental effects on their ability to successfully navigate the digital assessment.

Considerations when Designing Digital Assessments for Children. In instances where a child is expected to take a digital assessment without an adult, a child would be disadvantaged because they would be expected to both configure their computer accordingly and successfully complete the assessment without adult assistance, while potentially in a high-stakes testing situation. These disadvantages affect both Deaf and normative hearing students alike. The approach to resolving these computer-only disadvantages falls upon the designers of digital assessments to ensure that the user interfaces of the tests they design are (1) intuitive and/or can be easily taught on assessment day, and (2) that any computer issues are identified and resolved before the scored portion of the test begins.

Furthermore, to ensure that a digital assessment's interface can be readily understood on test day, there are a few different approaches that can be taken to compensate. One approach is to provide practice items prior to test day. This allows students to familiarize themselves with the assessment UI so that they are not disadvantaged on test day by any unfamiliarity they may have, either due to lack of experience with comparable UI, and/or any novel UI that is introduced as part of the measured constructs of the assessment. An example of one of these pre-assessment practice items can be found on the SmarterBalanced Consortium's website: https://smarterbalanced.org/our-system/students-and-families/samples/. Another approach to helping test-takers is a front-loaded tutorial, wherein the assessment's UI elements are individually called out and introduced to the student. This is the technique that NAEP employs.

On the topic of computer issue identification and issue resolution, this activity is traditionally done by adults on test day by Supervisors and Assessment Administrators. In recent years, there has been the creation of self-guided device-readiness checks that have been posted online for digital assessments, prior to the scored portion of the test. For example, the assessment company Questar provides a System Scan utility to ensure that your computer technically meets minimum system requirements: http://www.questarai.com/readiness/system-scan/. Additionally, they recommend third-party software like SpeedTest.net to determine download and upload speeds as part of its Testing Readiness Check[2] to ensure your internet connection does not disadvantage a student's assessment experience. While Questar itself administers K-8 (Kindergarten to Grade 8) assessments, it is not clear whether these readiness check solutions were designed with children in mind, in a self-service capacity.

[2] Testing Readiness Check - http://www.questarai.com/readiness/test-readiness/

Using a Readability Test Tool[3] provided by WebFX, it looks like the System Scan page has an "average grade level of about 8," which demonstrates that it may not be suitable as a computer-only solution for most middle-school-aged children. Understanding this, designers of DHH assessments may wish to consider lowering the reading level of any computer-only testing UI that is put in front of children. This consideration aligns with the recommendation proposed by the author in the previous section on Test Score Validity (Section 2.2).

3. Research Problem & Study Purpose

The literature review conducted in Section 2 suggests that the design of digital assessments for Deaf and Hard of Hearing students is inherently multi-disciplinary, as no single discipline can wholly encompass problem space. As such, the author identified the following research problem as one to be tackled by this study: Designers are being tasked today with creating work in the interdisciplinary space without the resources they need to best serve the needs of DHH students. Therefore, the purpose of this study is to understand how subject-matter experts (SMEs) and designers of products for Deaf and Hard-of-Hearing students perform their work; how they meet the needs of their audience; and whether there is an incongruence between their work process versus what they believe will *actually* serve their audience. Based on experience working with the NAEP assessment, the author surmises that business and operational constraints may not permit designers to follow their idealized work process. That is, companies may avoid user testing with DHH children because of budget and/or time constraints.

To address the research problem and study purpose, the author consulted experts in the related disciplines. These disciplines include user experience design, assessment design, child

[3] Readability Test Tool search query against Questar System Scan Page - https://www.webfx.com/tools/readable/check.php?tab=Test+By+Url&uri=http%3A%2F%2Fwww.questarai.com%2Freadiness%2Fsystem-scan%2F

computer interaction, digital accessibility, and Deaf-and-Hard-of-Hearing education. These experts participated in one-on-one interviews via online conference calls. Given that designers of different disciplines think differently about their design problems (e.g. user experience designers vs visual designers), the author expected that experts from different fields may conduct themselves differently (Kirsh 2019). For example, during a research interview, a user experience designer may share their computer screen to show how their designs relate to Nielsen's 10 Heuristics. Alternatively, a Deaf or Hard-of-Hearing subject-matter expert may request that the author share an exemplar assessment product to better contextualize their feedback on how it could be better designed and/or how a designer could better empathize with the DHH community.

The ideal outcome for this research study is to identify guidelines that can be instrumental in solving many of the user-centered design problems that developers and designers of DHH assessments face. As such, the overarching research questions are: "What guidelines exist today to support the needs of designers developing assessment experiences for middle school-aged Deaf of Hard-of-Hearing students? If none exist, how might new or alternative guidelines emerge?" For convenience sake, propositional guidelines will be henceforth referred to as Deaf and Hard-of-Hearing Middle-school Assessment Design Guidelines (or DHH-MSAD Guidelines).

4. Theoretical Framework and Research Significance

Using Casare et al's association between usability and accessibility (Section 2.3), it is useful to visualize where DHH-MSAD is situated when mapped alongside Nielsen-10, PUD, WCAG, and CCI (**Figure 1.7**). This study seeks to enhance the theoretical framework by

acknowledging the pure accessibility problems and pure usability problems that were described by Petrie & Kheir (2007) in Section 2.3. As such, the proposed DHH-MSAD Guidelines seek to do two things: (1) emphasize CCI-characteristics of PUD, and (2) expand the scope of Nielsen-10 to address DHH middle school students' pure accessibility needs.

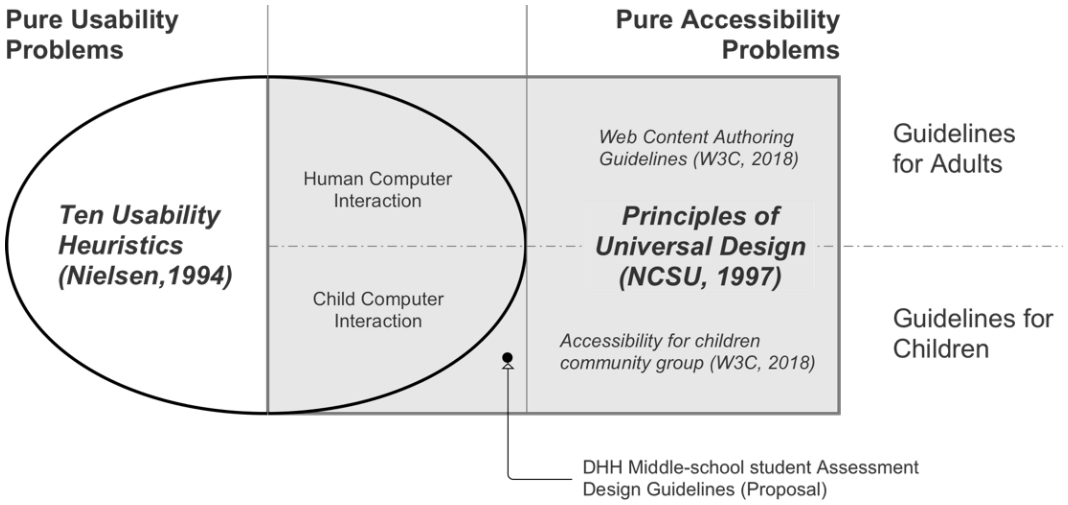

Figure 1.7 - Demonstrating the theoretical framework of the DHH Middle-school student Assessment Design Guidelines (Proposal)

Research Significance

This study is significant because the DHH population remains one of the least understood groups in the United States even as this population group grows year over year, with the US CDC reporting a 239% increase in the number of babies born Deaf or Hard-of-Hearing who were identified early (CDC 2022). New knowledge in this space will ensure that DHH children have a better chance at receiving equitable educational assessment, which will in turn allow for a better quality of life as these assessments can have federal funding implications for school districts, as is exemplified by legislation like the No Child Left Behind Act of 2001 and the Every Student Succeeds Act of 2015.

Additionally, research on Deaf and Hard-of-Hearing students' assessment is seemingly non-existent according to the American Annals of the Deaf (Pizzo and Chilvers 2016, p. 56). This claim is corroborated by the author's first-person investigation into potential user usability subcontractors as part of his NAEP responsibilities as the UX Lead of the NPD Contract from 2019-2022. Every vendor that was solicited by the author said that digital accessibility was either a) something they did not support, or b) a product roadmap item that was expected to be tackled at an undetermined date in the future.

Lastly, this study is significant because it ties into the larger problem of inaccessible digital content. As will be expanded upon in Chapter 3 of this dissertation, in 2023, over 96% of the top million websites globally triggered accessibility errors when checked against the 2.1 version of the Web Content Accessibility Guidelines (WCAG 2.1). To make matters worse, there is documentation proving that schools are ill-equipped and ill-informed about digital accessibility. This is supported by a study conducted in 2019 which discovered that products used in U.S. schools are often acquired by faculty members, without coordination of Information Technology (IT) staff that specialize in accessibility (McKenzie 2019). McKenzie's anecdote references the software procurement process employed by universities, like Iowa State University. The inference made here by the author is that middle schools follow a similar procurement process given that McKenzie referenced products from Pearson and McGraw-Hill, which are companies that work in both the K-12 and post-secondary education spaces. McKenzie goes on to say that the blind faith that schools have placed in these education companies have resulted in discrimination lawsuits, wherein Federal Courts have ruled that these schools have breached the Americans with Disabilities Act (ADA) of 1990. The expectation of this study is not that it will change software procurement processes in schools. Instead, the assumption is that

this research will empower designers to usher in the proliferation of more accessible products in the marketplace, thereby eliminating the need for lawsuits to bolster the equitability of DHH student assessments. Through new knowledge contributions like this paper, researchers can help make digital products more accessible for people with disabilities.

5. Design of the Dissertation

This dissertation is composed of five chapters as follows:

Chapter 1 describes the study rationale, summarizes the background and theoretical context, and the research objectives.

Chapters 2, 3, and 4 comprise three peer-reviewed, accepted, and published papers which expand on the elements of Chapter 1.

- **[Paper 1] Chapter 2:** Uncovering an Inclusion Gap in the Design of Digital Assessments for Middle school-aged Deaf and Hard-of-Hearing Students in the United States
- **[Paper 2] Chapter 3**: UX Design for Deaf children: How is it done now? Should it change?
- **[Paper 3] Chapter 4:** Multidisciplinary Framework for Creating the Next-generation of Human-centered Design Guidelines

Chapter 5 concludes the research study with a discussion that summarizes findings, provides the author's analysis and interpretation of the findings, describes limitations of the research study, and suggests future research opportunities revealed through the dissertation process.

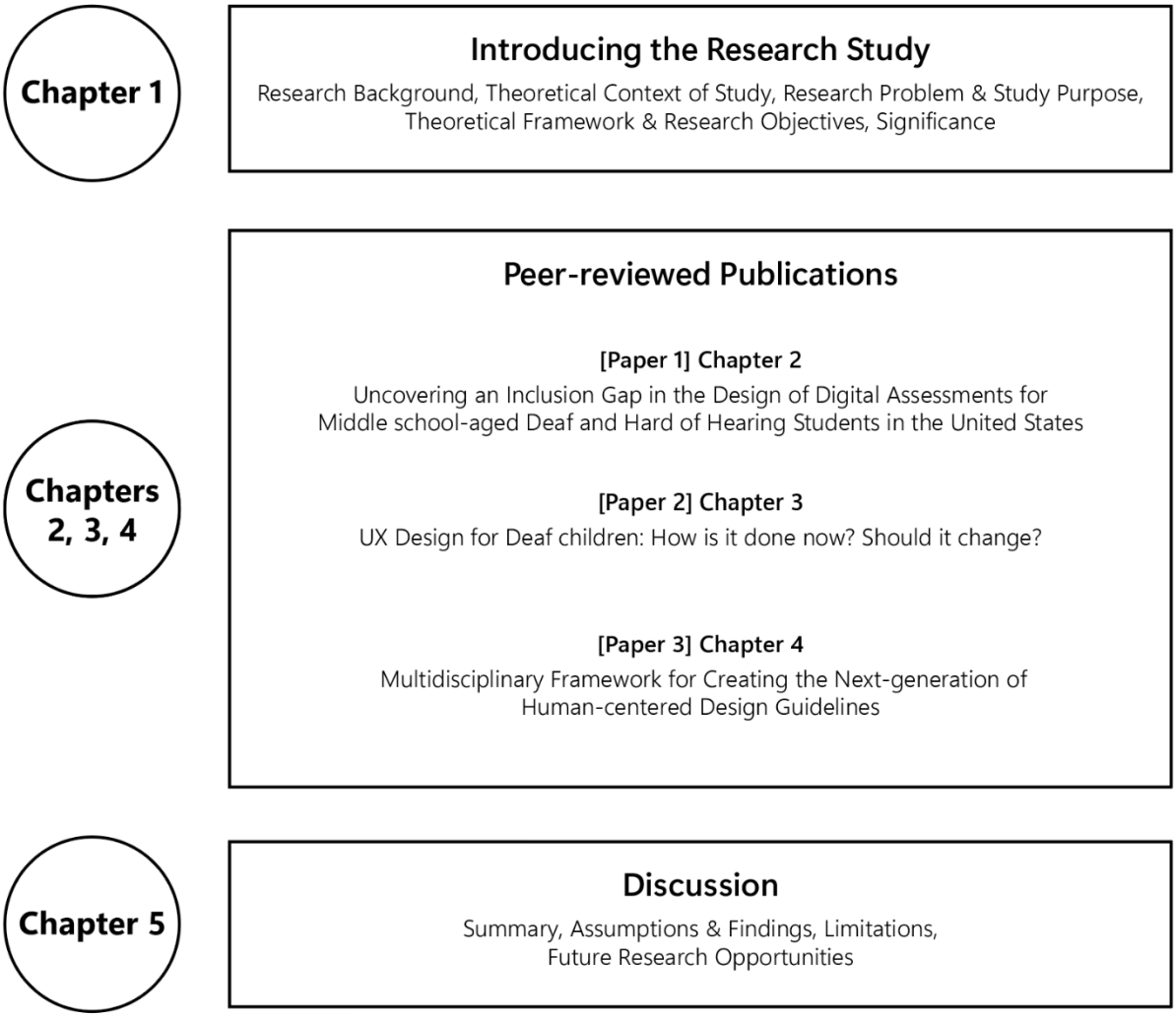

Figure 1.8 - Breakdown of dissertation chapters.

BRIDGE TO CHAPTER 2

The manuscript for the first paper (dissertation chapter 2), entitled "Uncovering an Inclusion Gap in the Design of Digital Assessments for Middle school-aged Deaf and Hard-of-Hearing Students in the United States," was submitted in January 2023 as both a Conference Paper and a Conference Proceeding presentation for the 14th International Conference on Applied Human Factors and Ergonomics. The objective of this paper was to place a spotlight on the high potential for inequity against DHH students who take educational assessments in the United States, as revealed through literature review, and to start a public discourse on this topic.[4] This paper was also intended to serve as a peer-reviewed foundation for the subsequent papers, that were intended to focus more on practitioner interviews and multi-disciplinary research outcomes.

[4] According to ResearchGate, this paper's Research Interest Score was higher than 39% of research items published in 2023.

CHAPTER 2 UNCOVERING AN INCLUSION GAP IN THE DESIGN OF DIGITAL ASSESSMENTS FOR MIDDLE SCHOOL-AGED DEAF AND HARD-OF-HEARING STUDENTS IN THE UNITED STATES

ABSTRACT

What does a score on a digital assessment mean? At its core, a score is a measurement of how a student matches up to a predefined construct. For example, a reading assessment may measure the construct of a student's reading fluency, comprehension, or both. This research seeks to challenge the legitimacy of digital assessment from the lens of Accessibility, User Experience (UX), Inclusive Design, and Marginalized Populations by focusing on the needs of the Deaf and Hard-of-Hearing (DHH) middle school-aged student in the United States.

DHH learners are among the least understood groups. Neither the US Census nor public schools recognize American Sign Language (ASL) as a non-English language used at home. For the sake of discussion, this research references a study by Goman from 2016 which estimates that 14.3% of all Americans aged 12 and older have some form of hearing loss, and a study from the U.S. National Center of Educational Statistics which estimated students with hearing impairment between ages 3-21 at 1% of all students.

These statistics are especially concerning when juxtaposed with how assessments are created. Two of the top educational companies in U.S. use a process called "pretesting" to determine the statistical relevance of the questions used in their assessments. This process involves trialing assessment items with a sample group similar to the population to be assessed. As assessments are increasingly delivered digitally, they overlap with other disciplines like UX Design. In UX, it is well documented that testing with five people finds most problems. If we assume that pretesting uses a similar sample size, it is a reasonable assumption that many items

would not be trialed with DHH students, i.e. this marginalized group isn't populous enough to be accounted for in a statistically relevant pretesting sample.

To provide legitimacy to this claim, this research used structured interviews with subject-matter experts (SMEs) in usability, accessibility, child-computer interaction, and DHH education. The responses provided by these SMEs lent credence to the idea that DHH learners were often not included in digital assessment design either due to being sampled out, a lack of accessibility awareness, and/or the absence of inclusive design guidelines for DHH students. For example, one interviewed Director at a prominent Deaf institution said, "In terms of my field, there isn't some tangible set of design principles that apply in [my] specific area. These things are developing as we go."

Keywords: Deaf, Hard-of-Hearing, design, accessibility, inclusion, education, assessment

INTRODUCTION

In the United States, it is estimated that 14.3% of all Americans aged 12 and older have some form of hearing loss (Goman 2016). Research on Deaf or Hard-of-Hearing students (DHH) is seemingly non-existent despite the number of DHH students in schools increasing each year. (Pizzo and Chilvers 2016). This lack of scientific literature and the novelty of digital assessments at large have caused me to investigate the efficacy of digital assessments that purport to quantifiably measure the proficiency of DHH students in the United States.

I. How are assessments measured?

Assessments provide a quantifiable measure for how a student matches up to a predefined construct. There are two primary methods of measuring student performance in an assessment: Classical Test Theory (CTT) and Item Response Theory (IRT). For the purposes of this article, these theories will not be disputed.

In short, the first method, Classical Test Theory, was developed in order to determine measurement error values to correct test scores. It sought to explain why the same construct could be measured multiple times with different results. (Steyer 2001, p. 1955) The primary concepts are “test observed” (X), “true score” (T), and “error score.” In CTT, a true score represents the average score a student would receive if they completed a test an infinite number of times. As a result, this approach is very dependent on the size of pretesting test-taker samples. Additionally, in order to properly calculate an assessment’s true score in the CTT model, the same question must be administered each time to appropriately determine the error score. (Zanon et al. 2016).

The second method, Item Response Theory, was formed in response to the CTT limitation of needing to deliver the same test items for analysis. Its key concepts are a person’s

"latent ability" and an "item characteristic curve (ICC)" (Yang and Solon 2014, "Basic measurement properties for IRT"). Latent ability refers to a construct being measured, and the item characteristic curve represents the probability of receiving a correct score based on a test-taker's latent ability. Item response theory is founded upon a few key assumptions— assumptions which have caused it to be the more prevalent model in educational assessment.

1) Monotonicity: as the latent trait increases, the probability of a correct answer will always increase.
2) Unidimensionality: there is one dominant trait that is being measured by an item, and it is the primary factor in all item scores.
3) Local independence: scores received on different items have no statistical bearing on one another; i.e. they are mutually independent.
4) Invariance: Item characteristics can be estimated from any point on the ICC. (Yang and Solon 2014, "Basic assumptions in IRT")

II. How do these test theories relate to students with disabilities?

Quite simply, the study and measurement of students with disabilities is novel. To lend credence to this claim, the National Assessment of Educational Progress (NAEP) --which is the largest nationally representative assessment of American students' proficiency in various subjects-- was first administered in 1969, but its first assessment with special needs accommodations was in 1998.

Unsurprisingly, the study of students with disabilities' performance on *digital assessments* is even more novel. Pointedly, it was only until *2017* that NAEP transitioned to a digitally based assessment (NCES, 2019). Furthermore, the transition to digital assessments has forced a new level of interdisciplinary design that did not exist. Namely, fields like user experience design, instructional design, and assessment design have been forced to co-exist, and often in ways that may seem contradictory. For example, in paper-based assessment (PBA)

things like line-wrapping, the font weights, and font sizes are elements that can be kept constant between test-takers—thereby eliminating the potential statistical impact these variations could have on a students' ability to score well on digital assessments.

In the world of digital design, however, it is generally assumed that there will be variance between test taking devices. For example, a web page may require considerably more scrolling on a smaller or low-resolution computer monitor than would be required on a larger, high-resolution computer monitor. While it is possible to lockdown the computer hardware that students use through heavily modifying a computer's firmware (e.g. altering the firmware of a Windows-based PC to prevent the use of hotkey that calls forth operating system functions that are not controlled by the digital assessment, like the Magnification tool). Unfortunately, this type of practice may unfairly punish students with disabilities who may rely on these types of operating system tools to be successful. As a result, assessments have employed the use of accommodations for students with disabilities. Fortunately, in the case of Magnification, it is a widely accepted accommodation.

That said, I have been unable to locate studies that demonstrate that this accommodation does not alter constructs being measured—which is how NCES defines an accommodation. As a result, it becomes difficult to assert why accommodations are not universally provided. This bolsters the argument that assessments that require accommodations may be inherently flawed due to the need to provide special exceptions to measure a construct. The theoretical argument to be made here is that if the assessment was designed from the beginning with the needs of all students in mind, accommodations would be unnecessary.

Fortunately, assessments like NAEP, have been adopting a universal design mindset which aligns with this argument. These universal design elements provide equitable assistance to

all students, regardless of ability—some examples include Text-to-Speech for directions or the ability to change the color contrast of assessment items. The downside to a purely universal design mindset, however, may force statistically small populations to be overlooked in the design process.

III. How statistically significant are Deaf and Hard-of-Hearing Middle-school aged students?

It is likely that anyone who has worked in accessibility in any capacity has been asked to quantify the problem or value proposition associated with persons of disabilities. It is an unfortunate question, because even a single problem for a person with disability is something that should be respected.

That said, I have attempted a best guess at quantifying the number of students in the United States that that know American Sign Language (ASL) to approximate the scope of my research statement. (Please note that not all Deaf students learn ASL. Parents/guardians of Deaf or Hard-of-Hearing children may have elected to provide their children with cochlear implants which allow for improved hearing ability.)

Here is the big caveat: Neither the US Census nor public schools (authorized by the Bilingual Education Act of 1968) list ASL as a non-English language used in the home (Michell et al., 2006, p. 306). As a result, the most relevant information I can find is a 2005 publication by Gallaudet University. Due to the aforementioned limitations of the U.S. Census, Gallaudet was forced to rely on Internet sources, which place ASL usage in the United States at somewhere between 100,000 and 15,000,000 people.

TABLE 2. Selected Internet Sources for Estimates of the Prevalence or Prevalence Ranking of ASL Use in the United States

Prevalence or Prevalence Ranking Estimate	Website where estimate was found
100,000–500,000	ERIC Digests (Wilcox and Peyton 1999) MSN Encarta (Wilcox 2004) Ethnologue.com (Ethnologue 2004)
250,000–500,000	American Sign Language Program at the University of Iowa (Department of Speech Pathology and Audiology 2004) ASLTA (NC ASLTA and NCAD Ad Hoc Committee 2004) Colorado Department of Human Services (Colorado Commission for the Deaf and Hard of Hearing, n.d.)
300,000–500,000	BarnesandNoble.com (Costello 1994) SignWriting.org (Rosenberg 1999)
500,000*	American Academy of Family Physicians (CDGAP 1997) ASLinfo.com (ASLinfo.com, n.d.) DEAF CAN! (Deaf Community Advocacy Network n.d.)
500,000–2,000,000	Brenda Schick (Schick 1998) DawnSignPress (DawnSignPress 2003) Gallaudet University Library (Harrington 2004)
15,000,000	Aetna InteliHealth (Gordon 2001)
Third most-used language	HandSpeak (HandSpeak.com n.d.) Health Literacy Consulting (Osborne 2003) Missouri Office of State Courts Administrator (Office of State Courts Administrator n.d.)
Fourth most-used language**	ASHA Leader Online (Scott and Lee 2003) Deaf Resource Library (Nakamura 2002) NIDCD (National Institute on Deafness and Other Communication Disorders 2000)
Third to tenth most-used language	Wikipedia (Wikimedia 2004)

*The sites listed here used the number 500,000 in similar but not identical ways, such as "approximately one-half million," "more than one-half million," or "more than 500,000."
**These sites include those that report that ASL is the third most-used non–English language.

Figure 2.1 - Prevalence of ASL Use in the United States. Reprinted from "How Many People Use ASL in the United States? Why Estimates Need Updating", by Michell, R. E., Young, T. A., Bachleda, B., & Karchmer, M. A. (2006), p. 315

Using these estimates, we can attempt a loose approximation of ASL usage in middle-school-aged students (MS students). Another reason why these numbers are not definitive is that the US National Center for Educational Statistics (NCES) aggregates PreK-8 as one dataset. According to NCES, the projected total number of students in PreK-8 in 2020 is 39,476,000.

(NCES, 2019). If we assume the following: 1) Prekindergarten and Kindergarten are distinct grade-levels, 2) Students have an equivalent distribution between all 10 grade levels (preK-8), 3) Middle-school represents grades 6-8; we can estimate that there are 11,842,000 students in middle school in the United States.

$$39{,}476{,}000\ students \div 10\ grade\ levels = 3{,}947{,}600\ \frac{students}{grade\ level}$$

$$3{,}947{,}600\ \frac{students}{gradelevel} \times 3\ grade\ levels = 11{,}842{,}000\ MS\ students$$

The last datapoint we need to come to an ASL in middle-school number is the total US population. According to Census.gov, the US Population is approximately 330,000,000 as of October 2020. If we divide the total middle school population by the US population, we can calculate the percentage of middle-school aged students at 3.59% of the total population.

$$11{,}842{,}000\ MS\ students \div 330{,}000{,}000\ US\ population\ = 0.0359\ \frac{MS\ students}{US\ population}$$

Another assumption that we will make is that the proportion of MS students to total US population is an equal proportion to that of ASL-speaking MS students of the US to the total ASL-speaking population of the US. This would place the ASL MS student population between 3,590 and 538,500.

$$0.0359 \frac{MS\ students}{US\ population} \times\ 100{,}000\ ASL \cdot US\ Population = 3{,}590\ ASL\ MS\ students$$

$$0.0359\ \frac{MS\ students}{US\ population}\ \times\ 15{,}000{,}000\ ASL \cdot US\ Population = 538{,}000\ ASL\ MS\ students$$

The reason for providing these numbers is not to lessen or diminish the lived-in experiences of these middle-school aged students. The goal is to demonstrate the duality of this population group as being one that measures in the thousands to hundreds of thousands, yet can at times be a rounding error that can cause these students to be ignored due to statistical rounding errors.

IV. How are test items developed today?

It is unlikely that every testing institution uses an identical process for creating items, but they do have many similarities. In this article, we will remark on the item creation process for two of the largest testing institutions globally- Educational Testing Services and Pearson. ETS's test development process follows seven steps (ETS 2020):

1. **Define Objectives:** "Educators, licensing boards or professional associations identify a need to measure certain skills or knowledge."

2. **Item Development Committees:** "The answers for the questions in Step 1 are usually completed with the help of item development committees, which typically consist of educators and/or other professionals appointed by ETS with the guidance of the sponsoring agency or association."

3. **Writing and Reviewing Questions:** "Each test question undergoes numerous reviews and revisions to ensure it is as clear as possible, that it has only one correct answer among the options provided on the test, and that it conforms to the style rules used throughout the test."

4. **The Pretest:** "After the questions have been written and reviewed, many are pretested with a sample group similar to the population to be tested."

5. **Detecting and Removing Unfair Questions:** "Trained reviewers must carefully inspect each individual test question, the test as a whole, and any descriptive or preparatory materials to ensure that language, symbols, words, phrases, and content generally regarded as sexist, racist, or otherwise inappropriate or offensive to any subgroup of the test-taking population are eliminated."

6. **Assembling the Test:** "After the test is assembled, it is reviewed by other specialists, committee members and sometimes other outside experts."

7. **Making Sure that the Test Questions are Functioning Properly:** "Even after the test has been administered, statisticians and test developers review to make sure that test questions are working as intended.

Pearson, conversely, provides a cyclical diagram to describe its test content creation process (see **Figure 2.2**). It is not immediately clear which is the first step in the process, but I believe it to be Periodic test and item analysis. Pearson describes these terms as follows:

> **Pretesting:** "Pre-testing (or trialing) of test items refers to the administration of the items solely to gather performance statistics to determine if the items should be included in the operational item pool from which items are selected for future administration & scoring." (Pearson, 2018, "Pre-testing")
>
> **Item Analysis:** "Statistical investigation of the performance of test items to obtain information about the quality of the items." (Pearson, 2018, "Glossary")
>
> **Item pool updated:** "Psychometric analysis of operational scored items should be conducted to evaluate which items should remain in the operational pool and which items should be retired from use." (Pearson, 2018, "Item analysis")

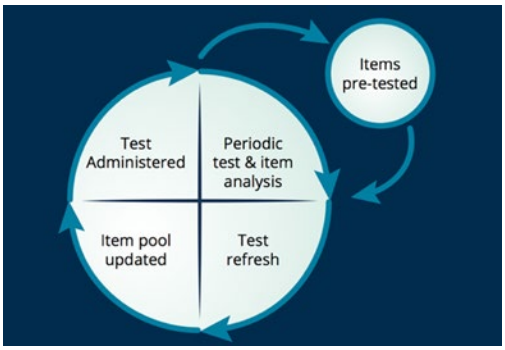

Figure 2.2 - Pearson's test content creation process (Pearson, 2018).

V. What's the problem?

The processes described by Educational Testing Services and Pearson appear are sufficient from the perspective of effective item development (ID). At a surface-level, they also

appear sufficient from an equity perspective-- from the understanding that items will be pretested and will have gone through various ID committees. The assumption here is that these committees are well-versed in the nuances of their represented populations.

Upon closer reflection, the problem lies in the pre-testing step that is common to both ETS and Pearson. Let us revisit ETS's definition of pre-testing (with **bold** representing my emphasis): "… many are pretested with **a sample group similar to the population to be tested**."

Unfortunately, students with disabilities represent a statistically small segment of the entire student population, especially so for DHH middle school-aged students, as was communicated in Section III (see **Figure 2.3**). As such, it is not unreasonable to assume that this population can and *will* be sampled out.

In the field of usability, for example, user-testing with five participants finds most problems, user-testing with 39 participants is at the high-end, and most companies use 11 participants per testing session (Nielsen 2012). If we use the industry recommendation of five test participants, that would mean that every student of every disability type, with the exception of "specific learning disability" would be omitted due to representing less than < 20% of a nationally-representative sample (NCES 2020).

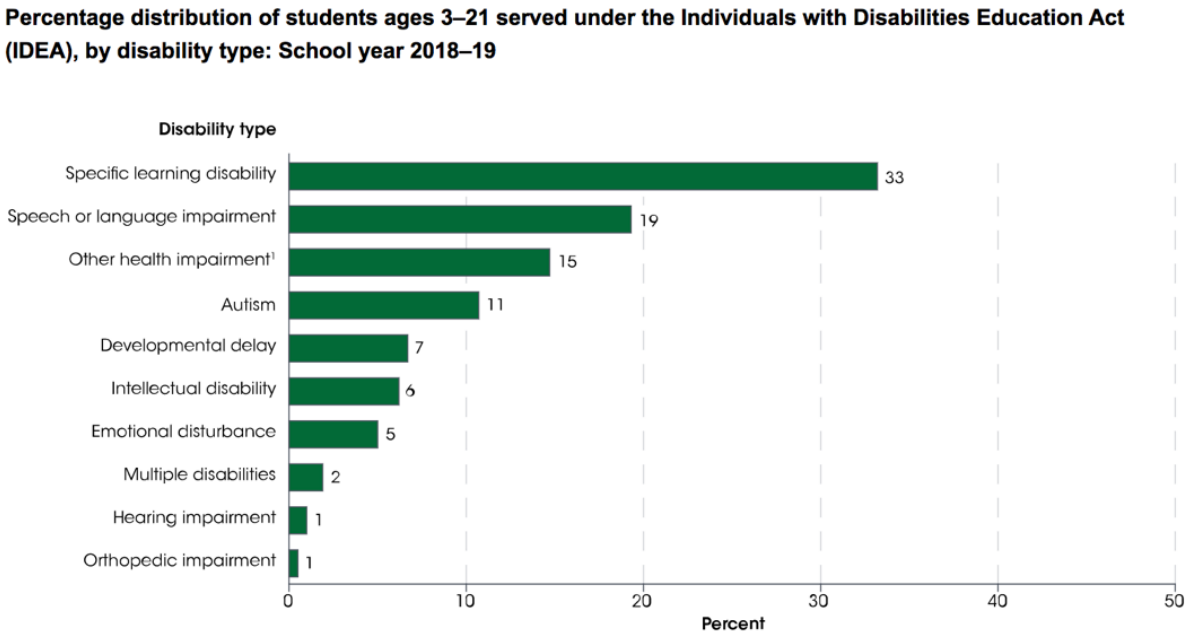

Figure 2.3 - Percentage of students with disabilities (NCES 2020).

VI. What can be done?

Before we can continue to discuss the problem facing DHH middle school-aged students, it is important to understand the larger context of digital accessibility in the United States.

In 1996, the US Department of Justice ruled that websites were public accommodations. (US DOJ 2012). As such, businesses were not allowed to discriminate against individuals with disabilities. In 1998, Section 508 of the Rehabilitation Act of 1973 required the federal sector (e.g. government agencies, federally funded non-profits, K-12 schools) to have accessible digital assets (Pan 2017). In the years that followed, the Web Content Accessibility Guidelines versions 1.0 and 2.0 that were developed by the World Wide Web Consortium (W3C) came out in 1999 and 2008, respectively, as a way of improving web accessibility. In a move that codified WCAG as a de-facto standard, the US Government then further modified Section 508 in 2018 to require WCAG 2.0 Level AA compliance (US GSA 2018).

Understanding all this history, one would expect ETS and Pearson to explicitly reference WCAG in their content creation guidelines, especially for work done on behalf of the federal government. Unfortunately, this is not the case as was exemplified in Section IV. I believe this to be unacceptable, and this needs to change. believe it to be unreasonable that both Pearson and ETS do not explicitly reference WCAG in terms of work process or review process in their respective content creation guidelines.

To give credit where it is due, ETS and Pearson are both members of the W3C. From the interviews I have conducted with individuals who work in the field of accessibility at these and other companies, I understand that it is a difficult battle to bring an entire company on-board with the idea that accessibility is the responsibility of every employee. As such, it is my hope that this article is perceived less as an attack on credibility, but more so a rally cry to get practitioners excited to work with their accessibility partners.

VII. What do experts have to say about this?

The theoretical constructs I sought to investigate were usability, accessibility, and child-computer interaction. Using structured interviews and the strategy of the Critical Decision Method (CDM), domain specific knowledge was elicited from subject matter experts (SMEs) in usability, assessment design, child computer interaction, accessibility, and Deaf or Hard-of-Hearing education.

CDM works by applying a set of probing questions as a framework for allowing experts to recall aspects of their decision-making process. Since its inception, it has gained notable use in the fields of instructional design, system development, and information technology (Taylor, 2006). The strength of CDM is rooted in empirical studies that have found that subject matter experts enjoy telling stories, and that “some practitioners learn on the job by sharing their ‘war

stories' and even report that they learn more that way than through formal instruction" (Hoffman et al. 1998, p. 271).

Among the most telling of the questions asked were: "What principles or guidelines are you aware of and use in your profession?" No list of existing guidelines were provided so as to not bias the answers provided. To my surprise, there was no single guideline that emerged as being universally used by all SMEs. A slight majority did emerge, however, with WCAG and Universal Design being mentioned by 4 out of the 7 SMEs, which represented a slight majority.

The quote which resonated the most with me came from Melissa Malzkuhn, the Director of the Motion Light Lab at Gallaudet University. Through an interpreter, she said:

> *"I'm not really clear on what you mean by that question in terms of design principles, and really it's interesting because when we developed these storybook apps for Deaf readers there was no precedent for them. ...In terms of my field there isn't some tangible standard set of design principles that apply in this specific area. These are things that we're developing as we go."*

Conclusion

At the onset of this research, I believed that a clear set of guidelines could emerge which would in turn be provided to the leading digital assessment companies in the United States to improve fair and equitable assessment for Deaf and Hard-of-Hearing middle school-aged students. Unfortunately, that outcome did not happen. What did happen, however, was that an inclusion gap in assessment pre-testing activities became apparent and the SMEs who I spoke to came to understand that even if they were not designers themselves, they had valuable expertise that could aid in the design of digital assessments and designerly output.

REFERENCES (Chapter 2)

ETS. (2020). *How Tests and Test Questions are Developed*. Retrieved Oct 2020, from ETS.org: https://origin-www.ets.org/understanding_testing/test_development

Goman, A. (2016). *HOW MANY PEOPLE HAVE HEARING LOSS IN THE UNITED STATES?* Retrieved from Johns Hopkins Bloomberg School of Public Health: https://www.jhucochlearcenter.org/how-many-people-have-hearing-loss-united-states.html

Hoffman, R. R., Crandall, B., & Shadbolt, N. (1998, June 1). *Use of the Critical Decision Method to Elicit Expert Knowledge: A Case Study in the Methodology of Cognitive Task Analysis*. Retrieved from SAGE journals: https://journals-sagepub-com.prox.lib.ncsu.edu/doi/abs/10.1518/001872098779480442

Michell, R. E., Young, T. A., Bachleda, B., & Karchmer, M. A. (2006). *How Many People Use ASL in the United States? Why Estimates Need Updating.* Retrieved from Project MUSE: https://muse-jhu-edu.prox.lib.ncsu.edu/article/197164

NCES. (2019). *Digest of Educational Statistics.* Retrieved from NCES: https://nces.ed.gov/programs/digest/d19/tables/dt19_105.20.asp

NCES. (2019). *NAEP History and Innovation*. Retrieved from Ed.gov: https://nces.ed.gov/nationsreportcard/about/timeline.aspx

NCES. (2020, May). *The Condition of Education: Students With Disabilities*. Retrieved from National Center for Education Statistics: https://nces.ed.gov/programs/coe/indicator_cgg.asp

Nielsen, J. (2012, June 3). *How Many Test Users in a Usability Study?* Retrieved from Nielsen Norman Group: https://www.nngroup.com/articles/how-many-test-users/

Pan, J. (2017, September 26). *508, ADA, WCAG: What's the difference?* Retrieved from LOGIC Solutions: https://www.logicsolutions.com/508-ada-wcag-accessibility-difference/

Pearson. (2018). *Evaluating performance*. Retrieved from Pearson VUE: https://home.pearsonvue.com/Documents/Security/pearson_vue_test_items_US.aspx

Pizzo, L., & Chilvers, A. (2016). *Assessment and d/Deaf and Hard-of-Hearing Multilingual Learners: Considerations and Promising Practices.* Retrieved from American Annals of the Deaf: http://www.jstor.org/stable/26235251

Taylor, H. (2006). Eliciting Tacit Knowledge Using the Critical Decision Interview Method. In M. E. Jennex, *Knowledge Management in Modern Organizations* (pp. 285-301). Idea Group, Inc.

US DOJ. (2012, Dec 7). *Nondiscrimination on the Basis of Disability; Accessibility of Web Information and Services of State and Local Government Entities and Public Accommodations*. Retrieved from ADA.gov: https://www.ada.gov/anprm2010/web%20anprm_2010.htm

US GSA. (2018, May). *Applicability & Conformance Requirements*. Retrieved from Section508.gov: https://www.section508.gov/create/applicability-conformance

Yang, F. M., & Solon, T. K. (2014, June). *Item response theory for measurement validity.* Retrieved from National Instutites of Health: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4118016/#B3

Zanon, C., Hutz, C. S., Yoo, H., & Hambleton, R. K. (2016, April 18). *An application of item response theory to psychological test development.* Retrieved from Springer Open: https://prc.springeropen.com/articles/10.1186/s41155-016-0040-x

BRIDGE TO CHAPTER 3

The second paper**,** "UX Design for Deaf Children: How is it done now? Should it change?" was published by SageJournals in the Journal of Educational Testing Systems on May 6, 2024. It begins with an expanded literature review of the origins of user experience design, child-computer interaction, and UX Design education. The author felt that this information was warranted to better illustrate the theoretical perspective of the research prior to describing the research interviews. The rationale for this decision was rooted in the observed uncertainty that some self-identified non-designer practitioners demonstrated when asked to describe their design process. These non-designers felt apprehensive about whether their responses would be of value to designers.

With an understanding that user experience design may have originated from non-designerly sources (e.g. computer programmers and product designers), the reader is better positioned to adopt the viewpoint of the author that all the participants' responses are valid to this design research process. That said, to respect the perspectives and self-identified roles of the research participants, this paper marks a point when the author begins to adopt a classification scheme of designers vs subject-matter experts (non-designers) which will continue to be used extensively throughout this dissertation.

Among the primary research objectives of this paper was to determine whether designers and subject-matter experts could shed light on what, if anything, should change in their design process when creating products for children, especially Deaf or Hard-of-Hearing (DHH) children. Additionally, this paper sought to determine if today's design education can sufficiently position designers to develop products for DHH children. Upon completion of the interviews, further opportunities for design research emerged and were shared with the reader.

CHAPTER 3 UX DESIGN FOR DEAF CHILDREN: HOW IS IT DONE NOW? SHOULD IT CHANGE?

ABSTRACT

The process by which user experiences (UX) for children are created is uncertain, especially for Deaf and Hard-of-Hearing (DHH) children. This paper seeks to I) describe the origins of UX and Child-computer interaction and to describe what is being taught to designers today; II) use the example of digital assessment to extract insights from practitioners about UX Design for DHH children; and III) describe opportunities for amending today's UX design curricula to foster more equitably designed products for children.

Interviewing 13 practitioners revealed the following findings: 1) Allowing non-designer participants to self-identify their profession makes them more likely to discuss design concepts. 2) The challenge of working with children is more about legal/risk assessment than any gaps in design education. 3) Practitioners who design products for children follow similar processes to designers who design for adults; age-specific language is the main difference between child and adult UX.

Keywords: Deaf, Hard-of-Hearing, design, user experience, child computer interaction

INTRODUCTION

What should user experience designers (henceforth shortened to "designers" or "UX Designers") do differently today when designing for Deaf or Hard-of-Hearing (DHH) children? Where does this question come from and what literature is out there? According to a 2023 study performed by WebAIM, 96.3% of the top 1,000,000 most-visited home pages globally triggered accessibility errors when checked against version 2.1 of the Web Content Accessibility Guidelines (WCAG 2.1), the world's leading set of accessible guidelines. This speaks to an overwhelming deficiency in applied knowledge about building accessible web pages across designers, developers, and other business professionals. It is therefore reasonable to assume that an even smaller number of designers understand the needs of DHH users, and even less understand the needs of DHH children.

Fortunately, it is projected that the global market for Diversity and Inclusion (sometimes described as DEI, or Diversity, Equity, and Inclusion) will grow to $15.4 Billion USD, which represents a 12.6% increase in compounded annual growth rate between 2022 and 2026. (Ali 2022). The authors' hope is that this increase in DEI spending, and an increased awareness of disability in general will spur subsequent increases in accessibility training and accessibility education for designers of all skill-levels. A further hope is this will spur educators and learning institutions to instruct designers on the creation of accessible products for children with disabilities versus only learning how to design for adults with disabilities.

This paper is organized into three parts. The author will first describe I) the origin of the User Experience Design (UXD) and Child-computer Interaction (CCI) disciplines and describe what is taught to designers today; II) use a set of interviews with UX, Accessibility, and Education practitioners to reveal findings that could help improve UX Design education and the

CCI field at large; and lastly III) offer opportunities for further research to build off the findings described in the previous two parts.

PART I. LITERATURE REVIEW

User Experience Design

What is user experience (UX) design? According to the Interaction Design Foundation, UX Design is the process that design teams use to create "meaningful and relevant experiences to users". It is generally assumed that the end-users of UX design are engaging with a digital product, but that is not always the case. That said, its scope varies widely between who you speak to. Part of this is due to it being a relatively new design discipline with an ambiguous history. Namely, it's widely disputed how old the field is and what it covers. For example, there are at least three founders of user experience design:

The first is Alan Cooper. In 1995, Cooper was the creator of the "Visual Basic" programming language that was later acquired by Microsoft and formed his own design consultancy "Cooper" which was founded on the idea of "interaction design" which Cooper claims he invented. This claim is bolstered by his biography in the Computer History Museum (CHM, 2017).

The second is Bill Moggeridge, a prolific product designer and co-founder of IDEO. In 2002, he claimed at DIS 2002 (the Designing Interactive Systems conference) that he and his colleagues at Xerox created the standards for graphical user interface design in 1982 (Verplank, Sutcliffe, Mackay, Amowitz, & Gaver, 2002).

The third is Don Norman, of the Nielsen/Norman Group. In 1988, he published a book entitled The Design of Everyday Things, and in this book he talks about "user-centered design"

(Norman, 1988). In addition to this book, in 1995, Norman convinced Apple to give him a job title with "user experience" in the name, making him the first person with user experience in their job title (NNgroup, 2016).

All of this is to say that the field of user experience design is relative newcomer to the design profession overall. At best, the field is approximately 20-50 years old. Its definition is fluid, and its constituents are as varied as these founders—programmer (Cooper), product designer (Moggeridge), software architect (Norman).

For the purposes of this paper, we will proceed with Norman's interpretation of user experience design, which he defines as a cohesive set of experiences from "from initial intentions through final reflections, from the first usage to help, service, and maintenance." (IDF n.d.). To put it in other words, user experience design is a discipline that is inextricably a marriage between business outcomes and the needs of some type of user, often the consumer of some kind of product.

Child-computer Interaction

The closest approximation to a degree that targets the user experience design of digital products for children is CCI, or Child-computer Interaction. CCI emerged from the human-computer interaction (HCI) field. Defined by the Association for Computing Machinery (ACM), HCI is defined as "a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them" (Hewett et al. 1992, p. 5).

CCI is about getting designers, developers, and other stakeholders to leverage their HCI expertise in support of the needs of the child. Unfortunately, the definition of "child" remains ambiguous. During the creation of a CCI Special Interest Group, ACM members sought to define

children as those who are five years or younger (Hourcade et al., 2016). This represents a strangely specific pivot from the body of research that defines CCI. In *The Nature of Child Computer Interaction*, Read & Bekker demonstrates that most work in Child-computer interaction focuses on children between the ages of five and eleven, with an understanding that ages three to sixteen are acceptable ranges for CCI research (Read and Bekker 2011, p. 3).

While it is acknowledged that teenagers represent a valid subset of CCI, it has been documented that research into this population group has been on the decline (Giannakos et al. 2020). One of the reasons posed for this is that authors may have failed to give their studies appropriate keywords. That said, teenagers represent a demographic that may lend itself to better test results when conducting usability studies, due to their greater ability to verbalize concerns. (Markopoulos and Bekker 2003).

Challenges with UX Design Education

An investigation into leading design programs reveals no direct mention of the age of users for which designers will be learning to create user experiences for. As of the publication of this paper, there are no known design degree programs or coursework in the United States that provides specialized instruction in Child-computer interaction. This was verified by reviewing the course offerings by the top 10 universities as ranked by U.S. News & World Report in 2023.[1] The only known degree program that provides design education for child users education is the University of Central Lancashire in England. That said, there remains an international push for increasing the amount of literature and exposure to children's needs in design. Much of it is rooted in the Interaction Design for Children (IDC) community that has been working on normalizing co-designing practices with children for 20 years (Lehnert et al. 2022).

[1] https://www.usnews.com/best-colleges/design-major-5004

Possible Courses for CCI education. Coursework which could be used to introduce CCI to academia are Practicum courses. These courses task college students with solving real-world problems and designing solutions for local businesses. These are not limited to design schools and are employed by colleges as varied as New York University's business and fashion schools to University of Notre Dame's Social Work program (Walker 2021; ECDC n.d.). These types of courses theoretically should overcome one of the most difficult aspects of designing for children, e.g., the recruitment of child users. The assumption is that these businesses would already have an installed user base that could be co-designed with.

Unfortunately, searching for documented outcomes of these practicum courses does not yield results. The lack of available publicly available Practicum outcomes may be due in part to:

1. limitations placed on student designers by the businesses seeking to safeguard their Intellectual Property Rights (Smith, 2016); and/or
2. Businesses and government entities seeking to reduce their levels of risk associated with things like parental consent, personally identifiable information for minors, and federal security clearance requirements.

Expanding one's search from CCI courses to CCI-adjacent degree programs results in degrees in Education. Examples include: 1) Rutgers University's Digital Learning Environments (DLE) concentration within the Doctor of Education (Ed.D) program alludes to the design of classrooms; student age is not explicitly stated, but the umbrella Ed.D program does mention K-12 Education. 2) North Carolina State University's Learning Design and Technology (LDT) major within the College of Education. This major recommends taking courses from the College of Design, namely Design Inquiry I and II. While these Design Inquiry courses describe teaching students to use human-centered methods, they too do not reference working with child users.

Where NCSU LDT differs from Rutgers DLE is that it assumes a need to have students participate in both Education and Design courses. Unfortunately, there is no quantifiable method to date to determine the efficacy of either approach as it pertains to CCI, especially in the absence of publicly available course outcomes.

What is being taught to designers today?

UX Designers generally follow a double-diamond model of discovery as defined by the UK Design Council and adapted by the Nielsen Norman Group (see **Figure 3.1**) (Rosala 2020). In this model, designers are challenged to research their problem space, clearly identify their target user, and validate their assumptions by testing their design outcomes with individuals who closely align with this target user.

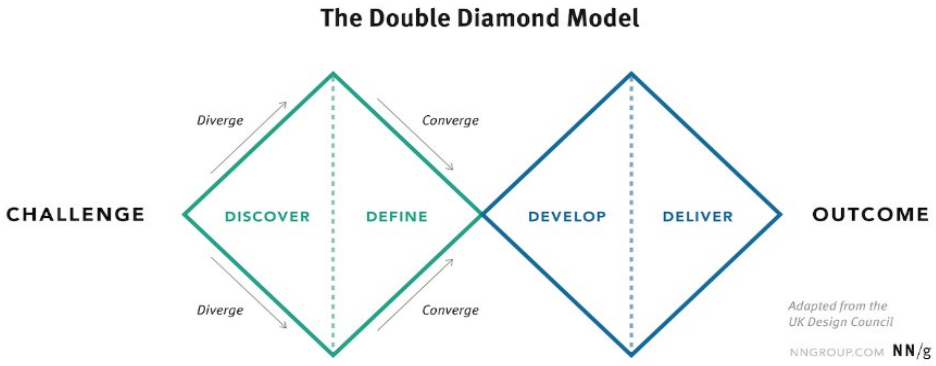

Figure 3.1 - Double Diamond model for design process, adapted by NN/group.

In the Discover phase, the idea is that a designer should strive to determine as many unknowns as possible of their project and the project's problem space through increasingly expansive UX methods of inquiry. These UX Methods can take the form of Stakeholder observations and interviews, user journey maps, affinity maps, persona generation, etc (Maguire 2022).

This Discover phase is where the biggest gaps and risks lay for designers of Deaf and Hard-of-Hearing children. Based on the relative sizes of the Challenge and Outcome diamonds in the diagram above, it is implied that the amount of effort needed to discover and design the Challenge is equivalent to the amount of effort needed to develop and deliver the designed Outcome.

In the case of designing for DHH children, this could not be further from the truth for two reasons: 1) the lack of CCI education as described in the previous section of this paper; and 2) research on Deaf or Hard-of-Hearing children being seemingly non-existent, even though the number of DHH students in American schools increases each year (Pizzo and Chilvers 2016).

One possible approach to address this lack of public case studies in CCI is the immediate dissemination of new knowledge that is made freely available upon publication. The author proposes interviewing practitioners who either directly work in the CCI space and/or practitioners with comparable domain knowledge as an example of this. In Part II of this paper, the author models this proposal by interviewing practitioners in the fields of UX, Accessibility, and Education to learn more about UX Design for Deaf and Hard-of-Hearing children.

PART II. RESEARCH STUDY

The authors' proposed solution to the lack of information regarding real-world Child-computer interaction case studies is to reach out to professionals, be they in academia or industry, who have experience designing products for children and to ask them to share their personal experiences and perspectives. It is only by accruing and publicly disseminating case studies can designers unfamiliar with designing for children can be sufficiently prepared.

Understanding the challenges with interviewing children in the US, this research sought second-hand testimonials from practitioners in the space of UX, Accessibility, and Education to

learn about how they perform their work. The word practitioner is used here to not discriminate between academic faculty, designers, software developers, content strategists, etc. Prior to interviewing these practitioners, the authors' hypothesis was that it is critical for designers to be educated in the practice of designing for children to ensure equitable products for children.

Study Design

This research consisted of 45-to-60-minute interviews with 13 practitioners in the space of UX, Accessibility, and Education on the topic of digital assessment for Deaf or Hard-of-Hearing students. Additionally, participants were directed to self-identify their job roles to prevent researcher bias (see **Table 3.1**).

Table 3.1 - Distribution of practitioners' self-identified job titles and experience working on products used by children.

Self-identified Job Titles			
Design Role	**Design Role & Products used by children**	**Subject-matter Expert**	**Subject-matter Expert & Products used by children**
Senior Software Engineer & Accessibility Designer	Lead R&D UX Researcher	Director of Curriculum Development	Research Scientist
UX Lead for Accessibility	Manager of Technical Content Design		Sr. Software Architect
	Senior Accessibility Manager		Certified Teacher of the Deaf
	Director of User Experience		Director Motion Light Lab
	UX Researcher		Teacher of Disabled Students

Of the 13 participants interviewed, 7 of the participants self-identified with design roles with 4 of those explicitly naming UX as part of their job title. 10 out of the 13 had experience working on products that were used by children.

This study took place over two years and used social media, email, and word-of-mouth recruitment techniques. Recruitment proved especially difficult, with a known factor being that participants were not paid for volunteering their time. While a common practice in professional user experience research, the research committee felt that compensation might add a degree of bias to this study.

The questions asked were as follows:

Interviewee Classification

- What is your job title?
- How do you define your field of expertise?
- Who do you design for?
    - (If not children) Do your products get used by children? Why or why not?
    - If age is not specified) Do you think middle school-aged students use your products?

Work Process

- What design principles are you aware of and use in your profession?
- How would you describe your daily tasks?
- Does your work affect a Deaf or Hard-of-Hearing audience?

Child-computer Interaction

- Does your work process vary based on whether your target is an adult or child?

    - (If yes) How does your design process vary from adult-targeted products to child-targeted products?
    - (If no) Why doesn't your design process vary from adult-targeted products to child-targeted products?
- What would you change about your current process, if anything, to suit child audiences?

Deaf or Hard-of-Hearing (DHH)

- Does your work process vary based on whether your target audience is Deaf or Hard-of-Hearing?
    - (If yes) How?
    - (If no) Why not?

Analyzing Responses

Critical Decision Method (CDM) was used to analyze how the practitioners made their decisions regarding their work processes. This method consists of recording and transcribing the interviews, and then using content analysis to place the transcripts into codes. (Zannieret al. 2007). It works by applying a set of probing questions as a framework for allowing experts to recall aspects of their decision-making process. Since its inception, it has gained notable use in the fields of instructional design, system development, and information technology. (Taylor H. 2006)

The strength of CDM is rooted in empirical studies that have found that subject matter experts (SMEs) enjoy telling stories, and that "some practitioners learn on the job by sharing

their ‘war stories’ and even report that they learn more that way than through formal instruction.” (Hoffman et al., p. 271)

Values Coding was used to enhance CDM due to its ability to both “explore cultural values, identity, intrapersonal and interpersonal participant experiences” and its ability to delve into the subjective nature of the human experience (Saldaña 2013, p. 111).

The authors’ goal with this methodology was to better understand the types of information that practitioners might need to successfully design products for DHH child students.

Exemplar Research Interview – Deborah Bilik[2]

The following represents a synopsis of one of these interviews to provide context for how this interview framework was used in this study, and to demonstrate how it could be used to elicit design collaboration in future DHH child products and studies. The interviews with the other research participants may be published as part of a larger dissertation study for North Carolina State University. The hope is that this framework could also be adapted by other researchers to investigate the designerly work processes of practitioners in the space of Child-computer interaction, especially as it pertains to other disabilities like visual impairment, motor skills, etc.

Interviewee Classification. Deborah Bilik is a Certified Teacher of the Deaf for grades Kindergarten to grade 12 (K-12), works in the public-school sector. She has been teaching for 25 years across various curriculums. When she first graduated from college, Deborah worked as an educational interpreter where she followed and signed for a Deaf student. Later in her career she was strictly a middle school teacher, but now she currently works as a Kindergarten to grade-5

[2] The exemplar interview is permitted to be shared in this publication due to the participant providing Broad Consent to future research uses of her identifiable information as part of the research interview consent process.

teacher at the Marie Katzenbach School for the Deaf in New Jersey (Katzenbach)[3]. When asked who Deborah designs for, she stated that she designs for her students and their diverse, individual needs. There are occasions where she will have to create things for her supervisor, but she estimates that 90% of her work is for her students.

Work Process. In terms of her work process, Deborah was unaware of formal design principles, but stated that her work process is largely informed by students' IEPs, or Individual Education Plans. Some examples of IEP accommodations for print or digital artifacts include enlarged print or different colored backgrounds (e.g. yellow paper with blue print); or in the physical space: different furniture and/or alternative seating arrangements in the classroom.

Her daily tasks include coordinating a reading intervention program for K-8 Deaf children where she sets up centers that work on specific skills for each students; tracking IEP goals for her reading and language students; designing the school's website and social media; running and playgroup for children under three years old; and teaching English to adults. When asked if COVID has affected her ability to teach her students, she says that her school didn't miss a beat and used Zoom as soon as schools shut down. Captions and the inherent nature of video conferencing lent itself well to teaching Deaf and Hard-of-Hearing students.

Child computer interaction. On the topic of child computer interaction, Deborah does believe that her process varies between whether she is performing work for a child vs an adult. She prefaced her response by stating that she doesn't design much for adults, but she was emphatic that there is a difference in access to language between adults and children. For that reason, it is especially important that any instructional language given to children has a corresponding ASL equivalent.

---

[3] According to the school's "About Us" page on NJ.gov, "Katzenbach has over 100 years of experience educating New Jersey's Deaf and Hard of Hearing children." https://www.nj.gov/education/mksd/about/

Deaf or Hard-of-Hearing (DHH). When asked to explicitly describe how her work affects a Deaf and Hard-of-Hearing/HH audience, Deborah's first answer was that she teaches children how to read. She says this is particularly a struggle because it often involves translating English to American Sign Language (ASL).

Among the reasons for this is that her students often lack the incidental knowledge that most other students receive at home. For example, her Deaf students may not have someone who signs at home, and so there may be gaps in those students' vocabulary that might not exist in households where a student is learning English both in school and from listening to their parents. That said, Deborah is flexible with her teaching methods. For example, she may focus more on ASL for her Deaf students, and English for Hard-of-Hearing students.

Findings

Finding #1 – A Repeatable Research Framework

One of the goals of this research was to create a repeatable framework of interview questions that could be asked of practitioners that strikes the right balance of acknowledging their expertise, while also building enough rapport to allow these participants to feel comfortable with engaging in a healthy and meaningful dialogue about design.

For example, a practitioner may have plenty of experience working with minors, but they may doubt how relevant their experience is to a designer, as was the case with Deborah Bilik. Conversely, a designer may have plenty of experience designing for adults but may question the relevancy of their experience to a researcher who is exploring the DHH experience.

Fortunately, these questions were successful in that every participant expressed doubt about the relevance of their experience to the interview process in the recruitment phase—but by

the conclusion of their interview, these same participants expressed a willingness to participate in future rounds of interviews. The Lead R&D UX Researcher offered that this style of interview may lend itself to a podcast, wherein practitioners can share their design experiences for other designers to learn from.

Finding #2 – Significance of Child-specific Design Education

No participants interviewed in this study cited a specific academic or licensed background in Child-computer interaction. That said, every participant could reference specific studies or products they worked on that affected children and could intuit how a UX designer could perform design for a child. In this case, they could see the value in serving as a partner to a designer who is emerging in the CCI space.

Unfortunately, in some cases, not every interview participant felt comfortable sharing their personally identifiable information; they especially did not feel comfortable sharing what company they worked for due to concerns about how their employer may respond. This affirms the authors' claims that important information about designing for children is willfully being withheld from the public, likely under the auspices of trade secrets.

Surprisingly, the practitioners who claimed to not have experience working with children followed design processes that were not dissimilar to the processes described by those who did have this experience. The processes employed by all the practitioners aligned with the Double Diamond process (**Figure 3.1**). Notably, the Discovery and Define phases seemed especially universal amongst the research participants, especially as it pertained to observing how users engaged with their products.

For example, when asked about his work process, the Director of Curriculum Development (a SME) said, "I would visit a classroom that was teaching our curriculum. … I

care about what [the instructors] were doing with our products, and how they're being taught, and are [they] using all of the things we provide them?"

That response was very similar to what the UX Researcher (a design practitioner) had to say about her ideal work process. She said, "When I first come up with a [user testing] script, [I would like to] test it out on a couple kids or a couple [users] of my target market….That way, when you get all of your questions and everything together you can make changes, 'Oh, we don't need this [question],' 'Oh, we can probe more on this one.'"

From a Values perspective, there was a strong desire amongst the participants to work with users as co-designers, even amongst the non-designers. This leads the authors to believe that practitioners who have not been educated in child-specific techniques are not inherently disqualified from designing products for children. What this does mean, however, is that the challenge of working with children largely lies within aspects of legal or financial risk and the compliance issues that come with it; it is less of a design challenge.

Using the NAEP Assessment as an example, individuals who will directly interface with minors are required to undergo Security Clearance applications with the U.S. Government (Geary 2019). As of today, these applications require U.S. Citizenship, background checks performed by federal investigators going as far back as 7 years (including but not limited to personal, academic, and professional connections), etc. None of these things speak to any sort of design qualification or expertise, beyond making the U.S. Government feel secure in their levels of risk. In the case of academia and professional institutions, researchers are beholden to specific policies and training which may similarly inhibit working with children (Duke HR 2012).

Finding #3 – What unique activities do practitioners who work with children perform?

The activities described by the research participants were similar amongst them all. They all generally attended design stand-ups, read a lot of emails, and conducted user studies when possible. As products become increasingly built using processes like Agile and Scrum development—which call for the openness and inputs of designers, developers, product owners, stakeholders, etc— these various constituent partners need to feel like they can speak to one another (Schwaber and Sutherland 2020). Fortunately, the common activities between these practitioners reveal the potential for compatibility of individuals across various disciplines.

Unfortunately, the reality revealed through these interviews is that many designers and accessibility practitioners are siloed off from each other organizationally within their companies. This creates a dynamic in which these compatible practitioners will never meet with each other. When probed further, the research participants acknowledged that even at huge, multinational companies (e.g., Fortune 500 companies), Accessibility teams had no designers in their direct organization chart, despite these teams being largely responsible for products that needed to be designed.

Additionally, through these interviews, an originally unforeseen party to the child UX design process was revealed: the content writer, as was exemplified by the Manager of Technical Content Design. This role is significant in that it does not exist in every organization, yet 10 out of 13 participants (all except for the Director of UX Design, the UX Researcher, and the Sr. Software Architect) stressed that the difference between designing for adults and children lie in a product's content (see **Figure 3.2**).

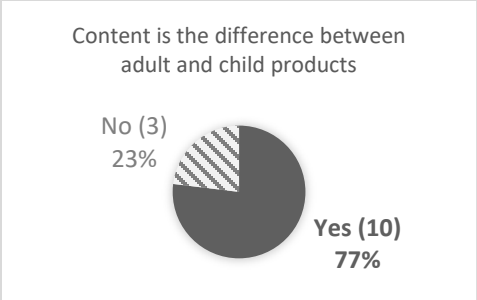

Figure 3.2 - Participant sentiment about the significance of content to the design of child products vs the design of adult products

Between these 10 participants, different aspects of content were called out. For example:

- The Certified Teacher of the Deaf stressed that she sought to compensate for some of her students' lack of incidental knowledge around English or signed content.
- The Content Manager and the Lead R&D UX Researcher stressed the importance of age-appropriate content, especially in parts of the user interface which may get overlooked as simply software engineer-managed elements.
- The "Senior Software Engineer & Accessibility Designer" and the UX Lead for Accessibility talked about making sure that content is captured in alternative formats (e.g., signed video, closed captions, etc.).

These examples of differing types of content do not appear to be deliverables that are generally attributable to designers, however.

As such, these participant responses infer that designers who do not have experience or education specific to designing for children can still be qualified for this work. They can learn on the job so long as they partner with experts in this space—e.g., accessibility specialists, engineers, content managers, or other experienced designers. If such experts are not present

within their immediate organization chart within their employer company, it is imperative that they demand interdisciplinary communications with others who have experience working with these child users. In due time, these practitioners can -in turn- become the de facto experts in this space and who should continue the mentorship cycle and collaborate with emerging practitioners who are new to the space of designing for children.

This process is one which is being employed by the World Wide Web Consortium (W3C) and came up in the research interviews. For example, in 2019, there was a call for participants for a new working group called Accessibility for Children. This group is composed of experts who have varying levels of experience with either design, child education, development and/or some mix of those or other skills. Through interviewing members of this working group, it was revealed that no singular person has all the answers related to designing for child accessibility, and yet this group has been able to provide accepted input to shape the standards which may one day make it into the next update of the Web Content Accessibility Guidelines, version 3.0. (Taylor & Stiernet, 2022). As an exemplar for this interdisciplinary model of spreading child design education, the Accessibility for Children Community group has recently published a Frequently Asked Questions document which was presented at TPAC 2022, W3C's annual conference. Building off that momentum, this group has also contributed to other W3C task forces, like the Pronunciation Task force (W3C 2022).

Unfortunately, like most of the interviewed participants, many of the members of this Accessibility for Children working group have not directly interacted with children in their work capacity for years. That said, these practitioners' past experiences still allow them to intuit how children may respond to the various products and guidelines that are referenced.

Part III. Post-Research Discussion: Opportunities for future research

As a result of these interviews, a few themes emerged which suggest that further in-depth study would be beneficial to the study of UX Design for children. These themes include:

1. The juxtaposition between the techniques employed by designers for English Language Learners (ELL) and the opportunities for equitable design for Deaf and Hard-of-Hearing (DHH) children;
2. The application of existing research into marginalized communities to DHH children; and
3. The value of open-source design and development to the advancement of equitable products for DHH children.

Opportunity 1 - Comparison between ELL and DHH design for children

For those unaware, American Sign Language (ASL) is its own complete language. It is insufficient to assume that closed captions are a wholly sufficient substitute for designing natively in ASL. This was raised by one of the research participants. To that end, it is worth considering the application of English Language Learner (ELL) techniques, for example those used for Spanish-speaking children, with DHH children. This changes the conversation from the oft-quoted statistic that most age-14 DHH students (80^{th} percentile or lower) on average place below a grade-4 reading level. The conversation instead becomes less of a statement about DHH students' low reading level, and more about questioning how digital products, especially digital assessments, have failed them.

Any researcher seeking to take up this research opportunity is recommended to read Holcomb et al.'s *ESL Literacy for a Linguistic Minority: The Deaf Experience (1992)*. It focuses

on providing possible explanations for DHH students' literacy difficulties. It also seeks to recontextualize these students' experiences from one of disability to one of bilingual/bicultural approaches.[4]

Opportunity 2 - Application of Marginalization Theories to DHH children

When the research participants were asked about what design principles they were aware of, one participant mentioned Minority Stress Theory. In the context of DHH child learners, it could provide an explanation for the disparity between the English reading proficiency of DHH children vs non-marginalized children. The idea behind this theory is that the stress placed upon a minority group is "unique…, chronic (related to social and cultural structures), and socially based (social processes, institutions and structures)" (Denato 2012).

Minority Stress Theory alone, however, may be insufficient to improve child students' ability within the classroom. It may need to be paired with similar theories, such as Critical Social Theory which has a goal of "advancing the emancipatory function of knowledge" (Leonardo 2004). The idea here is to not downplay the experiences of DHH children, but to instead provide a framework for which designers and those influencing the design of products can identify and dismantle the power structures that inhibit DHH children.

Further study into this research opportunity may open additional paths to empathizing with the DHH experience and may also spark research into the experience of other less researched disabilities.

Opportunity 3 - Effect of Open-Source Design and Development for DHH child products

This last research opportunity is one that may prove contentious in professional industries: unapologetically open-source design and development. Affirmed by these research

[4] According to Holcombe et al, Deafness was considered a pathological condition for centuries (i.e., a disease).

interviews, designers are held to contractual clauses that forbid them from retaining any work product they develop while employed. Assessment companies are no exception to this process. For example, one of the challenges with some digital assessments is also in their proprietary nature. For some assessment companies, their test items are closely guarded secrets under constant surveillance. Due to this secretive nature, it becomes more challenging to ensure that these items go through the same rigor that publicly available test items might be afforded.

One example of the alternative is the National Assessment of Educational Progress (NAEP). Known as the largest nationally representative assessment of American students' proficiency, NAEP has been around since 1969 (NCES 2019). What makes NAEP different from other assessments is that it releases a portion of their test items to the public after a certain number of years. While publishing test items is certainly valuable, without the larger context of the assessment user interface, the full test taker experience is not able to be examined.

For hearing students, the risk of public scrutiny is partially addressed by the publication of freely-available NAEP Tutorials and the open-source publication of the eNAEP platform delivery style guide, which provides some exemplary interaction types. These two open-source products were referenced by a few of the research participants. Unfortunately for Deaf students, the specialized version of the NAEP tutorial that allows for special functionality (like pausing the auto-playing segments so that sign language interpreters can translate audio instructions before they leave the screen) is not publicly available.

During the interviews, the participants lamented that external stressors like government approval and corporate secrecy (e.g., non-disclosure agreements) were detrimental to their work process. They went on to say that they were unable to keep NAEP's open-source style guide up to date, because it was incompatible with their employers' attempt at a next-generation design

system. This next-generation design system was conceived as being inherently more accessible than its predecessor style guide, but corporate interests prevented it from being shared as an open-source product. These points further exemplify the power struggle between competitive advantage and the advancement of design knowledge.[5]

Conclusion

In summary, UX Design for DHH children remains a critically underserved specialization of UX Design. As the overall field of UX Design seeks to define itself, design practitioners and subject matter experts have come to doubt the relevancy of their expertise to this emerging specialization.

After interviewing 13 practitioners in the space of UX, Accessibility, and Education it was revealed that the design process for creating products for children is not radically different than the design process employed in the creation of products for adults or other marginalized groups. The common process entails collecting background knowledge, forming hypotheses, testing designs, synthesizing feedback, and iteration.

The primary difference between designing for these different user groups (adults, children, other disabilities) is content, which is not uniquely specific to the design profession. For example, for fourth graders to be successful, the content needs to be written at that age-level; in the case of middle-school students in the United States, their assessments should be written at the sixth- to eighth-grade levels.

Lastly, to lend confidence to designers for current or emerging designers of products for children, it is important to allow other designers the opportunity to review their work with other practitioners. The challenge with this lies in how to provide this external review in a way that

[5] An anonymous peer reviewer suggests that this is a good point, and is an issue with learning on the job, hence the argument for degree or accredited programs that teach using case studies, collaboration with experts and users.

does not compromise the trade secrets of companies that are potentially competing for the same expertise and market share.

Fortunately, this research's findings support the idea that designing for DHH children does not need a radical overhaul of how UX design is taught. It does, however, require providing a structure that allows these designers to succeed, e.g. publicly available design examples, content writers, and the ability to user test with children. None of these structural elements are trivial, but they do cast a spotlight on the need for designers to become increasingly interdisciplinary in how they move forward with designing products for children, especially for Deaf or Hard-of-Hearing children in the United States.

References (Chapter 3)

Ali, Z. (2022, Feb 23). As Diversity, Equity & Inclusion Hits $9.3 Billion in Global Spending, Watch Out for These Key Trends in 2022. Retrieved from https://www.prnewswire.com/news-releases/as-diversity-equity--inclusion-hits-9-3-billion-in-global-spending-watch-out-for-these-key-trends-in-2022--301487159.html

CHM. (2017). ALAN COOPER. Retrieved from Computer History Museum: https://computerhistory.org/profile/alan-cooper/

Denato, M. P. (2012, April). The minority stress perspective. Retrieved from American Psychological Association: https://www.apa.org/pi/aids/resources/exchange/2012/04/minority-stress

Duke HR. (2012, May 16). Policy for Minors in Duke University Programs. Retrieved from Duke.edu: https://forms.hr.duke.edu/minors/training/training_workbook.pdf

ECDC. (n.d.). Practicum and Course Related Experiences. Retrieved 2023, from Notre Dame.

Geary, D. B. (2019, February 21). Federal Contract Opportunity for National Assessment of Educational Progress (NAEP) RFI 91990023Q0002. 215. U.S. Department of Education. Retrieved from https://govtribe.com/file/government-file/naep-alliance-2019-2024-pws-dot-pdf

Giannakos, M., Papamitsiou, Z., Markopoulos, P., Read, J., & Hourcade, J. P. (2020, June). Mapping child–computer interaction research through co-word analysis. Retrieved from Elsevier: https://www.sciencedirect.com/science/article/pii/S2212868920300015

Hewett, T. T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., . . . Verplank, W. (1992, January). ACM SIGCHI Curricula for Human-Computer Interaction. Retrieved from ACM: https://dl.acm.org/doi/book/10.1145/2594128

Hoffman, R. R., Crandall, B., & Shadbolt, N. (1998, June 1). Use of the Critical Decision Method to Elicit Expert Knowledge: A Case Study in the Methodology of Cognitive Task Analysis. Retrieved from SAGE journals: https://journals-sagepub-com.prox.lib.ncsu.edu/doi/abs/10.1518/001872098779480442

Holcomb, T., & Kreeft, J. (1992). ESL Literacy for a Linguistic Minority: The Deaf Experience. ERIC Digest. (ED353861). Retrieved from Ed.gov: https://files.eric.ed.gov/fulltext/ED353861.pdf

Hourcade, J. P., Revelle, G., Zeising, A., Iversen, O. S., Pares, N., Bekker, T., & Read , J. C. (2016, May). Child-Computer Interaction SIG: New Challenges and Opportunities. Retrieved from ACM: https://dl.acm.org/doi/abs/10.1145/2851581.2886433

IDF. (n.d.). What is User Experience (UX) Design? Retrieved from Interaction Design Foundation.

Lehnert, F. K., Niess, J., Lallemand , C., Markopoulos, P., Fischbach, A., & Koenig, V. (2022). Child–Computer Interaction: From a systematic review towards an integrated understanding of interaction design methods for children. International Journal of Child-Computer Interaction. doi:https://doi.org/10.1016/j.ijcci.2021.100398

Leonardo, Z. (2004, August). Critical Social Theory and Transformative Knowledge: The Functions of Criticism in Quality Education. doi:https://doi.org/10.3102/0013189X0330060

Maguire, M. (2022). Interface Design and Usability Evaluation of Voice-Based User Interfaces. In M. M. Soares, F. Rebelo, & T. Z. Ahram, Handbook of Usability and User-Experience (pp. 17-25). Taylor & Francis Group. doi:https://doi-org.prox.lib.ncsu.edu/10.1201/9780429343513

Markopoulos, P., & Bekker, M. (2003, April). On the assessment of usability testing methods for children. Retrieved from ScienceDirect: https://www.sciencedirect.com/science/article/pii/S0953543803000092

NCES. (2019). NAEP History and Innovation. Retrieved from Ed.gov: https://nces.ed.gov/nationsreportcard/about/timeline.aspx

NNgroup. (2016, July 2). Don Norman: The term "UX". Retrieved from YouTube: https://youtu.be/9BdtGjoIN4E

Norman, D. (1988). The psychology of everyday things. Basic Books.

Pizzo, L., & Chilvers, A. (2016). Assessment and d/Deaf and Hard-of-Hearing Multilingual Learners: Considerations and Promising Practices. Retrieved from American Annals of the Deaf: http://www.jstor.org/stable/26235251

Read, J., & Bekker, M. M. (2011). The nature of child computer interaction. BCS-HCI '11: Proceedings of the 25th BCS Conference on Human-Computer Interaction. ACM.

Rosala, M. (2020, March 15). The Discovery Phase in UX Projects. Retrieved from Nielsen Norman Group: https://www.nngroup.com/articles/discovery-phase/

Saldaña, J. (2013). Values Coding. In The coding manual for qualitative researchers (2 ed., pp. 110-115). SAGE.

Schwaber, K., & Sutherland, J. (2020). The Scrum Guide. Retrieved from ScrumGuides.org: https://scrumguides.org/scrum-guide.html#scrum-values

Smith, K. A. (2016, May 21). Recruiters requesting a UX portfolio can cause a breach of Intellectual Property Rights. Retrieved from LinkedIn: https://www.linkedin.com/pulse/recruiters-requesting-ux-portfolio-can-cause-breach-property-smith/

Taylor, H. (2006). Eliciting Tacit Knowledge Using the Critical Decision Interview Method. In M. E. Jennex, Knowledge Management in Modern Organizations (pp. 285-301). Idea Group, Inc.

Taylor, S., & Stiernet, M. (2022, March). RESULTS OF QUESTIONNAIRE COMMUNITY GROUP PRIORITIES. Retrieved from W3C.org: https://www.w3.org/2002/09/wbs/119612/poll_number_one/results

Traxler, C. B. (2000, September 01). The Stanford Achievement Test, 9th Edition: National Norming and Performance Standards for Deaf and Hard-of-Hearing Students. Retrieved from The Journal of Deaf Studies and Deaf Education: https://doi-org.prox.lib.ncsu.edu/10.1093/Deafed/5.4.337

Verplank, B., Sutcliffe, A., Mackay, W., Amowitz, J., & Gaver, W. (2002). DIS '02: Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques. DIS: Designing Interactive Systems. Association for Computing Machinery. doi:https://dl.acm.org/doi/proceedings/10.1145/778712

W3C. (2022, June 9). Github issues toward including Children's Needs in W3C Publications. Retrieved from W3C.org: https://www.w3.org/community/accessibility4children/wiki/Github_issues_toward_including_Children%27s_Needs_in_W3C_Publications

Walker, R. (2021, May 7). The Practicum Course: What it means to work and study all at once. Retrieved from Meet NYU: https://meet.nyu.edu/academics/the-practicum-course

Zannier, C., Chiasson, M., & Maurer, F. (2007, June). A model of design decision making based on empirical results of interviews with software designers. Retrieved from ScienceDirect: https://www-sciencedirect-com.prox.lib.ncsu.edu/science/article/pii/S0950584907000122#bib37

BRIDGE TO CHAPTER 4

Upon further review of the transcripts from the first round of interviews, the author came to realize that there may have been a researcher-research participant misalignment. The author's intention was to discover a set of guidelines that could help designers create products for DHH children, but the probe asked the participants to describe "design principles" that they used in their profession (Polanco Jr and Liu 2024).

To mitigate any lingering doubts, the author re-interviewed the same pool of participants and prompted them to describe what guidelines meant to them from the most general sense of the word down to their opinion on what they would expect to see out of Deaf and Hard-of-Hearing Middle-school Assessment Design Guidelines (DHH-MSAD). This second round of interviews reaffirmed the need for DHH-MSAD, reaffirmed that a multi-disciplinary approach is needed to create DHH-MSAD, and revealed biases that both designers and subject-matter experts have with regards to who they would or would not engage when creating design guidelines for Deaf or Hard-of-Hearing students.

The following paper "Multidisciplinary Framework for Creating the Next-Generation of Human-Centered Design Guidelines," is the result. It was accepted as both a Conference Paper and Conference Proceeding for the AHFE International Conference on Human Factors in Design, Engineering, and Computing (AHFE 2023 Hawaii Edition) in December 2023. As of June 2024, the paper has been visited 177 times and downloaded 82 times.

CHAPTER 4 MULTIDISCIPLINARY FRAMEWORK FOR CREATING THE NEXT-GENERATION OF HUMAN-CENTERED DESIGN GUIDELINES

ABSTRACT

Human-centered design for students with disabilities is an inherently multidisciplinary endeavor. It requires input from practitioners in the fields of user experience (UX), instructional design, accessibility, software engineering, and more. To that end, this paper seeks to reveal the perspectives of practitioners who have experience across these various disciplines, especially as they pertain to Deaf and Hard-of-Hearing (DHH) students.

In a previous study, the author asked practitioners to describe the design guidelines they use in their professions. To contextualize the concept of guidelines, the research participants were introduced to North Carolina State University's Principles of Universal Design and Jakob Nielsen's 10 Usability Heuristics. The outcome of that study revealed a wide spectrum of guidelines, with little commonality except for the Web Content Accessibility Guidelines and Universal Design.

This paper builds off that prior research by interviewing the same practitioners with a renewed focus on the definition of guidelines. In particular:

1. How do these practitioners define "guidelines", and do these definitions vary between the different disciplines of user experience, accessibility, etc.?
2. Do the practitioners' definitions of "guidelines" change as the end-user becomes more specifically designed? For example, do practitioners expect more, less, or something different when designing for hearing students vs DHH students?

3. If these practitioners were empowered to create a theoretical set of design guidelines for DHH middle school-aged students, what process would they follow?

The results of this new study demonstrate an awareness in the research participants that experts from different disciplines may need to be engaged in the process of creating design guidelines. More specifically, when asked to create DHH guidelines, the SMEs were often quick to say that they needed to collaborate with individuals who were embedded in the DHH community, i.e., individuals who are Deaf or Hard-of-Hearing, or individuals who have acquired second-hand knowledge through being immersed in Deaf culture.

Surprisingly, when interviewing these practitioners for a second time, designers were not always explicitly stated as integral to the creation of design guidelines for DHH middle school students. A few reasons emerged for why this is the case:

A. On multiple occasions, the research participants lamented that design education today often fails to prepare new designers in the space of accessibility.
B. In another instance, a senior software architect felt that engaging designers would be detrimental to the guideline creation process. He would later amend his response to say that he would include a designer only if they had a sufficient background in accessibility.
C. In a different instance, a designer took designer participation for granted and assumed that designers would obviously be part of the design guideline process. When asked to clarify why he did not include designers in his process when first asked, he said that his ego led him to assume that he would naturally be part of the process.

While both this paper and the previous paper focus on the DHH student experience, the overall goal of this research is to create a reusable framework that will allow subject matter experts (SMEs) to examine the guidelines they use today, and to learn how to create the next generation of design guidelines that will rise to the challenge of human-centered design becoming increasingly multidisciplinary.

Keywords: Design, user experience, Accessibility, Deaf, Hard-of-Hearing, guidelines, framework

INTRODUCTION

The author's initial research goal was the creation of guidelines for practitioners who were involved in the design of products for Deaf and Hard-of-Hearing (DHH) children. This goal was formed in response to the seemingly non-existent research into Deaf and Hard-of-Hearing children, despite marked increases in the Deaf child population every year (Pizzo and Chilvers 2016). To provide context on this growing population group, between 2005 and 2020, the United States Center for Disease Control and Prevention reported a 239% increase in the number of babies born Deaf or Hard-of-Hearing who were identified early (CDC 2022).

To accelerate the creation and release of much-needed DHH child design guidelines to the public, the author sought to narrow his focus to the field of Deaf and Hard-of-Hearing assessment design. This resulted in the paper, "Uncovering an Inclusion Gap in the Design of Digital Assessments for Middle school-aged Deaf and Hard-of-Hearing Students in the United States". (Polanco Jr and Liu 2023)

That paper revealed that leading testing institutions are prone to pretesting their assessment items without including students with disabilities for reasons including, but not limited to, the fact that these populations often fall below the statistical thresholds that would make them candidates for pretesting candidate selection, especially when using oft-quoted rules of thumbs for these activities (ETS 2020; Pearson, 2018). For example, it's known that interviewing at least five people reveals most usability problems (Nielsen 2012). Given that the number of students in the United States with disabilities generally fall below 20% of the total U.S. student population, it's very likely students with disabilities will he wholly omitted from the design process, unless they are explicitly sought out (NCES 2020).

Choosing to focus on the Deaf and Hard-of-Hearing student experience, the author interviewed 13 practitioners in the fields of user experience (UX), accessibility, and education and had them self-identify their roles, their work process, their familiarity with both DHH (Deaf and Hard-of-Hearing) users, and their professional opinion on the difference between designing for adults and designing for children.

This resulted in the paper "UX Design for children: How is it done now? Should it change?" (Polanco Jr and Liu 2024). At a high-level, three outcomes emerged from that study:

1. Designers questioned how relevant their experience is to subject matter experts in child computer interaction; subject matter experts doubted the value of their expertise in the design of products for children.
2. The design process for practitioners who worked with children did not vary significantly from those who only had experience designing for adults.
3. The primary difference between products for children and adults is the age-level appropriateness of the content. As such, it is of the utmost importance that at least one practitioner in the child-product creation process includes a content specialist who understands the difference between content for adults and content for children of different age-levels.

STUDY DESIGN

In the author's previous study, "UX Design for children", 13 practitioners in the space of user experience, accessibility, and education were interviewed for 45-to-60 minutes on the topic of digital assessment for Deaf and Hard-of-Hearing students. This study took place over two years and used social media, email, and word-of-mouth recruitment techniques. Recruitment

proved especially difficult, with a known factor being that participants were not paid for volunteering their time. Financial incentives would have resulted in significant increases in research participation, but the author's research committee was uncertain of the effects of incentives on the participants' responses (Abdelazeem et al. 2022).

The questions asked were as follows:

Interviewee Classification

- What is your job title?
- How do you define your field of expertise?
- Who do you design for?
    - If not children, do your products get used by children? Why or why not?
    - If age is not specified, do you think middle school-aged students use your products?

Work Process

- What design principles are you aware of and use in your profession?
- How would you describe your daily tasks?
- Does your work affect a Deaf or Hard-of-Hearing audience?

Child-computer Interaction

- Does your work process vary based on whether your target is an adult or child?
    - If yes, how does your design process vary from adult-targeted products to child-targeted products?
    - If no, why doesn't your design process vary from adult-targeted products to child-targeted products?
- What would you change about your current process, if anything, to suit child audiences?

Deaf or Hard-of-Hearing (DHH)

- Does your work process vary based on whether your target audience is Deaf or Hard-of-Hearing? If yes, how? If no, why not?

Evaluation of the interview transcripts revealed the possibility that one of the interview questions may have misled the participants. Notably, one question asked what design principles the practitioners were aware of and used in their profession. According to Merriam Webster (n.d.), a principle is a "comprehensive and fundamental doctrine, or a fact of nature" (para. 1a). This word was chosen as a means of juxtaposing the practitioners' domain expertise with

published works like North Carolina State University's Principles of Universal Design and Jakob Nielsen's 10 Usability Heuristics for User Interface Design (Mace et al. 1997; Nielsen 1994). Unfortunately, this may have been a leading question— which, in turn, caused the SMEs to doubt the relevancy of their domain expertise, as described in Outcome #1 of "UX Design for children".

In order to A) address this potential flaw of the previous study and to B) return to the author's initial research goal of creating guidelines for practitioners who are involved in the design of products for Deaf and Hard-of-Hearing children, the author designed a follow-up study which is described below:

Seeking out the previous 13 practitioners in the space of user experience, accessibility, and education, the author re-interviewed these research participants for 30-minutes using a very narrow set of probes which solely focused on the concept of human-centered design guidelines.

The interview questions existed on a scale of semantic probing which started with high-level questions about the participants' knowledge of guidelines and gradually worked towards domain-specific questions which sought to uncover the participants' knowledge and opinions about the ultimate target of the research study—e.g., Design guidelines for middle school-aged Deaf and Hard-of-Hearing students. The interview transcripts were then coded inductively using Values Coding in order to determine the values, attitudes, and beliefs of the participants. (Saldaña 2013, p. 111).

The author's goal with this methodology was to better understand the types of information that practitioners in user experience, accessibility, and education might need to successfully design products for Deaf and Hard-of-Hearing students. Through this method, the author sought to create a reusable framework that will allow practitioners to examine the

guidelines they use today, and to learn how to create the next generation of multidisciplinary-informed design guidelines.

The probes were as follows:

- **Q1.** When you hear the word "guidelines", what type of information would you expect to see?
- **Q2.** If you knew that "guidelines" were created for designers, would you expect to see different information? If so, what type of information?
- **Q3.** During the [previous] interview, we talked about empowering designers who create products for Deaf children. If you were reviewing "Deaf Design Guidelines", would you expect to see different information? If so, what type of information?
- **Q4.** If you were responsible for creating guidelines for designing Deaf middle school assessments, what type of process would you follow?

Research Participants

In alignment with the previous study design of *"UX Design for Children",* interview participants were categorized into two classifications: designers and subject matter experts. The hypothesis was that interviewees of the same classification would provide similar responses to each other.

It should be noted, however, that this delineation of interviewee categories is not a perfect one. For example, one interviewee self-identified as both an Engineer and a Designer. However, his described daily actions leaned more heavily towards development, not design—as such, he was classified as a subject matter expert.

Additionally, out of the 13 participants who were interviewed in the previous study, only 12 were able to be interviewed for this study (**Table 4.1**). The participant not re-interviewed for this study was the Director of the Motion Light Lab. All interviewed participants chose to retain their previously self-identified job titles for the purposes of this study, despite some of them changing job titles and employers between the previous study and this current study. For example, the Director of User Experience changed employers and is now the Head of Accessibility within a new employer's design organization, but still self-identified as the former.

Table 4.1 - The distribution of the participants' self-identified job titles.

Design Practitioners	Subject-matter experts
• Lead R&D UX Researcher • Director of Curriculum Development • UX Researcher • UX Lead for Accessibility • Director of User Experience • Senior Accessibility Manager • Manager of Technical Content Design	• Sr. Software Architect • Research Scientist • Certified Teacher of the Deaf K-12 • Senior Software Engineer / Accessibility Designer* • Special Education Teacher
**Could be classified as design practitioner or SME due to providing a hybrid job title.*	

RESEARCH RESULTS

The responses to the research interview have been aggregated at the question-level and have similarly been delineated by classification type.

Q1. When you hear the word guidelines, what type of information do you expect to see?

Amongst the designers, five out of seven saw guidelines as being a list of requirements or things that should be done, usually to reach some level of compliance. Two of the seven designers (the Lead R&D UX Researcher and the Director of Curriculum Development) saw guidelines as a type of referential material or starting point that could inform the design process.

The majority, however, saw guidelines as being a list of requirements or things to be done, usually to achieve some level of compliance. Due to the question being intentionally general, six out of the seven designers did not mention a specific compliance entity or standards body was mentioned. The one remaining participant (the Senior Software Engineer / Accessibility Designer) quickly began to ground his response in the context of the Web Content Accessibility Guidelines (WCAG). In his opinion, WCAG appeared to be more of a list of things that should not be done, whereas he would prefer guidelines to be a list of things that should be done.

Unlike the designers, most of the subject matter experts felt that guidelines were simply recommendations or steps to be successful; only one SME, the Sr. Software Architect, alluded to guidelines being used to satisfy an administrative person (i.e., compliance).

Q2. When you hear that design guidelines were created for designers, would you expect to see different information?

When asked to describe what kind of information they would expect from a set of design guidelines, the consensus across the designers and SMEs was “it depends.” The variance in expectations was similar, but slightly different between some of the respondents.

Amongst the designers, the UX Lead for Accessibility, the Director of User Experience, and the Senior Accessibility Manager alluded to the idea that different disciplines of design may need different design guidelines. For example, it was posited that UX design guidelines may need information about functionality, but visual design guidelines may need to provide information about color contrast.

When the same question was asked of subject matter experts, all but one SME felt that more details should be provided for design guidelines than the previously nondescript / general guidelines. The specifics of these additional details differed from SME to SME and were further different than those provided by the designers.

Examples include: the Research Scientist believed that the skill level of a designer may necessitate additional visual aids to explain more complex concepts. The Sr. Software Architect, alternatively, expected to see a lot of user stories. His thinking was that general guidelines that simply say what to do or not to do are insufficient, because they do not explain the reasons why these guidelines exist and who they impact. The Certified Teacher of the Deaf offered up a suggestion for additional information that no other practitioner offered: culture. Surprisingly, the Senior Software Engineer said that he expected no difference between design guidelines and general guidelines; he was the only respondent who expected no difference between general guidelines and design guidelines.

Q3. During the interview, we talked about empowering designers who create products for Deaf children. If you were reviewing "Deaf Design Guidelines", would you expect to see different information? If so, what type of information?

This third question gets closer to the author's expectations of what practitioners might be interested in from a hypothetical set of DHH Design Guidelines. Notably, both designers and SMEs made references to wanting to see information related to accessibility, with the UX Researcher making an explicit reference to WCAG.

Surprisingly, the Certified Teacher of the Deaf alluded to color contrast compliance as being necessary for a successful DHH student user experience. This was the first and only time

that visual design was described as an imperative for DHH users. She described how something as seemingly innocuous as the color difference between a scene (i.e., virtual backgrounds in video calls, or colored walls in physical spaces) and a signer's clothing could drastically inhibit a DHH student's comprehension of signed language, such as ASL. She would go on to express a desire to see this type of multi-disciplinary, accessibility-compliance thinking in the creation of Deaf Design Guidelines.

While not directly contradictory, it was an interesting juxtaposition to hear the UX Lead for Accessibility be dismayed that visual disabilities often appear to be overrepresented in today's accessibility guidelines and sought something different. In her estimation, Deaf Design Guidelines should be less concerned with how things look. It should be noted that of all the practitioners that were interviewed in this study, the UX Lead for Accessibility was the only interviewee who self-identified as being part of the DHH spectrum.

That said, the sentiment across both the designer and SME classification of participants was three-fold:

1. All but one SME expected different information from Deaf Design Guidelines than generic design guidelines.
2. They sought information that they could not intuit themselves without personally having a connection to the Deaf experience, and
3. There was a clear acknowledgement that the experience of being Deaf or Hard-of-Hearing is varied and should not be approached homogeneously.

As far as the first point goes, the one SME who did not expect different information was the Special Education Teacher. In this case, they did not challenge that different information

would be needed. Instead, they offered that it was outside the realm of their expectations and was emphatic that this was a question that DHH users should answer for themselves.

Regarding the second point, the Director of Curriculum development conveyed a desire to see something that could simulate the DHH experience, akin to how color blindness simulators exist for exploring designs for visually impaired users.

On the topic of the third point, the Sr. Software Architect wanted these Deaf Design Guidelines to either: A) have complete coverage for the entire range of the Deaf and Hard-of-Hearing experience, or B) if this was not possible, have the gaps in coverage be explicitly stated as a limitation of the guidelines. To that end, he believed that these theoretical Deaf Design Guidelines could serve as an exemplar for the creation and maintenance of future accessibility guidelines. This call for a gap analysis of the Deaf Design Guidelines and other accessibility guidelines was similarly echoed by another SME, the Research Scientist, who referred to this analysis as a "transparency statement." In their estimation, this would give greater credibility and validity to the Deaf Design Guidelines— as would use cases and exemplary designs. In these ways, both the Research Scientist and the Sr. Software Architect provided very similar responses.

Q4. If you were responsible for creating guidelines for the design of Deaf middle school assessments, what process would you follow?

Finally, with respect to this last question, every designer and SME stressed the importance of engaging experts who were knowledgeable of what it means to be Deaf or Hard-of-Hearing, either through first-hand experience or through being deeply embedded in the DHH community.

That said, while every practitioner alluded to engaging DHH experts, it is inconclusive as to whether every practitioner would work with DHH middle-school students directly. The interviewees were inconsistently probed to elaborate on what an expert meant to them. For examples:

- The Lead R&D UX Researcher stated that he took it for granted that designers would be considered amongst the list of possible DHH experts.
- The Director of User Experience and the Certified Teacher of the Deaf K-12 both stressed the importance of engaging with teachers who work with DHH students. The Director also claimed that every K-8 teacher she knew had to in essence become a designer out of necessity for meeting their students' needs.
- Only the Special Needs Teacher offered that she would personally work in a school for Deaf children so that she could directly ask DHH students what issues they face and what expectations they may have.
- Only the Manager of Technical Content Design described working with experts in assessment delivery platforms. Their reasoning was to facilitate stronger quality control through better understanding which assistive technologies are used within assessments.
- The Senior Software Engineer/Accessibility Designer uniquely felt like Deaf or Hard-of-Hearing Middle school Assessment Design (DHH-MSAD) Guidelines would be best expressed as a WCAG guideline. As a result, a DHH expert in his opinion would consist of an individual who is responsible for W3C (World Wide Web Consortium) standards, e.g. someone who is part of a W3C accessibility working group.

- When asked to elaborate, the Sr. Software Architect emphatically stated that he probably would not engage a designer in the creation of these DHH-MSAD Guidelines. He would later amend his response to say that he would only engage designers who were close to the DHH experience, but his experience is that designers tend to operate more from online sources versus first-hand experience with the communities they design for.

CONCLUSION

The author still believes that a set of design guidelines are necessary to ensure that equitable and human-centered approaches are used in the design of products for Deaf and Hard-of-Hearing users, especially in the design of middle school educational assessments. After interviewing 12 practitioners in the space of user experience, accessibility, and education for a second time, the author's belief in Deaf and Hard-of-Hearing Middle school Assessment Design (DHH-MSAD) guidelines remains steadfast. Unfortunately, the process by which these DHH-MSAD guidelines emerge still requires further investigation; notably because the amount and types of experts that need to be engaged to create these guidelines was not conclusively identified in this latest round of interviews.

Fortunately, the interviews proved that the process of developing DHH-MSAD requires a multi-disciplinary effort across multiple domains of expertise. To the extent that this result is clear, the author's research method of using a scale of semantic probing—which starts by eliciting general knowledge and works towards uncovering domain-specific outcomes—coupled with Values coding appeared to be successful.

To that end, this multidisciplinary framework may also prove to be compatible with other researchers who seek to develop the next-generation of human-centered design guidelines for other user populations.

REFERENCES (Chapter 4)

Abdelazeem, B., Abbas, K. S., Amin, M. A., El-Shahat, N. A., Malik, B., Kalantary, A., & Eltobgy, M. (2022). The effectiveness of incentives for research participation: A systematic review and meta-analysis of randomized controlled trials. PloS one. doi:https://doi.org/10.1371/journal.pone.0267534

CDC. (2022). EHDI Annual Data. Retrieved from CDC.gov: https://www.cdc.gov/ncbddd/hearingloss/ehdi-data.html

ETS. (2020). How Tests and Test Questions are Developed. Retrieved Oct 2020, from ETS.org: https://origin-www.ets.org/understanding_testing/test_development

Mace, R., Connell, B. R., Jones, M., Mueller, J., Mullick, A., Ostroff, E., . . . Vanderheiden, G. (1997, April 1). THE PRINCIPLES OF UNIVERSAL DESIGN. Retrieved from The Center for Universal Design: https://design.ncsu.edu/research/center-for-universal-design/

Merriam-Webster. (n.d.). Principle Definition & Meaning. Retrieved 2023, from Merriam-Webster.com dictionary: https://www.merriam-webster.com/dictionary/principle

NCES. (2020, May). The Condition of Education: Students With Disabilities. Retrieved from National Center for Education Statistics: https://nces.ed.gov/programs/coe/indicator_cgg.asp

Nielsen, J. (1994, April 24). 10 Usability Heuristics for User Interface Design. Retrieved from Nielsen Norman Group: https://www.nngroup.com/articles/ten-usability-heuristics/

Pearson. (2018). Evaluating performance. Retrieved from Pearson VUE: https://home.pearsonvue.com/Documents/Security/pearson_vue_test_items_US.aspx

Pizzo, L., & Chilvers, A. (2016). Assessment and d/Deaf and Hard-of-Hearing Multilingual Learners: Considerations and Promising Practices. Retrieved from American Annals of the Deaf: http://www.jstor.org/stable/26235251

Polanco Jr, A., & Liu, T. (n.d.). UX Design for Deaf children: How is it done now? Should it change?

Saldaña, J. (2013). Values Coding. In The coding manual for qualitative researchers (2 ed., pp. 110-115). SAGE.

CHAPTER 5 CONCLUDING DISCUSSION

This study sought to uncover a set of design guidelines (DHH-MSAD Guidelines) that could inform future design best practices in the space of DHH middle school assessment design. The research process began with an extensive literature review that encompassed the statistical relevance of the DHH population, educational assessments that have been and/or are currently given to middle school-aged students today, concepts for recognizing marginalized populations, and the origins of the knowledge disciplines that relate to DHH assessment—e.g. usability, child-computer interaction (CCI), accessibility, and education.

Summary. Prior to conducting research interviews, information from the U.S. National Center of Educational Statistics (NCES) on students with disabilities was gathered; and a case study of the National Assessment of Educational Progress (NAEP) tutorial was conducted. Through this analysis, the author discovered that DHH students were at high risk of not being represented in normative assessment research and, as a consequence, were also missing in the design process for testing instruments. This, in turn, results in DHH students being served inequitably by educational assessments, which were not designed with those students in mind. This assessment inequity would exacerbate poor research framing, and potentially lead to flawed but heavily cited studies that place DHH students at reading-levels far below that of their same-age hearing students.

Through the process of gathering research participants, the author noted that some subject-matter experts seemed skeptical that their interview responses could aid in the author's research goals. For example, the K-12 Certified Teacher of the Deaf thought her lack of design experience would make her a poor candidate for the research interviews. However, as was demonstrated in Chapter 3 Part II, she offered great insight into the space of DHH education, and

her responses are exemplary of the types of information that future designers need to be aware of when creating products for DHH child students.

Despite challenges that resulted in a protracted research recruitment process (September 2021 to November 2022), 13 interview participants were gathered. These participants had the requisite experience in usability, accessibility, and education to make them valuable respondents. Although none of the participants self-identified as experts in child-computer interaction, those participants who expressed apprehension about the relevance of their experience to the study were assuaged through the repeatable research framework that the author described in Chapter 3, Part II, Finding 1. This framework began by prompting the participants to self-identify their role and expertise, as a rapport-building tool that allowed participants to better juxtapose their own experience with the research questions.

Assumptions & Findings. Going into the study, the author believed there might be a difference in the work process used by practitioners who worked with DHH children vs. those who worked with adults only. Given Markopoulos and Bekker's third dimension for usability testing (Chapter 2, Section 2.4), the author thought that designers may need to employ different techniques to minimize negative influences on the design process. Surprisingly, this did not seem to be the case. Instead, 10 out of 13 participants felt that that content (e.g. age / grade-level vocabulary) was the deciding factor in adult vs child products. This outcome leads the author to strongly recommend that content specialists (like the Manager of Technical Content Design who was interviewed in this study) be included early and throughout the design process to ensure that digital assessments are maximally equitable for DHH child students.

Another finding that emerged is that there was little commonality in the design guidelines referenced by the various interview participants. In one instance a research participant, the

Director of the Motion Light Lab, stated through an interpreter that design guidelines did not exist when she was creating apps for Deaf readers; instead, she was creating guidelines for her specific product, in an ad-hoc manner. This finding was significant for multiple reasons: Firstly, this demonstrated a critical need for DHH guidelines that currently do not exist. Secondly, it led the author to believe that when experts in the field create their own guidelines out of necessity, these guidelines risk being hyper specialized to the expert's niche, might not be readily transferable to other applications, and thus might never be made available for general consumption. This supposition was confirmed when the Senior Accessibility Manager shared 1) that his role involves the use of many different principles and theories, and 2) he recently had to synthesize all these principles and theories into a consumable package so that another co-worker could perform their accessibility tasks. Third, this points to a need for a publicly accessible forum and/or repository for experts to share their findings as they pertain to DHH research. This would foster the type of multi-disciplinary collaboration that is needed to address the needs of the rapidly growing DHH child population. Next, when subject matter experts and designers were asked to elaborate on what types of information should be provided by guidelines, the difference in responses between the two groups was subtle. For example, designers saw guidelines as requirements, but subject-matter experts mostly saw them as recommendations. This subtle difference suggests that subject-matter experts and designers are compatible with each other, which should facilitate the type of collaboration that the author just described as being necessary in the previous paragraph.

Lastly, as the research probes became more specific (described in Chapter 4, "Study Design") there was a consensus between all participants that DHH design guidelines should 1) provide specialized knowledge, 2) that the information should not be composed of information

they could intuit themselves without being connected to the DHH experience, and 3) that the experience of being Deaf or Hard-of-Hearing should not be presented in a homogenous manner, so as not to diminish the value of DHH-MSAD. This phenomenon of consensus between SME and designer responses as the problem space became more specifically defined (i.e. when SMEs and designers were asked to focus specifically on the DHH child student experience) further affirms the author's recommendation that the design of digital assessments for middle school aged DHH students would benefit greatly from increased communication and knowledge sharing across the various disciplines of UX, accessibility, and education.

Limitations

This study was designed foremost to uncover guidelines and best practices for designing middle school assessments for Deaf and Hard-of-Hearing in the United States. While the author believes that the research framework utilized in this study could be widely applicable to study assessments administered more globally, this belief would need to be further verified through future research studies. At the moment, only one participant of this study was non-native to the United States (the Teacher of Disabled Students is a teacher in Poland). Another limitation of this study is that the author did not directly engage with DHH children directly, as is recommended in some CCI literature. This was a conscious decision by the author due to the logistical and legal challenges of recruiting and performing research with child participants in the United States.

Additionally, the lack of financial incentives (e.g. gift cards or cash payments) for participating in this research study may have vastly limited the candidate pool for research participants. Unfortunately, it is difficult to say in what way the candidate pool may have changed if incentives were provided.; it is possible that financial incentives may have resulted in

a larger number of candidates and/or a shorter recruitment period, but that is speculation. Lastly, the author does not use American Sign Language and required the use of a sign language interpreter when he interviewed a Deaf participant. Given that the research participant was permitted to provide their own interpreter, the author believes that the interpretation service utilized in this study was satisfactory. To further mitigate this limitation, transcripts were shared with the research participants via secure file transfer and the participants were allowed at any time to change, add, or redact information that their confidentiality, if that was their preference.

Future research opportunities

1. Exploring the connection between Marginalization and DHH child assessment.

One of the research participants, the Senior Accessibility Manager, listed Minority Stress Theory as one of the guidelines he uses to inform his work process. The author's literature review into this theory suggests that the link between marginalization and disability studies is an emergent field of study that may provide further insight into the field of DHH student child assessment. At the theory's core, the connection is not obvious.

First conceived by Meyer (2003), the Minority Stress Theory provides a framework for correlating the increased amount of physical and mental health problems that individuals of minority communities face in comparison to non-minority populations. Meyer rooted the theory largely in the experiences faced by sexual minorities (i.e. Lesbian, Gay, Bisexual, Transgender [LGBT]), but recent studies have sought to expand the scope of the theory to include other marginalized individuals. For example, Lund (2021) found that the minority stress provided a useful model for linking higher rates of suicide among people with disabilities, and

recommended research and policy to address ableism and disability-based discrimination. To address systemic ableism, Lund emphatically states the following:

> It is critical that psychologists interested in doing antiableist work listen to needs and concerns of colleagues, trainees, and community members with disabilities, especially those who are multiply marginalized…, and that they be willing to actively call out ableism in their colleagues and organizations to create an environment where disabled individuals feel free to speak. (Lund 2021, p, 188, "Systemic", para. 1)

The fervor with which Lund makes this statement mirrors the author's own sense of urgency that subject-matter experts and designers alike must engage with members of the Deaf and Hard-of-Hearing community so as to mitigate the inequities faced by DHH child students. Furthermore, while this study did not explore the effect of multiply marginalized children, the author sides with Lund that these children need additional attention when being designed for.

What exactly are marginalized communities? The verb marginalize is defined as "to relegate to an unimportant or powerless position within a society or group" (Merriam-Webster 2020). Marginalized populations have also been deemed synonymous with "vulnerable populations, oppressed populations, or undercounted populations" (Brutschy and Zachary 2014, p. 3771-3776). While there are clear overlaps between the two definitions, the lack of a definitive definition suggests to the author a need for additional research to prevent unintentional oversight of marginalized populations, which can have disastrous results–such as, failing to anticipate the needs of DHH children when designing an assessment (e.g. Chapter 1, Case Study of NAEP Tutorial).

An example where the two definitions may conflict is within the senior citizen population group (aged 65 or older). Accounting for about 16% of the total US population in 2018, senior citizens are slated to exceed the "Under 18" population group by 2035 (McPhillips 2019). If marginalization was solely measured by population count, senior citizens would no longer fall

under this classification once this population shift occurred. However, if systemic issues continued to force this population into a position of powerlessness, it could be argued that this demonstrated a qualifier for marginalization status. This argument against population-based marginalization holds even more true if we similarly look at US Census data. In 2022, female persons accounted for 50.4% of the US population and would be considered the majority group by the numbers, but at the same time Pew Research Center reported that this group was financially marginalized because they earned 82 cents per dollar as compared to male persons (Census.gov 2022, Kochhar 2023). As such, the author argues that majority-sized populations can still be considered marginalized due to historical and systemic injustice. This must hold true, because the size of the Deaf and Hard of Hearing population in the United States remains on the rise, but is still actively in need of advanced equity and support (CDC 2022).

Unfortunately, a definitive list of marginalized population qualifiers is not universally agreed upon. One example of qualifiers is provided by the US Department of State, which asks its grant applicants to demonstrate marginalized communities on the basis of age, disability status, ethnicity, religion, and sexual orientation. (US BUREAU OF DEMOCRACY, HUMAN RIGHTS, AND LABOR, 2014) Another set of qualifiers is provided by SAGE research methods, which includes financial status, political affiliation, or cultural affiliation-- though researchers acknowledge that even this is not an exhaustive list of factors (Given 2008). The author recommends that future research be specifically focused on disability-based marginalization. (Note that due to the lack of research of children with disabilities vs the plethora of research on adults [see Chapter 1, Background], the subsequent discussion will be primarily based on the experiences of adults with disabilities.)

Disability-based marginalization.

To further argue against purely numerical-based classification of marginalized groups, the author will look again to the senior citizen population. Within this population group, 49.8% of adults aged 75 or older self-reported having a disability (Bialik 2017). Despite their prevalence as a subgroup, these adults remain heavily discriminated against and systemically biased against, even to this day. To build context, before showing more examples, let us refer to *Olmstead v L.C. (1999),* wherein the US Supreme Court ruled that institutionalizing individuals with disabilities against their preference represented a serious form of discrimination. The majority opinion on this case was authored by the late Justice Ruth Bader Ginsburg and stated,

> States are required to provide community based treatment for persons with mental disabilities when the State's treatment professionals determine that such placement is appropriate, the affected persons do not oppose such treatment, and the placement can be reasonably accommodated, taking into account the resources available to the State and the needs of others with mental disabilities. (Olmstead v. L.C. 1999)

Despite this decision, disabled persons who depend on Medicaid for assistance still face a bias in favor of institutional care (e.g. nursing facilities). Under Medicaid, nursing facilities are a mandatory State-provided service, but home and community-based services (HCBS) (e.g. homemaker, home health aide, etc) are optional (Reaves and Musumeci 2015). Based on the *Olmstead opinion,* this gap in home-based care is a clear example of State-based discrimination. To make matters worse, this discrimination furthers inequity by restricting access to preventative care.

Another recent example of the marginalization of disabled peoples is the 2020 presidential election cycle (Wong 2020). Due to COVID-19 restrictions, visitors were disallowed by certain nursing homes, like The Davis Community nursing home in Wilmington, NC. In effect this prevented hundreds of thousands of disabled persons who required assistance with the

voting process from being able to cast their votes (McCarthy and Gillum 2020). Until August 27 of 2020, it was not known if voters with print disabilities (e.g. blind voters) would be able to privately vote in Virginia, because accessible, electronic absentee ballots were not an option (Gary et al v. Virginia Department of Elections et al 2020). It was only because a disability rights organization in Idaho wrote a letter that the public became aware of this issue. On August 21, the DisAbility Rights Idaho letter stated that IdahoVotes.gov remains incompatible with screen readers, leaving blind users unable to both "without the ability to register to vote or request an absentee ballot online" (DisAbility Rights Idaho 2020). That said, one of the most egregious examples of voter-based marginalization against disabled adults is the Mississippi Supreme Court ruling that which recently ruled that persons with pre-existing conditions (who were and still are at greater risk of contracting COVID-19) were not automatically granted absentee-ballot voting, and were expected to vote in-person (Gold 2020). These are just the examples of active endangerment and discrimination against disabled persons that made news headlines. The author surmises that there are many undocumented instances against children with disabilities that require research and investigation to uncover.

2. Increasing Digital Assessment Equity by Supporting ASL Video content.

Returning to the topic of DHH child students, the author believes that UI designers should explore novel methods of incorporating ASL into digital assessment content. For DHH students who use English as a second language, interfaces that lack ASL content risk further marginalizing these students. One approach to addressing this problem is to consider building in ASL video support early in the design discovery process (described in Chapter 3, Part I, "What is being taught to designers today?").

While it can be argued that ASL video support could affect the ability of certain reading constructs to be measured (described in Chapter 1, “Measuring Student Performance”), the author argues that the values of fairness and equity in assessment should take priority. Furthermore, assessment companies have already set precedents for feature incongruence between DHH and normative hearing students, as is demonstrated by the following examples:

Examples provided by the Smarter Balanced Assessment Consortium:

- **Text-to-speech:** Text is read aloud to the student via embedded text-to-speech technology. The student is able to control the speed as well as raise or lower the volume of the voice via a volume control.

- **Read aloud:** Text is read aloud to the student via an external screen reader or by a trained and qualified human reader who follows the administration guidelines provided in the Smarter Balanced Test Administration Manual and Read Aloud Guidelines. All or portions of the content may be read aloud. (SBAC 2020, p. 1)

Examples of universal design elements (UDEs) provided by NAEP:

- **Directions Read Aloud/Text-to-Speech (English):** All content in the assessment is text-to-speech enabled, including general directions, directions within the assessment, Tutorial, Help, and Student Questionnaire screens.

- **Read Aloud/Text-to-Speech (English) – Occasional or Most or All:** All content in the assessment is text-to-speech enabled. Follow subject-specific recommendations developed by NAEP Item Development and applied to the text-to-speech functionality for the NAEP DBA. Constructed responses added by students and the toolbar are not text-to-speech enabled. (NCES 2023, "Universal Design Elements")

To build a case for native ASL video support in digital assessment, the author recommends that future research be conducted to demonstrate the effect of ASL video support on digital assessment score. The author anticipates detractors to suggest that ASL video support may provide an unfair advantage to DHH students. In response to that hypothetical argument, the author counters that this support is more about leveling the playing field, so as to prevent digital assessments from marginalizing DHH students on the basis of their disability.

3. Additional Recommendations for Future Research

As described in Chapter 3, the author believes that this study could serve as a basis for the following future research opportunities:

- Comparing the English Language Learner literature with Deaf and Hard-of-Hearing literature (i.e. juxtapose how these population groups address concepts like student literacy; educational scaffolding techniques and pedagogy; and the historical and current inequities.)
- Further investigate the connection between Marginalization Theories (e.g. Minority Stress Theory) and students with disabilities.
- Explore the potential and real impact of open-source design and development to the enhancement and uplifting of people with disabilities.

Beyond these opportunities, the author hopes that other researchers can adopt the scale of semantic probing that was used in Chapter 4 to gather a larger and more encompassing research population. For example, this framework could be extended to research the needs of students with other disabilities (e.g. visual impairment, cognitive impairment) both within the United States and abroad.

In conclusion, this research began as investigation into the specific needs of Deaf and Hard-of-Hearing children because the author observed inequity in educational assessment in the United States. While the DHH-MSAD Guidelines did not emerge from this study, the confirmation that DHH studies is inherently multi-disciplinary; the discovery that experts in relevant disciplines either do not feel empowered and/or do not feel confident in collaborating with other disciplines; and the fact that these experts do not share a common set of design guidelines are all examples of new knowledge that have been added to the field of design.

COMPREHENSIVE REFERENCE LIST

Abdelazeem, B., K. S. Abbas, M. A. Amin, N. A. El-Shahat, B. Malik, A. Kalantary, and M. Eltobgy. 2022. The effectiveness of incentives for research participation: A systematic review and meta-analysis of randomized controlled trials. PloS one. doi:https://doi.org/10.1371/journal.pone.0267534.

Ali, Zak. 2022. As Diversity, Equity & Inclusion Hits $9.3 Billion in Global Spending, Watch Out for These Key Trends in 2022. Feb 23. https://www.prnewswire.com/news-releases/as-diversity-equity--inclusion-hits-9-3-billion-in-global-spending-watch-out-for-these-key-trends-in-2022--301487159.html.

Awopeju, O. A., E. R. I. Afolabi, and O. A. Opesemowo. 2017. "AN INVESTIGATION OF INVARIANCE PROPERTIES OF ONE, TWO AND THREE PARAMETER LOGISTIC ITEM RESPONSE THEORY MODELS." ProQuest. https://search-proquest-com.prox.lib.ncsu.edu/docview/2076932436/fulltextPDF/818A7C38E6CE466FPQ/1?accountid=12725.

Bialik, Kristen. 2017. "7 facts about Americans with disabilities." Pew Research Center. July 27. https://www.pewresearch.org/fact-tank/2017/07/27/7-facts-about-americans-with-disabilities/#:~:text=About%20half%20of%20Americans%20ages,said%20they%20had%20a%20disability.

Brutschy, Susan, and Deanna Zachary . 2014. "Marginalized Communities." In Encyclopedia of Quality of Life and Well-Being Research, by Alex C. Michalos, 3771-3776. Springer, Dordrecht. doi:https://doi.org/10.1007/978-94-007-0753-5_1725.

Camilli, G. 2017. "The Scaling Constant D in Item Response Theory." Scientific Research. https://www.scirp.org/html/3-1240938_79609.htm.

Campbell, Alastair, Charles Adams, Rachael Bradley Montgomery, and Kevin White. 2024. Introduction to Understanding WCAG. May. https://www.w3.org/WAI/WCAG22/Understanding/intro#understanding-the-four-principles-of-accessibility.

Casare, Andreia R, Celmar G da Silva, Paulo S Martins, and Regina L. O. Moraes. 2016. "Usability Heuristics and Accessibility Guidelines: a Comparison of Heuristic Evaluation and WCAG." ACM. April. https://dl.acm.org/doi/abs/10.1145/2851613.2851913.

CDC. 2022. EHDI Annual Data. https://www.cdc.gov/ncbddd/hearingloss/ehdi-data.html.

Census.gov. 2022. QuickFacts. https://www.census.gov/quickfacts/fact/table/US/LFE046221.

CHM. 2017. ALAN COOPER. https://computerhistory.org/profile/alan-cooper/.

Denato, Michael P. 2012. "The minority stress perspective." American Psychological Association. Psychology and AIDS Exchange Newsletter. April. https://www.apa.org/pi/aids/resources/exchange/2012/04/minority-stress.

DisAbility Rights Idaho. 2020. "DRI and the National Federation of the Blind of Idaho Issue Joint Statement Regarding Idaho Election Accessibility Concerns." DisAbility Rights Idaho. August 21. https://disabilityrightsidaho.org/.

Duke HR. 2012. "Policy for Minors in Duke University Programs." Duke.edu. May 16. https://forms.hr.duke.edu/minors/training/training_workbook.pdf.

ECDC. n.d. Practicum and Course Related Experiences. Accessed 2023.

ETS. 2020. How Tests and Test Questions are Developed. Accessed Oct 2020. https://origin-www.ets.org/understanding_testing/test_development.

Francis, James. 2017. "The Effects Of Technology On Student Motivation And Engagement In Classroom-Based Learning." Une.edu. June 1. https://dune.une.edu/cgi/viewcontent.cgi?article=1120&context=theses.

Gary et al v. Virginia Department of Elections et al. 2020. 1:20-CV-860 (UNITED STATES DISTRICT COURT FOR THE EASTERN DISTRICT OF VIRGINIA, August 27).

Gbollie, Charles, and Harriet Pearl Keamu. 2017. "Student Academic Performance: The Role of Motivation, Strategies, and Perceived Factors Hindering Liberian Junior and Senior High School Students Learning." Hindawi. Educational Research International. March 20. https://www.hindawi.com/journals/edri/2017/1789084/.

Geary, David B. 2019. "Federal Contract Opportunity for National Assessment of Educational Progress (NAEP) RFI 91990023Q0002." U.S. Department of Education, February 21. 215. https://govtribe.com/file/government-file/naep-alliance-2019-2024-pws-dot-pdf.

Giannakos, Michail, Zacharoula Papamitsiou, Panos Markopoulos, Janet Read, and Juan Pablo Hourcade. 2020. Mapping child–computer interaction research through co-word analysis. June. https://www.sciencedirect.com/science/article/pii/S2212868920300015.

Given, Lisa A. 2008. "Marginalized Populations." SAGE Research Methods. https://methods-sagepub-com.prox.lib.ncsu.edu/reference/sage-encyc-qualitative-research-methods/n252.xml?fromsearch=true.

Gold, Michael. 2020. "Mississippi Supreme Court says Covid-19 risk doesn't qualify voters for absentee voting." NY Times. September 18. https://www.nytimes.com/2020/09/18/us/elections/mississippi-supreme-court-says-covid-19-risk-doesnt-qualify-voters-for-absentee-voting.html.

Goman, Adele. 2016. "HOW MANY PEOPLE HAVE HEARING LOSS IN THE UNITED STATES?" Johns Hopkins Bloomberg School of Public Health. https://www.jhucochlearcenter.org/how-many-people-have-hearing-loss-united-states.html.

Harris, Jennifer L, Mohammed T. Al-Bataineh, and Adel Al-Bataineh. 2016. "One to One Technology and its Effect on Student Academic Achievement and Motivation." Ed.gov. https://files.eric.ed.gov/fulltext/EJ1117604.pdf.

Hewett, Thomas T, Ronald Baecker, Stuart Card, Tom Carey, Jean Gasen, Marilyn Mantei, Gary Perlman, Gary Strong, and William Verplank. 1992. "ACM SIGCHI Curricula for Human-Computer Interaction." ACM. January. https://dl.acm.org/doi/book/10.1145/2594128.

Hoffman, Robert R, Beth Crandall, and Nigel Shadbolt. 1998. Use of the Critical Decision Method to Elicit Expert Knowledge: A Case Study in the Methodology of Cognitive Task

Analysis. June 1. https://journals-sagepub-com.prox.lib.ncsu.edu/doi/abs/10.1518/001872098779480442.

Holcomb, Tom, and Joy Kreeft. 1992. ESL Literacy for a Linguistic Minority: The Deaf Experience. ERIC Digest. (ED353861). https://files.eric.ed.gov/fulltext/ED353861.pdf.

Hourcade, Juan Pablo, Glenda Revelle, Anja Zeising, Ole Sejer Iversen, Narcis Pares, Tilde Bekker, and Janet C. Read . 2016. "Child-Computer Interaction SIG: New Challenges and Opportunities." ACM. May. https://dl.acm.org/doi/abs/10.1145/2851581.2886433.

IDF. n.d. What is User Experience (UX) Design?

IEC, ISO, ITU. 2014. "IEC/ISO/ITU Policy on Standardization and accessibility." ISO.org. October. https://www.iso.org/files/live/sites/isoorg/files/archive/pdf/en/iec_iso_itu_joint_policy_statement.pdf.

ISO. 2018. "ISO 9241-11:2018(en) Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts." ISO.org. https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en.

Kirsh, David. 2019. "Do Architects and Designers Think about Interactivity Differently?" ACM Transactions on Computer-Human Interaction. April. https://doi-org.prox.lib.ncsu.edu/10.1145/3301425.

Kochhar, Rakesh. 2023. The Enduring Grip of the Gender Pay Gap. March 1. https://www.pewresearch.org/social-trends/2023/03/01/the-enduring-grip-of-the-gender-pay-gap/.

Le, Thi. 2020. US INDUSTRY (NAICS) REPORT 61171 Testing & Educational Support in the US. September. https://my-ibisworld-com.prox.lib.ncsu.edu/us/en/industry/61171/competitive-landscape.

Lehnert, Florence Kristin , Jasmin Niess, Carine Lallemand , Panos Markopoulos, Antoine Fischbach, and Vincent Koenig. 2022. "Child–Computer Interaction: From a systematic review towards an integrated understanding of interaction design methods for children." International Journal of Child-Computer Interaction. doi:https://doi.org/10.1016/j.ijcci.2021.100398.

Leonardo, Zeus. 2004. "Critical Social Theory and Transformative Knowledge: The Functions of Criticism in Quality Education." SAGE journals. August. doi:https://doi.org/10.3102/0013189X0330060.

Lollis, Jana, and Carol LaSasso. 2008. "The Appropriateness of the NC State-Mandated Reading Competency Test for Deaf Students as a Criterion for High School Graduation." The Journal of Deaf Studies and Deaf Education. June 05. https://doi.org/10.1093/deafed/enn017.

Lund, Emily M. 2021. "Examining the potential applicability of the minority stress model for explaining suicidality in individuals with disabilities." Rehabilitation Psychology 66 (2): 183–191. doi:10.1037/rep0000378.

Mace, Ron, Bettye Rose Connell, Mike Jones, Jim Mueller, Abir Mullick, Elaine Ostroff, Jon Sanford, Ed Steinfield, Molly Story, and Gregg Vanderheiden. 1997. "THE PRINCIPLES OF UNIVERSAL DESIGN." The Center for Universal Design. April 1. https://design.ncsu.edu/research/center-for-universal-design/.

Magno, Carlo. 2009. "Demonstrating the Difference between Classical Test Theory and Item Response Theory Using Derived Test Data." Ed.gov. April. https://files.eric.ed.gov/fulltext/ED506058.pdf.

Maguire, Martin. 2022. "Interface Design and Usability Evaluation of Voice-Based User Interfaces." In Handbook of Usability and User-Experience, by Marcelo M. Soares, Francisco Rebelo and Tareq Z. Ahram, 17-25. Taylor & Francis Group. doi:https://doi-org.prox.lib.ncsu.edu/10.1201/9780429343513.

Markopoulos, Panos, and Mathilde Bekker. 2003. On the assessment of usability testing methods for children. April. https://www.sciencedirect.com/science/article/pii/S0953543803000092.

Mayer, Richard E. 2001. 12 Principles of Multimedia Learning. https://www.hartford.edu/faculty-staff/faculty/fcld/_files/12%20Principles%20of%20Multimedia%20Learning.pdf.

McCarthy, Ryan, and Jack Gillum. 2020. "Hundreds of Thousands of Nursing Home Residents May Not Be Able to Vote in November Because of the Pandemic." ProPublica. August 26. https://www.propublica.org/article/hundreds-of-thousands-of-nursing-home-residents-may-not-be-able-to-vote-in-november-because-of-the-pandemic.

McKenzie, Lindsay. 2019. "The Digital Courseware Accessibility Problem." Inside Higher Ed. December 2. https://www.insidehighered.com/digital-learning/article/2019/12/02/professors-colleges-and-companies-struggle-make-digital.

McPhillips, Diedre. 2019. "Aging in America, in 5 Charts." US News & World Report. September 30. https://www.usnews.com/news/best-states/articles/2019-09-30/aging-in-america-in-5-charts.

Merriam-Webster. 2020. "Marginalize." Merriam-Webster. Oct 26. https://www.merriam-webster.com/dictionary/marginalize.

—. n.d. Principle Definition & Meaning. Accessed 2023. https://www.merriam-webster.com/dictionary/principle.

Michell, Ross E, Travas A Young, Bellamie Bachleda, and Michael A Karchmer. 2006. "How Many People Use ASL in the United States? Why Estimates Need Updating." Project MUSE. https://muse-jhu-edu.prox.lib.ncsu.edu/article/197164.

National Assessment Governing Board. 2022. Reading Assessment Framework for the 2022 and 2024 National Assessment of Educational Progress. September. https://www.nagb.gov/content/dam/nagb/en/documents/publications/frameworks/reading/2022-nagb-reading-framework-508.pdf.

National Deaf Center. 2019. Interpreters as a Reasonable Accommodation for Testing. https://www.nationaldeafcenter.org/sites/default/files/Interpreters%20as%20a%20Reasonable%20Accommodation%20for%20Testing.pdf.

NC State University. 1997. "THE PRINCIPLES OF UNIVERSAL DESIGN." The Center for Universal Design. April 1. https://projects.ncsu.edu/ncsu/design/cud/about_ud/udprinciples.htm.

NCES. 2023. Digitally Based Assessments UDE and Accommodation Descriptions. April 13. https://nces.ed.gov/nationsreportcard/about/accom_ude_descriptions.aspx.

—. 2019. "Digest of Educational Statistics." NCES. https://nces.ed.gov/programs/digest/d19/tables/dt19_105.20.asp.

—. 2019. NAEP History and Innovation. https://nces.ed.gov/nationsreportcard/about/timeline.aspx.

—. 2020. The Condition of Education: Students With Disabilities. May. https://nces.ed.gov/programs/coe/indicator_cgg.asp.

Nielsen, Jakob. 1994. "10 Usability Heuristics for User Interface Design." Nielsen Norman Group. April 24. https://www.nngroup.com/articles/ten-usability-heuristics/.

—. 1994. "10 Usability Heuristics for User Interface Design." NN/g Nielsen Norman Group. April 24. https://www.nngroup.com/articles/ten-usability-heuristics/.

—. 2012. How Many Test Users in a Usability Study? June 3. https://www.nngroup.com/articles/how-many-test-users/.

NNgroup. 2016. Don Norman: The term "UX". July 2. https://youtu.be/9BdtGjoIN4E.

Norman, Don. 1988. The psychology of everyday things. Basic Books.

Olmstead v. L.C. 1999. 98–536 (U.S. Supreme Court, June 22).

Pan, Jennifer. 2017. 508, ADA, WCAG: What's the difference? September 26. https://www.logicsolutions.com/508-ada-wcag-accessibility-difference/.

Pearson. 2018. Evaluating performance. https://home.pearsonvue.com/Documents/Security/pearson_vue_test_items_US.aspx.

Petrie, Helen, and Omar Kheir. 2007. "The relationship between accessibility and usability of websites." CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 397-406. doi:10.1145/1240624.1240688.

Pizzo, Lianna, and Amanda Chilvers. 2016. "Assessment and d/Deaf and Hard of Hearing Multilingual Learners: Considerations and Promising Practices." American Annals of the Deaf. Gallaudet University Press. http://www.jstor.org/stable/26235251.

Polanco Jr, Alexis, and Tsailu Liu. 2024. "UX Design for deaf children: How is it done now? Should it change?"

Qi, Sen, and Ross E. Mitchell. 2012. "Large-Scale Academic Achievement Testing of Deaf and Hard-of-Hearing Students: Past, Present, and Future." The Journal of Deaf Studies and Deaf Education,. Oxford Academic. June 28. https://academic-oup-com.prox.lib.ncsu.edu/jdsde/article/17/1/1/359085.

Read, Janet, and Mathilde M. Bekker. 2011. "The nature of child computer interaction." BCS-HCI '11: Proceedings of the 25th BCS Conference on Human-Computer Interaction. ACM.

Reaves, Erica L, and MaryBeth Musumeci. 2015. "Medicaid and Long-Term Services and Supports: A Primer." KFF.org. December 15. https://www.kff.org/medicaid/report/medicaid-and-long-term-services-and-supports-a-primer/.

Rodrigues, Filipa M., Ana Maria Abreu, Holmström Ingela, and Mineiro Ana. 2022. "E-learning is a burden for the deaf and hard of hearing." Scientific reports. doi:https://doi.org/10.1038/s41598-022-13542-1.

Rosala, Maria. 2020. The Discovery Phase in UX Projects. March 15. https://www.nngroup.com/articles/discovery-phase/.

Saldaña, J. 2013. "Values Coding." In The coding manual for qualitative researchers, 110-115. SAGE.

SBAC. 2020. "GUIDELINES FOR CHOOSING TEXT-TO-SPEECH OR READ ALOUD ACCOMMODATIONS FOR SMARTER BALANCED ELASUMMATIVE AND INTERIM ASSESSMENT READING PASSAGES FOR STUDENTS WITH DISABILITIES IN GRADES3-5." SmarterBalanced.org. March 5. https://portal.smarterbalanced.org/library/en/guidelines-for-choosing-tts-or-read-aloud-in-grades-3-5.pdf.

Schwaber, Ken, and Jeff Sutherland. 2020. The Scrum Guide. https://scrumguides.org/scrum-guide.html#scrum-values.

Section508.gov. 2023. Universal Design and Accessibility. August. https://www.section508.gov/develop/universal-design/.

Smith, Karl A L. 2016. Recruiters requesting a UX portfolio can cause a breach of Intellectual Property Rights. May 21. https://www.linkedin.com/pulse/recruiters-requesting-ux-portfolio-can-cause-breach-property-smith/.

Steyer, R. 2001. "Classical (Psychometric) Test Theory." In International Encyclopedia of the Social & Behavioral Sciences, by Neil J Smelser and Paul B Baltes, 1955-1962. Pergamon. https://www.sciencedirect.com/science/article/pii/B008043076700721X.

Taylor, Hazel. 2006. "Eliciting Tacit Knowledge Using the Critical Decision Interview Method." In Knowledge Management in Modern Organizations, by Murray E Jennex, 285-301. Idea Group, Inc.

Taylor, Suzanne, and Maud Stiernet. 2022. RESULTS OF QUESTIONNAIRE COMMUNITY GROUP PRIORITIES. March. https://www.w3.org/2002/09/wbs/119612/poll_number_one/results.

The College Board. 2015. "Test Specifications for the Redesigned SAT® ." Collegeboard.org. https://collegereadiness.collegeboard.org/pdf/test-specifications-redesigned-sat-1.pdf.

Traxler, Carol Bloomquist. 2000. "The Stanford Achievement Test, 9th Edition: National Norming and Performance Standards for Deaf and Hard-of-Hearing Students." The Journal of Deaf Studies and Deaf Education. September 01. https://doi-org.prox.lib.ncsu.edu/10.1093/deafed/5.4.337.

Trochim, M. K. n.d. Construct Validity. Accessed March 10, 2020. https://conjointly.com/kb/construct-validity/.

US BUREAU OF DEMOCRACY, HUMAN RIGHTS, AND LABOR. 2014. "The Bureau of Democracy, Human Rights, and Labor Potential Grant Applicants Workshop: Marginalized Populations." US Department of State. October 14. https://2009-2017.state.gov/j/drl/p/234757.htm.

US DOJ. 2012. Nondiscrimination on the Basis of Disability; Accessibility of Web Information and Services of State and Local Government Entities and Public Accommodations. Dec 7. https://www.ada.gov/anprm2010/web%20anprm_2010.htm.

US GSA. 2018. Applicability & Conformance Requirements. May. https://www.section508.gov/create/applicability-conformance.

Verplank, Bill, Alistair Sutcliffe, Wendy Mackay, Jonathan Amowitz, and William Gaver. 2002. "DIS '02: Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques." DIS: Designing Interactive Systems. Association for Computing Machinery. doi:https://dl.acm.org/doi/proceedings/10.1145/778712.

W3C. 2022. Github issues toward including Children's Needs in W3C Publications. June 9. https://www.w3.org/community/accessibility4children/wiki/Github_issues_toward_including_Children%27s_Needs_in_W3C_Publications.

—. 2018. "How WAI Develops Accessibility Standards through the W3C Process: Milestones and Opportunities to Contribute." W3.org. February 16. https://www.w3.org/WAI/standards-guidelines/w3c-process/.

W3C Team. 2019. "ACCESSIBILITY FOR CHILDREN COMMUNITY GROUP." W3C Community & Business Groups. December 6.

https://www.w3.org/community/accessibility4children/2019/12/06/call-for-participation-in-accessibility-for-children-community-group/.

W3C. 2023. What's New in WCAG 2.2. October 5. https://www.w3.org/WAI/standards-guidelines/wcag/new-in-22/.

Walker, Ryan. 2021. The Practicum Course: What it means to work and study all at once. May 7. https://meet.nyu.edu/academics/the-practicum-course.

Wong, Alice. 2020. "Voters With Disabilities Face Barriers in 2020 From COVID-19." Teen Vogue. September 16. https://www.teenvogue.com/story/voters-with-disabilities-barriers-covid.

Yang, Frances M, and T Kao Solon. 2014. "Item response theory for measurement validity." National Instutites of Health. June. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4118016/#B3.

Young, John W. 2009. Ensuring Valid Content Tests for English Language Learners. https://www.ets.org/Media/Research/pdf/RD_Connections8.pdf.

Zannier, Carmen, Mike Chiasson, and Frank Maurer. 2007. A model of design decision making based on empirical results of interviews with software designers. June. https://www-sciencedirect-com.prox.lib.ncsu.edu/science/article/pii/S0950584907000122#bib37.

Zanon, Cristian, Claudio S Hutz, Hanwook (Henry) Yoo, and Ronald K Hambleton. 2016. "An application of item response theory to psychological test development." Springer Open. April 18. https://prc.springeropen.com/articles/10.1186/s41155-016-0040-x.