Using video recordings as a COVID-19 adaptation to study the inter-rater reliability of the AccessTools app

Qussai M. Obiedat[1], Jaclyn Schwartz[2], Rochelle Mendonca[3], Suzanne Burns[4], Mason Drake[1], Laryn O'Donnell[1], and Roger O. Smith[1]

[1]Rehabilitation Research Design & Disability (R_2D_2) Center, University of Wisconsin-Milwaukee, [2]Florida International University, [3]Columbia University, [4]Texas Women's University.

ABSTRACT

Accessibility is an issue embracing all public buildings in society and is of vital importance to people with disabilities societal participation. Valid, reliable, and comprehensive accessibility assessment tools are scarce. AccessTools is a newly developed comprehensive and efficient accessibility assessment tool. This pilot study evaluates the inter-rater reliability (IRR) of the AccessTools assessment for trained student raters while using detailed video recordings of restaurants. Fleiss' kappa statistics calculations resulted in a slight to fair agreement between raters, while the Gwet's AC1 agreement coefficient calculations resulted in moderate agreement. Using video recordings for accessibility evaluation of buildings is not optimal, but it represents a helpful option under the current pandemic situation.

INTRODUCTION

Public buildings' accessibility is of societal importance for all individuals, especially people with disabilities (PWD). The creation of the Americans with Disabilities Act (ADA) in 1990 and its update in 2010 (ADA-ABA) [1], enabled thousands of PWD to gain access to formerly inaccessible public buildings. Despite all of the efforts that have been made on the societal and community levels, PWD are still limited from participating in the community due to inaccessible environments [2]. To empower PWD to engage and participate within their communities, it is essential to comprehensively assess the accessibility of public buildings and address any barriers that can hinder their participation.

The availability of valid, reliable, and comprehensive assessments for assessing and analyzing accessibility barriers of public buildings is still limited, and time consuming [3]. Additionally, although many accessibility assessment tools have been developed, the majority of the available tools focus on the physical aspect of accessibility and do not consider accessibility requirements for individuals that may have cognitive, visual, or auditory impairments.

AccessTools is a novel, comprehensive, and efficient accessibility assessment tool that has been recently developed to be used by trained assessors. Studying the inter-rater reliability (IRR) property is of high importance to ensure that the assessment is reliable when used by multiple raters. Typically, to study IRR of the assessment, multiple raters need to assess the same set of buildings in person and then the agreement of their ratings should be studied. However, as part of the toll of the COVID-19 pandemic, physical assessment of buildings is almost impossible. As an adaptation, video recordings of buildings that include detailed information regarding the accessibility related measurements can be used. The aim of this pilot study is to evaluate the inter-rater reliability (IRR) of the AccessTools assessment for trained student raters while using detailed video recordings of restaurants.

METHODS

AccessTools App

AccessTools is an iOS app that is designed to be used by trained assessors to assess, document, and quantify the accessibility details of building elements. The main level is comprised of 10 elements: Health Safety Measures; Parking & Valet Parking; Main Entrance/Exterior Doorway(s); Other Entrance(s)/Emergency Exit(s); Reception & General Information; Indoor Routes; Seating; Restroom(s); Other Interior Doorway(s); Specialty Features (Restaurants). Each element branches into several sub-elements until reaching the last level (leaf level) to provide more specific and detailed questions to help the assessors make informed decisions about the accessibility of the building. For example, the "Main Entrance/Exterior Doorway(s)" element branches into: Main Entrance Level Changes; Size of Doorway(s); Floor; Opening & Closing; Ease of Lock; Visibility; Automatic

Doors; Automatic Door Switch; Door Stops; and Signage providing the assessor with more details regarding the different sub-elements that need to be assessed.

In total, the assessment contains 2624 questions. In order to make the assessment more efficient, the AccessTools software utilizes a Trichotomous Tailored, Sub-branching Scoring (TTSS) system to elicit accessibility scores of each element and sub-element [4]. For each question, the software asks if a building feature is '[2] Fully Accessible / Yes', '[1] Somewhat Accessible', '[0] Not Accessible / No', '[X] Not Applicable') and uses skip logic to auto-advance through hundreds of questions about the building to optimize user efficiency. Figure 1 depicts the expanding outline and scoring screen.

Three embedded mini-apps were imbedded in the AccessTools app to reduce the time needed to perform the assessment. These apps will also be available as stand-alone apps, and include: AccessSlope, AccessSound, and AccessRuler. These mini-apps leverage the use of the iPad sensors to quickly measure inclines, decibels and distances in the context of accessibility, removing the need for separate tools, such as a tape measure, level, and clipboard. The app also enables assessors to take photos and videos of specific building elements in order to provide a complete picture of an identified accessibility barrier.
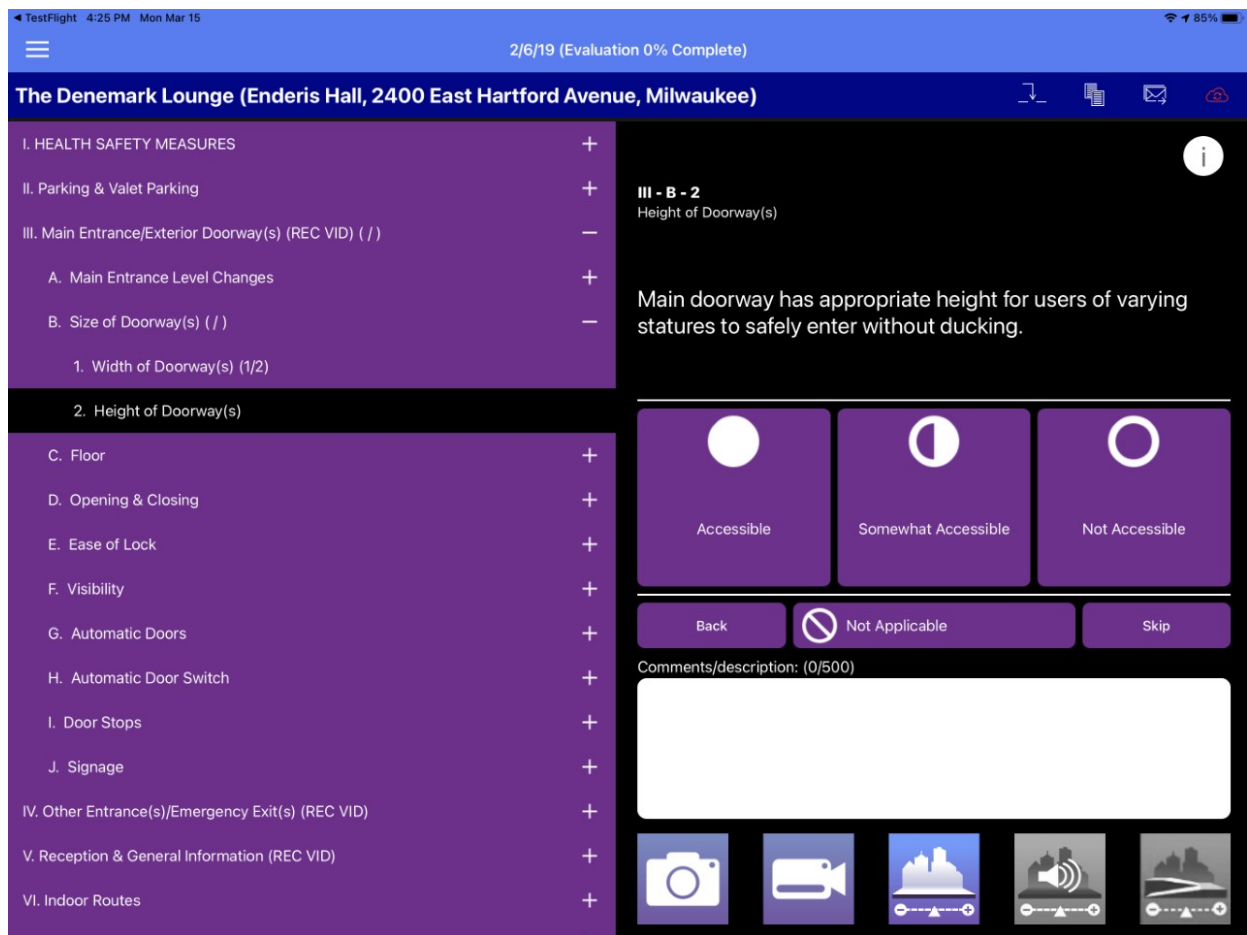

Figure 1: AccessTool app branching question system

Procedure

13 students participated in the data collection for this pilot study. All students received the AccessTools training before performing the restaurant assessments. The training included the completion of a self-directed learning module which was hosted on the learning management system Canvas and consisted of narrated presentations,

videos of people with disabilities experiencing environmental barriers, pictures, reflective questions, and links to key resources.

Short video recordings were created for three restaurants (A, B, C), while providing detailed information regarding the accessibility related measurements. For example, yard sticks were used to provide the dimensions of doors, tables, chairs, etc... Students were randomly assigned to assess two of the three restaurants using the AccessTools app. In each assessment, the following four elements were assessed: Main Entrance/Exterior Doorways, Indoor Routes, Seating, and Restrooms.

RESULTS

The raters' agreement calculations were performed based on the answers of the leaf level of the four elements that were assessed, which resulted in a total of 1553 questions (subjects). A total of 26 assessments were completed, 9 assessments for restaurant A, 9 assessments for restaurant B, and 7 assessments for restaurant C. One assessment for restaurant C was omitted from the analysis due to incompletion. The answers were considered nominal and the Fleiss' kappa statistics [5] and Gwet's AC1 agreement coefficient [6] were calculated for each restaurant in SAS 9.4 (SAS Institute, Cary NC) using the MAGREE procedure for computing estimates and tests of agreement among multiple raters [7]. The results for the Fleiss' kappa statistics are presented in Table 1, and Gwet's AC1 agreement coefficient are presented in Table 2.

Table 1: Kappa statistics for nominal response

Restaurant	Y	Kappa	Standard Error|H0	Z	Prob>|Z|	Standard Error	Lower Confidence Limit	Upper Confidence Limit
A	Overall	0.20343	.003098036	65.6640	<.0001	.005233094	0.19317	0.21369
B	Overall	0.15238	.003422251	44.5270	<.0001	.006907171	0.13884	0.16592
C	Overall	0.15300	.004344840	35.2153	<.0001	.007319714	0.13866	0.16735

Table 2: Gwet's Agreement Coefficient

Restaurant	Fixed	AC1	Standard Error	Z Value	Pr > |Z|	Lower Confidence Limit	Upper Confidence Limit
A	Raters	0.43026	0.007834	54.9241	<.0001	0.41490	0.44561
	Items	0.43026	0.066607	6.4596	<.0001	0.29971	0.56080
B	Raters	0.47486	0.004742	100.130	<.0001	0.46556	0.48415
	Items	0.47486	0.084772	5.602	<.0001	0.30871	0.64101
C	Raters	0.49545	0.006534	75.8296	<.0001	0.48264	0.50826
	Items	0.49545	0.095943	5.1640	<.0001	0.30741	0.68349

DISCUSSION

The calculations for the Fleiss' kappa statistics for the 3 restaurants resulted in a slight to fair agreement, while the Gwet's AC1 agreement coefficient calculations resulted in moderate agreement [8]. The lower results from the kappa calculations could be attributed to the 'κ paradox', which is considered as a disadvantage of kappa statistic [9]. The relatively low IRR agreement results were expected due to the limited ability of video recordings to capture all aspects of building elements that could potentially lower the agreement between raters. Some of these

issues can be resolved through providing better quality and more detailed recordings addressing the related accessibility measurements, while some issues will persevere due to the limitation of recordings to capture them as raters' perceptions are required to make informed decisions to assess them. For example, a rater will not be able to judge if the floor is slippery or not, or if it has glare just from watching the recordings.

Although the use of video recordings for accessibility evaluation of buildings is not optimal, it represents a helpful option under the current pandemic situation. Using recordings for accessibility assessments could also be of substantial importance for raters training, through providing initial exposure to the assessment tool and defining standardized minimum requirements for IRR agreement to be considered as a professional rater. Additionally, completing the AccessTools assessments using recordings resulted in a promising effect on the students' learning composite, which will be reported and discussed in future papers.

ACKNOWLEDGMENT

REFERENCES

[1] United States Department of Justice. 2010 ADA Standards for Accessible Design. Title II. 2010;(Journal Article):279–279.
[2] Hammel J, Magasi S, Heinemann A, Gray DB, Stark S, Kisala P, et al. Environmental barriers and supports to everyday participation: A qualitative insider perspective from people with disabilities. ArchPhysMedRehabil. 2015;96(4):578–88.
[3] Steinfeld E, Danford GS, Danford G. Enabling environments: Measuring the impact of environment on disability and rehabilitation. Springer Science & Business Media; 1999.
[4] Smith RO. OT FACT: A multi-level performance-based software instrument with an assistive technology outcomes assessment protocol. Technology and Disability. 2002;14(3):133–9.
[5] Fleiss JL, Levin B, Paik MC. Statistical methods for rates and proportions. john wiley & sons; 2013.
[6] Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. British Journal of Mathematical and Statistical Psychology. 2008;61(1):29–48.
[7] Compute estimates and tests of agreement among multiple raters [Internet]. [cited 2021 Mar 15]. Available from: https://support.sas.com/kb/25/006.html
[8] Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. 3rd, editor. Upper Saddle River, New Jersey: Prentice Hall; 2009.
[9] Walsh P, Thornton J, Asato J, Walker N, McCoy G, Baal J, et al. Approaches to describing inter-rater reliability of the overall clinical appearance of febrile infants and toddlers in the emergency department. PeerJ. 2014;2:e651.