Development of International Assessment Tools for Evaluating the Competency of Wheelchair Service Providers: Establishment of psychometric evidence of a basic knowledge test and a self-assessment questionnaire of provision skills

by

Anhue Yohali Burrola Méndez

Physical Therapy, American British Cowdray Hospital, 2010

Master of Science, University of Pittsburgh, 2014

Submitted to the Graduate Faculty of the

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Anhue Yohali Burrola Méndez

It was defended on

August 14, 2020

and approved by

Jonathan Pearlman, Chair and Associate Professor, Dept of Rehabilitation Science and Technology

Lauren Terhorst, Co-Director of the Physical Therapy Data Center, and Professor, Dept of Occupational Therapy

R. Lee Kirby, Professor, Division of Physical Medicine and Rehabilitation, Dept of Medicine, Dept of Community Health and Epidemiology, Dalhousie University

Dissertation Director: Mary Goldberg, Associate Professor, Dept of Rehabilitation Science and Technology

Development of International Assessment Tools for Evaluating the Competency of Wheelchair Service Providers: Establishment of psychometric evidence of a basic knowledge test and a self-assessment questionnaire of provision skills

Anhue Yohali Burrola Méndez, PT, MS

University of Pittsburgh, 2020

The limited training of wheelchair service providers has been associated with inappropriate wheelchair service provision. The World Health Organization (WHO) developed the Guidelines on the provision of manual wheelchairs and a series of Wheelchair Service Training Packages to support the training of personnel; the International Society of Wheelchair Professionals (ISWP) developed the ISWP Basic Test to assess wheelchair service knowledge. In terms of assessment of skills, there are currently no tools that evaluate all wheelchair service provision skills. The purpose of this dissertation was 1) the revision, evaluation, and update of the ISWP Basic Test to develop a new version; and 2) the development and establishment of psychometric evidence of a self-assessment questionnaire of basic wheelchair service skills.

Objective 1. We analyzed 943 successful first attempts of the ISWP Basic Test from 2015-2020. Passing rates and the questions' performance were obtained: item difficulty (*p*-values), and index of discrimination (*IDIs*). The questions that did not meet cutoffs were reviewed by an international group of stakeholders resulting in dropping 33 questions (22.7%), updating 112 (77.2%), and adding 61 new questions. The new set of questions was pilot tested by a group of 80 participants and the results indicated 61 (35.26%) met the *p*-values and *IDI* criteria, 62 (35.8%) met one of the criteria, and 50 (28.9%) did not meet both criteria. The pilot testing design and the sample analyzed led to expanding the cutoffs and retaining more questions.

Objective 2. A self-reported survey of 31 questions was designed and pilot tested with a group of Physical Therapy and Occupational Therapy students. The exploratory factorial analysis indicated that 5 factors explained 75.7% of the total variability of the scale. Items were grouped considering the factor load matrix; all of them had a Cronbach's alpha above 0.85. Results from the Spearman correlation indicated a strong, statistically significant positive association between times.

This work contributes to the development of international assessment tools to evaluate competency in wheelchair service providers, a sector's priority action to develop a competent workforce, and support good practice.

Table of Contents

List of Tables

List of Figures

Preface

This work would have not been possible without the support and guidance of many people who helped me complete this research and academic degree. First and foremost, I would like to thank my advisor, Dr. Mary Goldberg for her advice, guidance, time, and commitment thought this journey. My Dissertation Committee members, Dr. Lauren Terhorst, Dr. Jonathan Pearlman, and Dr. R. Lee Kirby for their support and constructive feedback. Thank you to ISWP staff Nancy Augustine, Krithika Kandavel, and Kim Robinson, your enthusiasm and willingness to help are inspiring. I am grateful to the talented group of colleagues, Paula Rushton, Padmaja Kankipati, María Luisa Toro, Norma Jiménez, Silvana Contepomi, Amira Tawashy, Tchai Xavier, Lee Kirby, Adriana Núñez, and Delia Ortega for their insights, contributions, and critical support for this work. Many thanks to all participants who volunteered their time for these research projects. Thank you to all my friends that became family in my years at Pittsburgh. Lastly, I want to thank the School of Health and Rehabilitation Sciences at the University of Pittsburgh, it has been an honor and privilege to be trained in a school that promotes health, diversity, and inclusion.

I would like to dedicate this work to my parents, Soraya Méndez and Iván Burrola, for their immense love and support and for being my greatest example of perseverance and hard work; to my beloved sister and best friend, Soraya Burrola; and to my husband Gibrán Álvarez who makes, with his love, everything possible.

Table 1 Nomenclature

Abbreviations and Acronyms	
ATC	Assessment Tools Committee
EFA	Exploratory Factor Analysis
IC	Internal Consistency
IDI	Index of Discrimination
IQR	Interquartile range
ISWP	International Society of Wheelchair Professionals
ISWP Basic Test	International Society of Wheelchair Professionals Wheelchair Service Provision Basic Test
ISWP WSP	International Society of Wheelchair Professionals Wheelchair Service Provider Certification
ITC	Item-total correlation
KMO	Kaiser-Meyer-Olkin
KR-20	Kuder-Richardson Formula 20
LMICs	Low and middle-income countries
MSH	Management Science for Health
OT	Occupational Therapy
PCA	Principal Component Analysis
PT	Physical Therapy
p-values	Items difficulty values
UN-CRPD	United Nations Convention on the Rights of People with Disabilities
USAID	United States Agency for International Development
WHO	World Health Organization
WHO Guidelines	World Health Organization Guidelines on the provision of manual wheelchairs in less-resourced settings
WHO WSTP-B	World Health Organization Wheelchair Service Training Packages - Basic level
WHO WSTPs	World Health Organization Wheelchair Service Training Packages
WMTP	Wheelchair Maintenance Training Program
WSP CAPS-B	Wheelchair service provision self-report capacity skills questionnaire – basic level
WST	Wheelchair Skills Test
WSTP	Wheelchair Skills Training Program
WST-Q	Wheelchair Skills Test - questionnaire version
Symbols	
α	Cronbach's alpha
r	Pearson’s correlation coefficient
W	Shapiro-Wilk test

1.0 Introduction

1.1 The Current State of Global Wheelchair Service Provision

The World Health Organization (WHO) and the World Bank estimate less than 20% of the 115 million people in the world who need a wheelchair for mobility and function have access to an appropriate wheelchair that meets their needs (1-3). The lack of access to an appropriate wheelchair has adverse effects on the life, safety, health, and other basic human rights of people with disabilities (2, 4-9). In addition, when a wheelchair does not meet the wheelchair user's needs it may result in underutilization or abandonment of the wheelchair (10, 11). This situation may be more problematic in low and middle-income countries (LMICs) where disability and poverty are interconnected, the incidence of disability is higher, people with disabilities often are marginalized, there is less availability of skilled health personnel, and there is a limited range of good quality and affordable wheelchairs (1, 12-16).

The United Nations Convention on the Rights of People with Disabilities (UN-CRPD), which promotes effective human rights for all persons with disabilities, indicates the importance of an appropriate wheelchair delivered by trained clinicians in six of its Articles (4, 9, 20, 25, 26, and 32) (17). In particular, Article 20 emphasizes the need to promote personal mobility with the greatest independence by providing training to persons with disabilities and personnel providing services to them (9). Nevertheless, the shortage of education and training in wheelchair service provision globally contributes to inappropriate wheelchair provision (14, 18, 19). This situation suggests that countries, where inappropriate wheelchair service delivery is occurring, are not fulfilling the promise of the UN-CRPD.

In 2018, the United States Agency for International Development (USAID), World Learning, Management Science for Health (MSH), the International Society of Wheelchair Professionals (ISWP), and the WHO planned the 'Wheelchairs Stakeholders Meeting' to reflect on past achievements and challenges, discuss current initiatives, and identify priority actions for the next five years to achieve the sector goal (20). The wheelchair sector goal was defined as: "By 2023, 10 countries have new or strengthened evidence-based, adequately-resourced, integrated wheelchair services supported by policies, competent personnel, and a range of appropriate wheelchairs" (20). In order to achieve the goal, the stakeholders' priority actions are: Build awareness, conduct research, establish global service standards, establish product standards, foster innovation, improve wheelchair supply, promote policy, stimulate collaboration, support competency development, and support good practice. The focus of this work is on competency development, an important priority action that will subsequentially support good practice, research, build awareness, and the establishment of global service standards.

1.2 Competency in Wheelchair Service Provision

Competency has been described as a multidimensional construct that involves a complex interaction of cognitive concepts related to the gathering of information and the processing of that information for translation into action. Thus, competency is a complex know-how that is based on combining knowledge, skills, and abilities with external resources and then applying them appropriately to specific types of situations based on the context (21, 22).

International organizations and research teams have developed wheelchair service

provision training materials, open-source, in multiple languages, that are available to download from their websites to help establish competency in wheelchair service provision and assist nations in fulfilling the UN-CRPD mandate of supporting providers' (2, 23). One such organization, the WHO, has published the Guidelines on the provision of manual wheelchairs in less-resourced settings (WHO Guidelines) and a series of open-source Wheelchair Service Training Packages (WHO WSTPs), in multiple languages, with the support of USAID (2, 23-26). These training materials have advocated an 8-step service delivery process: (1) referral and appointment; (2) assessment; (3) prescription; (4) funding and ordering; (5) product preparation; (6) fitting; (7) user training; and (8) follow-up, maintenance, and repairs (2). Professional rehabilitation programs have endorsed the use of the WHO 8-steps and research teams have begun to report evidence that supports this service delivery process (5, 19, 27-32).

The first training package of the WHO WSTPs series is the Basic level (WHO WSTP-B). This training package includes the knowledge and skills required to provide wheelchair service to people with mobility impairments who do not need additional postural support to sit upright (23). The training's target audience is personnel and volunteers, with or without a professional degree in the field, who are expected to carry out basic level wheelchair service delivery in their workplace (33). According to the WHO WSTPs, appropriate wheelchair service provision involves the development of skills and knowledge of the personnel involved in wheelchair service delivery. Therefore, the assessment of knowledge and skills using valid, reliable, and contextually appropriate assessment tools is a key component of understanding groups' specific needs, developing learning programs, testing the effectiveness of different learning approaches, and enhancing quality research.

The following sections will describe current assessment tools available to assess knowledge and skills in wheelchair service provision.

1.2.1 Wheelchair Service Provision Knowledge: ISWP Basic Test

1.2.1.1 Overview and dissemination

In 2015, ISWP developed a Wheelchair Service Provision Basic Test (ISWP Basic Test), an online test that measures basic level wheelchair service provision knowledge independent of geographic location (34). The test is aligned with the WHO Guidelines (2) and the WHO WSTP-B, and was developed as a first step to help assess wheelchair service knowledge worldwide. In 2018, after three years of development and free distribution of the ISWP Basic Test, the low pass rates confirmed the need to promote training of wheelchair service providers worldwide with 41% of test-takers in Africa passing the test; 44% in Asia; 46% in Latin America; 47% in Europe; 48 % in Australia and Oceania, and 55 % in North America (35). Recently, studies have used the ISWP Basic Test to evaluate baseline knowledge in rehabilitation professions, and the impact of the WHO WSTP-B training using different learning methodologies such as in-person and hybrid courses (36-39).

1.2.1.2 Development and limitations

The ISWP Basic Test consists of two sections: a demographic questionnaire and a multiple-choice test. The demographic questionnaire includes 19 questions regarding the participants' sociodemographic characteristics such as age, gender, education level, profession, employment status, years of experience in wheelchair service provision, work setting, and motivation to take the training. The questions related to the work setting, age group served, and the motivation to take

the training allow participants to select all applicable options. The multiple-choice test includes 75 questions from 7 domains of wheelchair service delivery knowledge: 1) assessment, 2) prescription, 3) fitting, 4) production, 5) user training, 6) process, and 7) follow-up and maintenance. A job task analysis was conducted with an international stakeholder group by The Institute for Performance Improvement to determine the weighting of each domain. The domains have different weights based on the pre-set number of questions that each domain was allocated; for instance, 19 questions are allocated to assessment while there are only 4 about follow-up and maintenance. Table 2 presents the pre-set number of questions allocated to each of the test's domains.

Table 2 ISWP Basic Test questions' distribution

Test Domain	Total number of questions
Assessment	19
Prescription	12
Fitting	10
Production	5
Users' training	15
Process	10
Follow up	4

Each domain has a pool of questions from which only a subset is drawn on each test. This reduces the likelihood of receiving the same question when taking the test multiple times. In addition, the test settings include: 1) random distribution of questions from each domain's pool of questions, 2) forced completion requiring participants to complete the test in one-time entry; and 3) immediate scoring of the test with the opportunity to review both correct and incorrect answers (34). Test scores greater than or equal to 53 points (70% of the total points) are considered passing scores.

The development of the ISWP Basic Test was conducted by a small international group of subject matter experts (34). The process included the selection of domains, the creation of questions, revision, and testing (36). The test has limited psychometric evidence and so far, no reliability analysis has been conducted. The methodology implemented for its development has several limitations such as the unknown characteristics and sample size of the international group of subject matter experts that guided the process; the unknown rationale for the selection of the domains and their weighting, the unknown methods used for developing, reviewing, and assessing the clarity and relevance of questions; the unknown procedure used to assess Internet bandwidth requirements; the exclusion of content from evidence-based training programs that are not included in the WHO WSTP-B; the unspecified sample size and characteristics of the alpha and beta testing pilot; and the lack of validity, reliability, and generalizability measures and evidence presented. To enhance the quality of research and to ensure the credibility of the findings in relation to global wheelchair service provision knowledge, the ISWP Basic Test needs to be revised, evaluated, and updated considering its items' performance and the inclusion of other relevant training packages.

1.2.2 Wheelchair Service Provision Skills Testing

The development of the ISWP Basic Test has been a first step to assess global knowledge level; however, the development of assessment methods for all wheelchair service provision skills is an area of research that has received little attention to date (14). Researchers and sector stakeholders have validated standardized assessment tools and training protocols for some wheelchair service provision steps, like wheelchair skills (aligned with the user training step) and maintenance and repairs (40, 41). For example, the Wheelchair Skills Test (WST) is a set of

standardized evaluation methods to document the capacity of a subject to perform wheelchair skills in a standardized setting (32 skills for manual wheelchairs, 28 for powered wheelchairs) (42). Each skill is scored 0-3 or 'NP' (No part) when the wheelchair does not have the parts to allow the safe performance of the skills. In addition to the scores for each skill, the total percentage score is calculated (42). The questionnaire version of the WST (WST-Q) is a set of questions related to the same set of manual and power wheelchair skills that compromise the WST. The WST-Q allows one to assess confidence, frequency, and performance of the wheelchair skills. (42). Multiple studies have evaluated the psychometric properties of the WST/WST-Q or have used them as outcome measures(43). The correlation between the total WST capacity and WST-Q performance scores is high although the WST-Q scores tend to be slightly higher. The WST and the WST-Q assessment tools are used interchangeably when the administration of one is preferable or more feasible over the other (44). Along with those efforts, a research team from the University of Pittsburgh and a collaborative team from Midwest Regional Spinal Cord Injury Care System, the Northern New Jersey Spinal Cord Injury System, and the South Florida Spinal Cord Injury Model System have developed the Wheelchair Maintenance Training Program (WMTP), a tool to train clinicians and caregivers (when applicable) to perform basic wheelchair maintenance (40, 45).

More recently, Ardianuari et. al (46) explored three basic skills assessment modalities (online mock-client case study quiz, in-person skills assessment, and video conference skills assessment) with a small sample of wheelchair service providers. The results suggest that the 3 modalities are potentially feasible but not comparable to the ISWP Basic Test. The mock-client case study quiz reported the fewest challenges encounter by the study team and participants thus it is the modality suggested for the ISWP Wheelchair Service Provider Certification (ISWP WSP) (46). Although this study provided insights about remote testing modalities, a bigger sample size,

and further pilot testing is required to identify barriers and possible facilitators that support the development of effective remote skills assessments tools.

The development of self-report assessment tools that do not require specific equipment, physical space, or additional human resources expands the possibilities and settings where they can be used. LMICs or resourced constrained places where funding is limited, there are fewer-available skilled personnel who can facilitate objective testing, and there is a limited range of quality and affordable wheelchairs, may particularly benefit from these types of assessment tools. In particular, the development of a self-report questionnaire to measure perceived capacity on all basic wheelchair service provision skills included in the WHO WSTP-B could be a practical solution to assess skills in wheelchair service provision in international settings and may contribute to assess competency in wheelchair service provision worldwide. This is the first step towards improving providers' awareness of their skills and a potential alternative for assessment where in-person skills tests are impractical.

The purpose of this dissertation is centered on 1) the revision, evaluation, and update of the ISWP Basic Test to develop a new version (2.0); and 2) the development, psychometric evaluation, and pilot test of a wheelchair service provision self-report capacity skills questionnaire – basic level (WSP CAPS-B) based on the WHO WSTP-B. These efforts are aligned with three priority actions of the 2018 Wheelchair Stakeholders Meeting, conduct research, support competency development, and support good practice to strengthen the wheelchair sector globally by 2023 (20).

2.0 Specific Aim 1. To review, evaluate, and update the International Society of Wheelchair Professionals Wheelchair Service Provision Basic Test, English version (ISWP Basic Test).

2.1 Methods

2.1.1 ISWP Basic Test Dataset analysis

2.1.1.1 Dataset

We analyzed the ISWP Basic Test data, English version, hosted in Test.com from January 1st, 2015 to January 24, 2020. The inclusion criteria included: 1) successful test attempts, test-takers who completed the demographic and the multiple-choice section, and 2) test-taker first attempt. We excluded all unsuccessful attempts and sub-sequent attempts of test-takers.

2.1.1.1.1 Statistical approach

A) Investigate passing rates and possible relations between test-takers' characteristics and total test scores. The sample (test scores) was analyzed for normality of distribution using the Shapiro-Wilk test (W) and visually inspected for normal distribution assumption using a histogram. Descriptive statistics, measures of central tendency, and measures of spread were calculated to summarize and describe the data and to determine the percentage of test-takers who passed the test. Bivariate linear regressions were performed to determine the possible relationships between a set of independent variables (e.g. age, educational level, years of experience in wheelchair service delivery, profession, country) and the dependent variable (total ISWP Basic Test scores). The variables with a significant association with the dependent variable, identified through simple bivariate linear regression, will be combined in a multiple linear regression model. Statistical significance was set at a 0.05 alpha level.

B) Investigate questions and domains performance. The Classical Test Theory (47, 48) was used as the framework to evaluate this project's data. The following item level statistics were obtained:

i. **Item difficulty (*p*)**. For dichotomously scored items, *p* is defined as the proportion of examinees who answer an item correctly (49, 50).

	$p_j = \frac{Number\ of\ people\ who\ correctly\ answer\ item_j}{N}$	**Equation 1**

This proportion is obtained by dividing the number of test-takers who respond correctly to an item (*j*) by the total number of respondents (*N*). *P* – values may range from .00 to 1.00. Lower *p*-values are indicative of more difficult items while higher

p-values suggest easier items (49). When all items are dichotomously scored, item variance is:

	$\sigma_j^2 = p_j q_j$	**Equation 2**

where $q_j = (1 - p_j)$. For this reason, the item difficulty level controls the item variance.

Assuming a total degree of correlation between items, total test score variance will be maximized when $p = 0.50$. In multiple-choice questions, the format used in the ISWP Basic Test, the observed *p*-value is affected by the examinees' true score and by guessing. Considering the random guessing assumption, the observed *p*-value is supposed to be the sum of the "true" *p*-value (examinees who know the answer) and 1/m the proportion of examinees who don't know the answer, when m is the number of choices (50) (Table 3).

Table 3 Observed difficulty (Po) with different number of answer options

Items' number of options	Proportion who know the answer	Proportion who guess the answer	Observed difficulty, Po	Lord's Po
4	0.5	0.50/4	0.50+(0.50/4) = 0.62	0.74
3	0.5	0.50/3	0.50+(0.50/3) = 0.67	0.77
2	0.5	0.50/2	0.50+(0.50/2) = 0.75	0.85

Table adapted from Crocker L, Algina J. Introduction to classical and modern test theory (50)

Frederick Lord (51) demonstrated that reliability is improved by choosing items with *p*-values higher than those obtained from random guessing assumption. Lord's values have been widely adopted and are considered more reasonable (50). Crocker

et al. (50) recommend using Lord's values because they recognize that many examinees have some partial knowledge that allows them to eliminate answer options before guessing, which make their possibilities even higher than 1/m (50). For the purpose of this study, we will used Lord's guidelines to analyze items' *p*-values. We will retain items with a *p*-value ≤0.74 for items with 4-answer options and ≤0.77 for items with 3-answer options.

ii. **Item discrimination.** The purpose of many tests is to differentiate between examinees based on the criterion of interest, for instance, between examinees who know the criterion and those who do not (49, 50). For dichotomously scored items, like the items in the ISWP Basic Test, the parameter is called **index of discrimination (IDI)**. The index is obtained by organizing the group's total test scores in descending order and dividing it in upper and lower scores. Kelly et al (52) identified that a sensitive and stable discrimination index can be obtained by using the upper and lower 27%. After the groups have been identified, the *IDI* is obtained:

	$IDI = p_u - p_l$	**Equation 3**

P_u is the proportion in the upper group who answer the item correctly and p_l the proportion in the lower group who answer correctly. Values of *IDI* may range between -1.00 to 1.00. Positive values indicate that the value discriminates in favor of the upper group, and negative values favors the lower-scoring group, indicating

it is a reverse discriminator (50). When the total test score (passing or failing) is the criterion, the literature (53) suggests using the guidelines presented in Table 4.

Table 4 Index of Discrimination Guidelines

IDI	**Item's interpretation**
≥ 0.40	Functioning satisfactorily
0.30 - 0.39	Little or no revision is required
0.20 - 0.29	Marginal, needs revision
≤ 0.19	Should be eliminated or completely revised

For the purpose of this study, we considered a cutoff of *IDI* ≥0.30, items with an index of discrimination smaller were flagged for revision.

iii. **Domains' reliability.** We used the Kuder-Richardson Formula 20 (KR-20) to estimate internal consistency reliability. KR-20 is commonly used to test internal consistency of an achievement test where items are scored dichotomously, and the test measures the unidimensional trait (54-56).

	$$r_{KR20} = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum pq}{\sigma^2}\right)$$	**Equation 4**

- r_{KR20} is the Kuder-Richardson formula 20
- k is the total number of test items
- Σ indicates to sum
- p is the proportion of the test-takers who pass an item
- q is the proportion of test-takers who fail an item
- σ^2 is the variation of the entire test

To meet the unidimensional condition, KR-20 coefficient was calculated and analyzed for each domain separately. KR-20 coefficients range from 0.0 to 1.0 (56). Values ≥0.7 will be considered 'good' internal consistency estimators.

2.1.2 Expert panel revision

2.1.2.1 Committee formation

ISWP emailed 10 invitation letters describing the scope of the project to a purposive sample of stakeholders, researchers, clinicians, and wheelchair service providers to form the Assessment Tools Committee (ATC). The inclusion criteria included 1) personnel with at least 5 years of experience in wheelchair service delivery, 2) familiar with the WHO WSTPs, WSP, and WMTP, 3) who are actively engaged in the wheelchair sector. Potential members who agreed to be part of the committee responded to a demographic survey that included questions exploring the familiarity with the content of the three training packages that will be used to review and develop the new version of the test, the WHO WSTP, the WSP, and the WMTP. In addition, the survey asked members to rate how confident they feel with interpreting components of a psychometrics analysis, correlations, and research methodology.

2.1.2.2 Revision process

The ATC received evidence-based training developed and delivered by this dissertation author about the best guidelines for the development of questions, the use of the assessment forms, and the methodology for the revision of the items. An online repository with relevant materials (e.g. articles, books, videos, presentations, etc.) was made available for all members.

The ATC was split into 7 teams of two members based on the total number of participants, familiarity with the training packages, and area of expertise. The member with the most years of experience and this dissertation author participated in all groups forming groups of 4 participants. The revision process consisted of four phases:

Phase 1: Analysis and revision. This phase included two steps. First, members dichotomously rated questions' text and answer options based on best practice guidelines on the development of questions. The author of this dissertation created the Question's Guidelines Table, a matrix based on evidence-based guidelines including 5 criteria for developing question text and 5 for developing answer options (Appendix A). After members completed the rating, they completed the Domain Feedback form which includes:

- First section: A domain's summary table consisted of the following statistics: total number of items in the domain, the total number of items flagged for revision, and the domains' reliability.
- Second and Third section: Flagged and unflagged questions with their *p*-values, *IDIs*, frequency, and percentage of answer options' selection.
- Four section: Table for new questions.

Members were asked to submit the Domains Feedback Form prior to Phase 2. The results from this phase were summarized and prepare for discussion in the next phase.

Phase 2: Synchronous Meetings. During the sessions, members discussed the results from Phase 1 and proposed changes. When new items were created, members rated them using the same methods described in Phase 1. Once consensus was reached, a pre-final draft of all items was created.

Phase 3: Pre-final version. The domain's pre-final draft was distributed to the rest of the ATC members for their revision. If members disagreed with any item, they emailed their concerns to the team who met again to review the questions. When consensus was reached, the pool of questions was finalized and considered ready for the classification phase. All meetings were recorded and made available for all ATC members.

Phase 4: Classification. Two independent reviewers (MG and AYBM), categorized all questions independently using the WHO WSTP-B' content as a reference. They met to discuss the discrepancies and do a consensus decision. Once consensus was reached, the primary author grouped content and proposed the recategorization of domains aligned to the WHO-WSTP-B.

Phase 5: Pilot preparation. ATC members made the last revision and provided feedback to all questions including their new categories. The final version was approved for pilot testing.

2.1.3 Pilot testing

2.1.3.1 Participants and recruitment strategy

A convenience sampling method was used to recruit wheelchair service providers interested in volunteering their time to evaluate the test. ISWP sent emails to its full contacts distribution list and social media posts to solicit volunteers. The messages sent included the link to the project description document. The document described the purpose of the project, the revision process, the instructions on how to complete the test, the average time it takes, the technical requirements, the acknowledgment for participants' time, a confidentiality statement, the research staff's contact information for questions or technical problems, and the link to access the test.

The inclusion criteria included: 1) wheelchair service providers who are native English speakers or proficient in English; 2) who have access to a smartphone, tablet or computer with Internet, 3) who can complete the revision, approximately 2 hours, between June 15th to June 29th.

2.1.3.2 Data collection

The new test was divided into two sections: demographic and knowledge questions. Participants were asked to respond to the demographic section before accessing the knowledge section. The evaluation process was the following:

1. Participants received a knowledge question and were asked to select the option that they considered correct. Afterward, they were prompted to indicate the perceived difficulty level from that question using a continuous grading scale from 0-10.
2. Once both questions were answered, the system automatically scored the questions and indicated the correct answer right after each question. Subsequently, an optional open-ended question was available for participants to provide feedback.

Participants could complete the questionnaire in multiple attempts but without the opportunity of reviewing or changing the questions to which they already responded. The survey was hosted in Qualtrics®, a survey software system that the University of Pittsburgh has a license to use, and could be completed from a tablet, computer, or smartphone connected to the Internet.

2.1.3.3 Statistical Approach

Descriptive statistics, measures of central tendency, and measures of spread were calculated to summarize and describe the characteristics of the sample analyzed. We implemented the same statistical approach to obtain the *p*-values and *IDIs* described in 2.1.1.1.1. Similarly, the threshold to flagged question remained consistent; questions with a *p*-value >0.74 and an IDI <0.30

were flagged for revision. The comments received were retained for individual revision as long as they provided feedback about the questions.

2.2 Results

2.2.1 ISWP Basic Test Dataset analysis

A total of 1276 test attempts of the ISWP Basic Test English version were completed between January 1, 2015 to January 24, 2020; 1108 were successful (completed) and represent 947 unique users on their first test attempt. Four users were removed from the analysis due to incomplete attempts that were not captured in the first exclusion. The sample size retained for analysis consisted of 943 test-takers. Table 5 includes the characteristics of the population, all total test-takers, and test-takers who passed and did not pass the test.

Table 5 Characteristics of test-takers and test scores.

Characteristics	**Total test-takers (n=943)**	**Pass (n=715)**	**Fail (n=228)**
Gender, n (%)			
Men	497 (52.7)	361 (50.49)	136 (59.65)
Women	446 (47.3)	354 (49.51)	92 (40.35)
Age, mean (SD)* {SEM}$^{\phi}$	34.46 (10.89) $^{\mu}$	34.58 {0.41}	34.06 {0.69}
English spoken at home, n (%)	315 (31.28) $^{\epsilon}$	263 (36.78)	52 (22.81)
Regions, n (%)$^{\alpha}$			
Asia	510 (54.25)	370 (51.97)	140 (61.4)
Africa	168 (17.88)	112 (15.73)	56 (24.56)

Table 5(continued)

North America	163 (17.34)	147 (20.65)	16 (7.02)
Oceania	38 (4.04)	35 (4.91)	3 (1.31)
Europe	35 (3.72)	29 (4.07)	6 (2.63)
Latin America	26 (2.76)	19 (2.67)	7 (3.07)
Educational level, n (%)			
Some College	147 (15.59)	95 (13.29)	52 (22.81)
High School	59 (6.26)	34 (4.76)	25 (10.96)
2-year degree/associates degree	111 (11.77)	74 (10.35)	37 (16.23)
4-year degree/bachelor's degree	380 (40.3)	309 (43.22)	71 (31.14)
Graduate degree/masters' level	213 (22.59)	173 (24.2)	40 (17.54)
Graduate degree/MD, PhD	33 (3.5)	30 (4.2)	3 (1.32)
Last year of formal training, n (%)			
Still attending	210 (22.27)	171 (23.92)	39 (17.11)
Less than 4 years	344 (36.48)	257 (35.94)	87 (38.16)
5-8 years	134 (14.21)	97 (13.57)	37 (16.23)
More than 8 years	255 (27.04)	190 (26.57)	65 (28.51)
Previous wheelchair training, n (%)			
No	305 (32.24)	211 (29.51)	94 (41.23)
Yes	638 (67.66)	504 (70.49)	134 (58.77)
Certification, n (%)			
Physical therapist	201 (39.33)	169 (40.82)	32 (32.99)
Occupational therapist	144 (28.18)	113 (27.29)	31 (31.96)
Prosthetics and Orthotics	76 (14.87)	59 (14.25)	17 (17.53)
WHO WSTP-B	64 (12.52)	48 (11.59)	16 (16.49)
WHO WSTP-I	4 (0.78)	3 (0.72)	1 (1.03)
ATP	22 (4.31)	22 (5.31)	0 (0)
Employment status, n (%)			
Unemployed	202 (21.42)	150 (20.98)	52 (22.81)

Table 5(continued)

Part-time 20hrs/week or less	125 (13.26)	100 (13.99)	25 (10.96)
Full time	616 (65.32)	465 (65.03)	151 (66.23)
Work setting: yes, n (%)			
Academic	344 (26.81)	270 (27.36)	74 (25)
In-patient	457 (35.62)	343 (34.75)	114 (38.51)
Outpatient	434 (33.83)	343 (34.75)	91 (30.74)
Other	48 (3.74)	31 (3.14)	17 (5.74)
Age group served: yes, n (%)			
Early childhood	475 (22.51)	378 (22.49)	97 (22.61)
Adolescents	478 (22.65)	389 (23.14)	89 (20.75)
Adults	714 (33.84)	551 (32.78)	163 (38)
Older adults	443 (21)	363 (21.59)	80 (18.65)
Motivation for training: yes, n (%)			
Professional growth	762 (61.55)	604 (62.92)	158 (56.83)
Personal growth	253 (20.44)	204 (16.48)	49 (3.96)
Required by academic program	137 (11.07)	96 (7.75)	41 (3.31)
Required by employer	86 (6.95)	56 (4.52)	30 (2.42)
Wheelchair service provision experience, years, n (%)			
<3 years	571 (60.55)	420 (58.74)	151 (66.23)
4-7 years	169 (17.92)	133 (18.6)	36 (15.79)
8 years or more	203 (21.53)	162 (22.66)	41 (17.98)
Wheelchair service provision, hours, n (%)			
Less than 3 hours/week	332 (35.21)	265 (37.06)	67 (29.39)
3 -20 hours/week	431 (45.71)	309 (43.22)	122 (53.51)
More than 20 hours/week	180 (19.09)	141 (19.72)	39 (17.11)
Member of an organization, n (%)			
No	275 (29.16)	196 (27.41)	79 (34.65)
Yes	668 (70.84)	519 (72.59)	149 (65.35)

Table 5(continued)

Mean Scale scores, mean (SD) {SEM}			
Total Wheelchair Service Basic Test (75)	56.48 (9.35)	60.58 {0.17}	43.6 {0.58}
Assessment	15.63 (2.5)	16.52 {0.1}	12.84 {0.2}
Prescription	8.93 (1.9)	9.54 {0.05}	7 {0.13}
Fitting	5.6 (1.86)	6.15 {0.06}	3.91 {0.11}
Production	3.6 (1.21)	3.9 {0.04}	2.7 {0.09}
User's training	11.25 (2.54)	12.12 {0.06}	8.53 {0.19}
Process	8.4 (1.85)	9.03 {0.04}	6.5 {0.16}
Follow up and maintenance	3.02 (1.01)	3.3 {0.03}	2.16 {0.08}

*SD: Standard deviation

ϕSEM: Standard error of the mean

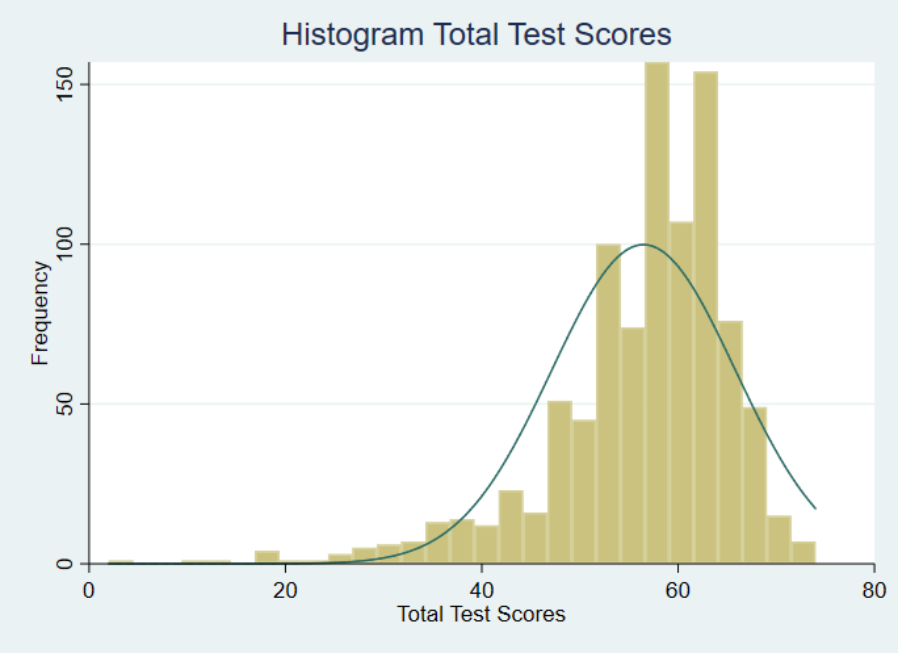
μN=938, 5 participants did not respond

ϵ Participants may enter multiple languages.

α The countries that composed the regions were the following. Asia: Bhutan, Cambodia, China, Hong Kong, India, Indonesia, Iraq, Israel, Japan, Jordan, Lebanon, Malaysia, Myanmar, Nepal, Pakistan, Palestine, Philippines, Saudia Arabia, Sri Lanka, Tajikistan, Thailand, United Arab Emirates, and Yemen. Africa: Ethiopia, Ghana, Kenya, Malawi, Nigeria, Sierra Leone, South Africa, Sudan, Tanzania, Uganda, and Zimbabwe. North America: Canada and the United States of America. Oceania: Australia, New Zealand, and Papua New Guinea. Europe: Belgium, Czech Republic, Finland, Hungary, Italy, Norway, Romania, Spain, Switzerland, and the United Kingdom. Latin America: Argentina, Brazil, Colombia, El Salvador, Guatemala, Haiti, Mexico, Nicaragua, and Peru.

The sample of total test scores was not normally distributed ($W=0.89$, $p<0.0001$). The histogram showed that the data is skewed to the left (Figure 1).

Figure 1 Histogram of Total Test Scores



Bivariate linear regression between independent variables and the dependent variables were not statistically significant. Considering these findings, multivariate linear regressions were not explored.

2.2.1.1 **Investigate questions and domains performance.**

A total of 145 questions were reviewed, 29 (20%) met the *p*-values and *IDI* criteria, 26 (17.9%) did not meet the *IDI* criteria, 33 (22.8%) did not meet the *p*-value criteria, and 57 (39.3%) did not meet both criteria. Appendix B presents the tables with the domains' summary and the results from the item analysis per domain.

2.2.2 Expert panel revision and pilot testing

2.2.2.1 Formation of the expert panel group

A total of eight people who met the inclusion criteria accepted our invitation to join the ATC. Also, the primary author of this dissertation and her advisor joined the ATC to form a group of ten members. Table 6 includes committee members' characteristics. Two members, one from Mexico and another one from the Philippines dropped out of the group due to an increase of workload that prohibited them to continue contributing to this project.

Table 6 Assessment Tool Committee Demographics

Characteristics	**Assessment Tool Committee (n=10)**
Gender, Female, n (%)	9 (90)
Age, mean (SD)*	44.5 (12.45)
Country of origin, n (%)	
Argentina	1 (10)
Canada	3 (30)
Colombia	1 (10)
India	1 (10)
Mexico	2 (20)
Philippines	1 (10)
United States of America	1 (10)
Primary language, n, (%)[µ]	
English	5 (45.45)
Spanish	4 (36.36)
Tagalog	1 (9.09)
Tegulu	1 (9.09)
Secondary language, n (%)	
English	4 (40)
French	2 (20)
Hindi	1 (10)
Italian	1 (10)
Kannada	1 (10)
Spanish	1 (10)

Table 6 (continued)

Educational level, n (%)	
4-year degree/bachelor's degree	1 (10)
Graduate degree/masters' level	4 (10)
Graduate degree/MD, PhD	5 (50)
Profession, n (%)	
Biomedical Engineer	2 (20)
Educator	1 (10)
Occupational Therapist	3 (30)
Physical Therapist	3 (30)
Physician	1 (10)
Primary Occupation, n (%)	
Consultant	4 (40)
Professor/Researcher	4 (40)
Service Provider	1 (10)
Student	1 (10)
Years involved in wheelchair sector, n (SD)	16.1 (10.74)
Wheelchair service providers, n (%)	7 (70)
Years providing wheelchair services	13.86 (13.87)
Have taken the ISWP Basic Test: yes, n (%)	7 (70)
Have taken the ISWP Basic Test: yes, n (%)	3 (30)

* SD: Standard deviation

μ : N= 11, one participant entered 2 primary languages

2.2.2.2 Revision process

The total pool of questions consisted of 145 items, 116 (80%) flagged, and 29 (20%) unflagged, and was reviewed by the ATC. Table 7 presents the average of agreement per domain based on the best practice guidelines for the development of questions (Appendix A). The second row of Table 7 indicates the total number of reviewers and the row below the total number of questions reviewed. The numbers bolded indicate the three domains with the lowest average percentages of agreement per criteria. The stem criteria with the most frequent lowest average percentages of agreement were *'stem is meaningful by itself and presents a definite problem'* and *'stem does not*

contain irrelevant material' while in the answer options' criteria were *'alternatives are free from clues'* and *'alternatives are mutually exclusive'*.

Table 7 Questions' criteria average of agreement per domain

Criteria	Assessment	Prescription	Fitting	Production	User's Training	Process	Follow-up & Maintenance
Reviewers, n*	3	4	4	3	4	3	3
Sample, n[μ]	34	17	22	12	26	25	9
Stem							
Stem is meaningful by itself and presents a definite problem.	87%	90%	**68%**	**67%**	83%	**77%**	89%
Stem does not contain irrelevant material.	90%	**78%**	77%	81%	**81%**	88%	**70%**
Stem is a question or partial sentence.	95%	91%	98%	94%	98%	98%	100%
Stem is negatively stated (*italics*/CAPS) only when significant learning outcomes require it.	**83%**	N/A[¥]	N/A	N/A	N/A	N/A	N/A
Stem targets a specific cognitive process.	100%	94%	90%	86%	96%	90%	93%
Answers' options							
All alternatives are plausible.	71%	79%	85%	**61%**	86%	81%	93%
Alternatives are stated clearly and concisely.	81%	81%	75%	67%	**77%**	79%	**81%**
Alternatives are mutually exclusive.	72%	**72%**	**67%**	75%	78%	**71%**	89%
Alternatives are homogenous in content.	85%	94%	90%	83%	92%	96%	100%
Alternatives are free from clues	**69%**	**72%**	82%	81%	93%	77%	**81%**

* Rev: total number of reviewers.

[μ]: N: total sample of questions per domain.

[¥] N/A: Non applicable.

A total of ten two-hour meetings were held between members to discuss and share the results from the domains' analysis. The results from this multiple revision process resulted in (1) dropping 33 questions (22.7%); (2) updating 112 (77.2%); and (3) adding 61 new questions. A total of 173 multiple-choice questions were retained for pilot testing. Table 8 includes the new classification based on the construct of interest, subconstruct, components, and number of questions per section. Appendix C includes all pilot tested questions including their classification.

Table 8 Pilot test version constructs, subconstructs, and components.

Construct	Subconstruct, n (%)	Components, n (%)		
		1st level	2nd level	3rd level
Basic Wheelchair Service Provision	Core Knowledge, 29 (16.8)	Wheelchair users, 9 (5.2)	Sitting upright, 3 (1.7)	
			Pressure injury, 6 (3.5)	Causes & risk factors, 3 (1.7)
				Risk areas, 2 (1.2)
				Stages, 1 (0.6)
		Wheelchair services 20 (11.6)	Appropriate wheelchair, 5 (2.9)	
			Cushions, 7 (4.1)	
			Documentation, 4 (2.3)	
			General information, 4 (2.3)	
	Wheelchair Service Steps, 144 (83.2)	Assessment, 35 (20.2)	General information, 2 (1.2)	
			Interview, 13 (7.4)	
			Physical assessment, 20 (11.6)	Body and wheelchair measurements, 6 (3.5)
				General info, 4 (2.3)
				Wheelchair Skills, 10 (5.8)
		Prescription and product preparation 26 (15)	Adjustments, 1 (0.6)	
			Cushion fabrication, 2 (1.2)	

Table 8 (continued)

			General information, 4 (2.3)	
			Wheelchair features and user needs, 9 (5.2)	
			Wheelchair measurements and features, 8 (4.6)	
			Wheelchair safe and ready, 2 (1.2)	
		Fitting, 28 (16.2)	General information, 2 (1.2)	
			Posture, 6 (3.5)	
			Pressure, 11 (6.4)	
			Problem-solving, 2 (1.2)	
			Size and adjustments, 4 (2.3)	
			While moving, 3 (1.7)	
		User training, 38 (22)	Pressure injury prevention, 2 (1.2)	
			Transfers, 4 (2.3)	
			Wheelchair skills, 31 (17.9)	
			Wheelchair transportation, 1 (0.6)	
		Follow-up and repairs, 17 (9.8)	Follow-up actions, 11 (6.4)	
			Repairs, 1 (0.6)	
			Wheelchair maintenance, 5 (2.9)	

*N= the total number of questions included per construct (total test) subconstruct, and component; (%) proportion of each question relative to the total pool of the test.

2.2.3 Pilot testing

2.2.3.1 Characteristics of the Sample Analyzed

A total of 163 test responses were received, 83 incomplete (78 tests with <50% of completion), and 80 complete test attempts that were retained for analysis. The demographic profiles from volunteers reported that 44 (55%) were male and 36 (45%) female with a mean age of 42.3 years old (SD=13.87). Less than half of or sample (45%) indicated English as their first language but all volunteers identified themselves as proficient in the language. Volunteers' country of birth was geographically distributed in 5 regions, with the majority of participants from Asia and North America (N=26, 32.5%, each) followed by Africa (N=14, 17.5%). The sample's most common professions were Physical Therapy (PT) (N=30, 37.5%), and Occupational Therapy (OT) (N=25, 31.25%); and the number of years involved in wheelchair sector showed a distribution of novice and veteran subjects with 50 (63.29%) with less than 10 years of experience and 29 (36.7%) with more than 10 years. Table 9 presents the characteristics of the sample analyzed. The demographic section included a question related to self-perceived confidence in knowledge in wheelchair training packages. In response to the questions pertaining to the training packages considered for the development of the new test, participants reported high knowledge confidence levels (fairly confident and completely confident) with 61 (76.25%) test-takers selecting those ranks for the WHO Wheelchair Service Training Package Basic Level, 40 (50%) for the Wheelchair Skills Program, and 33 (41.25%) for the Wheelchair Maintenance Training Program (Table 10).

Table 9 Characteristics of pilot test sample

Characteristics	**Total test-takers (n=80)**
Gender, n (%)	
Men	44 (55)
Women	36 (45)
Age, mean (SD)*	42.3 (13.87)
English spoken at home, n (%)	36 (45)
Country of birth, n (%)[α]	
Asia	26 (32.5)
Africa	14 (17.5)
North America	26 (32.5)
Europe	5 (6.25)
Latin America and the Caribbean	9 (11.25)
Profession, n (%)	
Physical therapist[ϕ]	30 (37.5)
Occupational therapist[ϕ]	25 (31.25)
Prosthetics and Orthotics	5 (6.25)
Orthotic Technician	4 (5)
Orthopedic Technologist	2 (2.5)
Wheelchair Technician	2 (2.5)
Other	12 (15)
Work setting, n (%)	
Academic	6 (7.5)
Clinician (in-patient)	14 (17.5)
Clinician (out-patient)	23 (28.75)
Industry	2 (2.5)
Self-employed	6 (7.5)

Table 9 (continued)

Other	29 (36.25)
Community Based Rehabilitation	4 (13.8)
Educational level, n (%)	
High School graduate	2 (2.5)
Some College but not degree	8 (10)
Associates degree/ 2-year degree	5 (6.25)
Bachelor's degree/4-year degree	26 (32.5)
Masters' degree	32 (40)
Professional degree/ JD, MD	4 (5)
Doctoral degree	3 (3.75)
Last year of formal training, n (%)	
Less than 5 years	25 (31.25)
5 – 10 years	13 (16.25)
11 – 20 years	25 (31.25)
More than 20 years	17 (21.25)
Years involved in wheelchair sector, n (%)[µ]	
Less than 5 years	29 (36.7)
5 – 10 years	21 (26.58)
11 – 20 years	16 (20.25)
More than 20 years	13 (16.45)
Wheelchair service provision experience, years, n (%)[¥]	
Less than 5 years	29 (40.84)
5 – 10 years	16 (22.53)
11 – 20 years	17 (23.94)
More than 20 years	9 (12.67)
Have taken ISWP Basic Test, n (%)	
No	28 (35)

Table 9 (continued)

Yes	52 (65)
Pass	46 (88.46)
Have taken ISWP Intermediate Test, n (%)	
No	70 (87.5)
Yes	10 (12.5)
Pass	8 (80)

*SD: Standard deviation

α Countries included. Latin America and the Caribbean: Mexico, Colombia, Argentina, and Haiti. North America: Canada and the United States of America. Africa: Kenya, Uganda, United Republic of Tanzania, Nigeria, Cameroon, Zimbabwe, Somalia, Sierra Leone, South Africa, Ethiopia, and Malawi. Asia: India, Nepal, Pakistan, Philippines, Bhutan, Hong Kong (S.A.R.), and Bangladesh. Europe: United Kingdom of Great Britain and Northern Ireland and Italy.

ϕIncludes students

µN=79, 1 participant did not respond

¥ N=71

Table 10 Self-reported confidence in knowledge of training packages

Wheelchair Training Packages	**Not confident**	**Slightly confident**	**Somewhat confident**	**Fairly confident**	**Completely confident**
World Health Organization, n (%)					
Wheelchair Service Training Package Basic Level	7 (8.75)	4 (5)	8 (10)	28 (35)	33 (41.25)
Wheelchair Service Training Package Intermediate Level	20 (25)	12 (15)	9 (11.25)	23 (28.75)	16 (20)
Wheelchair Service Training Package Managers	26 (32.5)	13 (16.25)	13 (16.25)	18 (22.5)	10 (12.5)
Wheelchair Service Training Package Stakeholders	22 (27.5)	22 (27.5)	13 (16.25)	15 (18.75)	8 (10)
Wheelchair Service Training of Trainers Package	22 (27.5)	17 (21.25)	11 (13.75)	15 (18.75)	15 (18.75)

Table 10 (continued)

Wheelchair Skills Program, n (%)	12 (15)	9 (11.25)	19 (23.75)	22 (27.5)	18 (22.5)
Wheelchair Maintenance Training Program, n (%)	19 (23.75)	13 (16.25)	15 (18.75)	20 (25)	13 (16.25)

2.2.3.2 Items performance

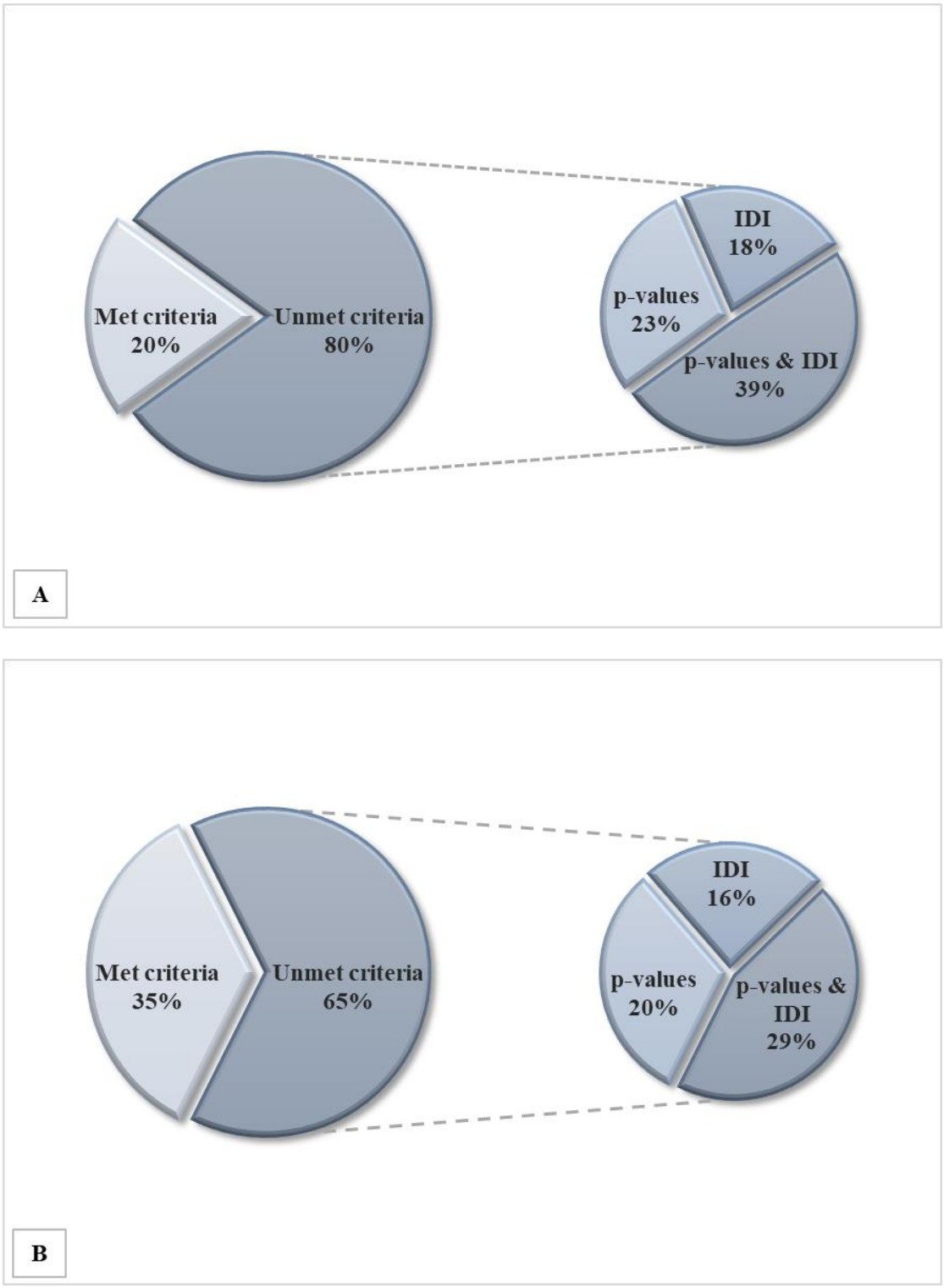
A total of 173 multiple-choice questions were pilot tested, 61 (35.26%) met the *p*-values and *IDI* criteria, 27 (15.6%) did not meet the *IDI* criteria, 35 (20.2%) did not meet the *p*-value criteria, and 50 (28.9%) did not meet both criteria. Table 11 and Figure 2 compares the results from the item level analysis between the ISWP Basic Test and the pilot set of question. Appendix D includes the results of all pilot tested questions.

Table 11 ISWP Basic Test and Pilot set of questions item level analysis results

	ISWP Basic Test Sample, n (%) N= 943*	Pilot tested questions, n (%) N=80
Meet criteria		
p-values & IDI	29 (20)	61 (35.26)
Unmeet criteria		
p-values	33 (22.76)	35 (20.2)
IDI	26 (15.93)	27 (15.6)
p-values & IDI	57 (39.31)	50 (28.9)
Total questions	145	173

*N= total sample of test-takers analyzed

Figure 2 Item level analysis results: ISWP Basi c Test and Pilot set of questions



A: ISWP Basic Test, total sample of 943 test-takers.

B: Pilot test of questions, total sample of 80 test-takers.

In view of providing a list of suggestions for future work, the results from the item performance were classified in 3 sets of questions: Set A, includes the questions that met the *p*-values and *IDI*s criteria; Set B incorporates questions that did not meet the *p*-values or the *IDIs*; and Set C is integrated by questions that did not meet the *p*-values nor the *IDIs*. Considering this grouping, Table 12 presents the question sets' frequency based on the subdomain and 1st level component classification. Set C had the highest flagged number of questions in the subdomain 'Core knowledge' section when compared to the other set of questions. In contrast, the same set reported the fewest frequency in flagged questions from wheelchair service steps when contrasted with the other sets.

Table 12 Question sets' frequency based on subdomain classification

Classification			**Set A**	**Set B**	**Set C**	
			Met p-values **and** IDIs	Did not meet p-values **or** IDIs	Did not meet p-values **nor** IDIs	**Total**
Core Knowledge	Wheelchair users	Sitting upright		1	2	3
		Pressure injury	3		3	6
	Wheelchair services	Appropriate wheelchair	1	3	1	5
		Cushions	2		5	7
		Documentation	2	1	1	4
		General information		2	2	4
Subtotal			**8**	**7**	**14**	**29**
Wheelchair Service Steps	Assessment	General information		1	1	2
		Interview		2	11	13
		Physical assessment	9	7	4	20
	Prescription and product preparation	Adjustments	1			1
		Cushion fabrication	2			2

Table 12 (continued)

		General information	1	1	2	4
		Wheelchair features and user needs	2	5	2	9
		Wheelchair measurements and features	3	5		8
		Wheelchair safe and ready		1	1	2
	Fitting	General information			2	2
		Posture	4	2		6
		Pressure	7	3	1	11
		Problem-solving		2		2
		Size and adjustments	1	1	2	4
		While moving	1	1	1	3
	User training	Pressure injury prevention			2	2
		Transfers	3	1		4
		Wheelchair skills	15	16		31
		Wheelchair transportation	1			1
	Follow-up and repairs	Follow-up actions	2	5	4	11
		Repairs			1	1
		Wheelchair maintenance	1	2	2	5
Subtotal			**53**	**55**	**36**	**144**
Total			61	62	50	173

A total of 2,568 comments were received. We retained 1463 (56.97%) that contained feedback about the questions for further analyses. Appendix E presents all comments retained for revision.

2.3 Discussion

2.3.1 Summary of results

This project reviewed and evaluated the ISWP Basic Test and created a new set of questions using a systematic approach guided by an international group of subject matter experts. The results from the revision show that only 20% of the questions included in the ISWP Basic Test met both the difficulty values and the index of discrimination cutoffs considered in the literature as standard principles for effective test development. The pilot results from the new set of questions developed revealed that 28.9% did not meet the difficulty values nor the discrimination cutoffs, 35.8% met one of the two item level statistics, and 35.2% met both the difficulty values and the discrimination cutoffs. The pilot testing design and the characteristics of the sample analyzed are important considerations when analyzing the results. A higher percentage of questions may need to be retained to ensure the new ISWP Basic Test considers a cross-cultural adaptation for use in international settings.

2.3.2 ISWP Basic Test analysis and results

2.3.2.1 Item level analysis

We analyzed a total of 943 completed first attempts of English tests received in a time frame of 5 years, the results indicated that more than two-thirds of the test-takers (75.8%) passed the test. These findings differ from the ongoing sector's call to advocate for more training due to limited competency in wheelchair service provision worldwide. In 2018, ISWP published on its

website the passing rates of the ISWP Basic Test showing that 6 out of the 7 regions included had passing rates below 50%, with only North America reporting a 55% passing rate (35). The report lacks the methodological description and details to contrast it with this project findings but reports overall significantly lower passing rates. In 2017, Fung KH et. al conducted a cross-sectional study to investigate the current state of wheelchair service education in academia (18). The results suggested limited training time allocated to wheelchair service provision in professional rehabilitation programs in low- to high-income countries (18). Considering this situation, the Wheelchair Stakeholders' Meeting held in Bangalore India in 2018 included strengthen personnel competency and the development of evidence-based practices as elements of the sector's goal by 2023 (20).

The high passing rates and test scores of the ISWP Basic test may not be reflecting competency in basic wheelchair service provision, instead, they could be explained by the unmet basic principles for effective questions design found in the test. The results from the item level analysis reported that 80% of the total test's pool (116 questions) did not meet the item difficulty values and index of discrimination suggested in the literature (47, 50). In particular, *p*-values were extremely high with 89 questions (61.4%) reporting scores >0.74, the maximum *p*-value suggested by the literature for multiple-choice questions with 4 answer options. This situation indicates that more than two-thirds of the questions could be responded to intuitively without testing any particular knowledge of wheelchair service provision. Similarly, the indices of discrimination outlined 34 (23.45%) questions ≤ 0.19 interpreter by the literature as '*the question should be eliminated or completely revised'* and 48 (33.1%) questions with scores between 0.20 – 0.29, indicating a '*marginal score, the question needs revision'*(50). These results suggest that the questions below the threshold do not provide information about individual differences on the

construct purportedly measured (50), in other words, the question does not discriminate between subjects who know the material and those who do not (50). Although these item statistics are sample dependent, the Classical Test Theory establishes a minimum sample size of 200 to mitigate this limitation (48). We had an average items' sample size of 487.8 (SD= 107.7) per question that exceeds the minimum size for item parameter estimation. Moreover, these findings are consistent with the first phase of the expert panel revision in which questions' quality was reviewed using a template that integrated 10 evidence-based guidelines (Appendix A). Reviewers indicated that the two criteria with the lowest average concurrence of opinion were '*stem is meaningful by itself and presents a definite problem'* and '*stem does not contain irrelevant material*'; in terms of the answer options, the two criteria most conflictive were '*alternatives are mutually exclusive*' and '*alternatives are free from clues*'. This exercise helped members throughout the revision process by focusing on fixing the aspects that were reported as problematic.

2.3.2.2 Revision process

We had a highly motivated group of reviewers that suggested evaluating the entire pool of questions by including in the revision the 29 questions (20%) that met the item statistics cutoffs. During the exercise, members identified the following main issues: 1) Many questions were misplaced in the domains impacting the domains' pool of questions. For instance, the question *'(code: 501490) What fitting tasks are best done when the user is sitting in the wheelchair?'* was classified as a 'Production domain' question instead of 'Fitting domain'; or the question *'(code: 501523) What activities would you not teach during standard wheelchair maintenance training?'* was included in 'User training domain' instead of 'Maintenance and repairs'. Additionally, some questions could not be placed under any current domain leaving them unclassified; for example, *'(code: 501333) What is the first sign of a pressure sore that appears 30 minutes after pressure*

has been removed from an area?' this question was included in the 'Assessment domain' although refers to the Core Knowledge section of the WHO WSTP-B and not to the Wheelchair Service Steps. 2) A total of 9 questions were repeated which overestimated the domains' pool of questions. 3) Questions did not use person-first language; for example, *'(code:501317) Your client is a 24--year--old paraplegic who has returned to your service with a pressure sore. What is NOT likely to be the cause of the pressure sore?'*. 4) There was a lack of consistency in the use of terms throughout the total pool of questions. These issues increased the members' workload and exceed the time commitment described in the invitation letter. The two members who dropped out of the group referred to not being able to complete all activities due to the amount of work. This situation not only impacted the members' group size but also the scope of the work leaving less time to develop new items and to determine the weighting of the domains.

A total of 61 new items were created, being the majority, 22, related to the User's training domain. The inclination towards the aforementioned domain could be explained by the members' profile and area of expertise. Four members of the ATC were part of the Editorial Committee of the Wheelchair Skills Program including the primary author of the training package. The new items developed used evidence-based content included in the WSP that was not covered in the WHO WSTP-B. The other training package considered for updating the test was the WMTP. To our surprise, one of the primary developers of this training package, also a member of the ATC, considered it inappropriate to include questions from that training resource. The rationale was that the WMTP includes recommendations for the frequency to complete maintenance tasks that were selected based on a group of content matter experts from the United States of America thus may not be generalizable to other contexts. Not all members agreed with that approach but since the testing phase was about to start consensus was not reached. Future ATC meetings should discuss

the inclusion or exclusion of the WMTP until a consensus decision is reached. As explained in the next section, the determination of domains' weighting was not completed due to time constraints. Future testing revisions could consider economic retributions to compensate members' time and a shorter revision period between versions. Since the development and distribution of the ISWP Basic Test in 2015, this has been the first formal revision conducted to the test. Periodic revisions of the test should be integrated into the ongoing organization practices to assist in maintaining, evaluating, and updating the assessment tool.

2.3.2.3 Domains' classification

The last two phases of the expert panel revision resulted in a new classification of subdomains and components of the test (Table 12). The new categorization is more aligned with the content of the WHO WSTP-B, the primary training package used for the development of the test, than the previous ISWP Basic Test domains' classification (34). In this new version, we included the subdomain 'Core Knowledge' along with its first level components, 'wheelchair users' and 'wheelchair services'. The subdomain of 'wheelchair service steps' is integrated by 5 components that represent 6 wheelchair service steps, the steps of 'prescription' and 'product preparation' were combined in one component. The WHO WSTP-B steps referral and appointment and funding and ordering lack of sufficient content to create a pool of questions of a standing component. This new classification recognizes that the construct domain of basic wheelchair service provision is integrated by two subdomains: Core Knowledge, that includes the background knowledge to undertake an appropriate wheelchair service delivery; and Wheelchair Service Steps, the interconnected actions to provide an appropriate wheelchair service delivery (33). The inclusion of the 'Core Knowledge' section is particularly important when considering that the target audience of the WHO WSTP-B includes personnel without a health care professional degree

such as community-based workers, technicians, and local craftsmen (33). It is important to consider that despite the ISWP Basic Test and the new set of questions were developed using the same training package, the test versions are not parallel in form. The new set of questions significantly changed and there was not a single question that remained the same between versions. The test domains should not be compared across versions unless test equating is conducted (57). Future work should include the discussion of the domains' weighting and the determination of the pool of questions required per domain.

2.3.2.4 Demographic section

We encountered several challenges during the analysis of the ISWP Basic Test demographic section that may be diminished by reformulating some questions, rethinking the categories, selecting other collecting data procedures, and updating the questions and categories to improve their clarity and to better represent participants' identities. Many questions such as, *'In what country do you live'*, *'What language do you mainly speak at home'*, *'What certifications or licensures do you currently hold if any? Please list below'*, used and open-ended type of question that allowed participants to type their answers. This situation impacted the data cleaning and management of the dataset and can potentially be overcome by changing the question to a closed-ended type. To collect participants' gender identity, the section included the question *'What is your gender?'* with two possible answers 'female' or 'male'. Studies have suggested updating the coding list by including transgender women, transgender male, other, and prefer not to answer option (58, 59). In terms of improving clarity to some questions, we noticed that the question *'What certifications or licensures do you currently hold if any? Please list below'* was confusing; some test-takers wrote their profession (e.g. PT, OT) while others included their continuing education certificates from a wide variety of fields (e.g. Yoga instructor, Cardiopulmonary resuscitation

certificate, Assistive Technology Professional, WHO WSTPs, cure tapping professional, etc); also, participants used many acronyms in the section what were difficult to interpret. We were interested in identifying participants' professional background to evaluate the feasibility of a subgroup analysis; unfortunately, due to the variability and unclarity of responses, we were not able to do it. We consider that the frequency and percentages of participants' professions reported in Table 4, may be underestimated. The future version of the test should contemplate a specific question related to professional background. We recommend including the following questions in the new version of the test:

1) '*What is your country of birth'* in addition to the current question *'In what country do you live?'* that may or may not reflect the place where participants were born.

2) *'What score do you expect to obtain in this test?*' providing options ranging from 0-10, being 10 the highest score possible and 0 the lowest. This question could help explore the possible correlations between participants' self-perceived knowledge and actual knowledge as measure by the ISWP Basic Test. Literature suggests that perceived knowledge is as important as actual knowledge (factual information content) for problem-solving tasks and during the process of acquiring new information and skills (60). Collecting participants' self-perceived knowledge and actual knowledge could provide insights to trainers and educators that may assist in adapting training materials. For instance, if the ISWP Basic Test is used as a pre-test before a training intervention, the trainers could analyze the possible alignment or misalignment of both types of knowledge and plan their training methods accordingly.

3) *'Have you taken this version test before?'* giving three possible answers 'yes', 'no', and 'I do not know'; this question could save some time in grouping subsets of participants.

4) A Likert-type question inquiring confidence levels of knowledge about the training packages used for the development of the test, similar to the question included in the pilot testing (section 2.2.3.2).

2.3.3 Pilot tested set of questions

We received a total of 80 complete test responses. The characteristics of the population showed a sample of medium-age participants, highly educated with 65 (81.5%) of test-takers with at least a bachelor's degree, being PT and OT the most common professions. Participants were geographically distributed in Asia and North America, in contrast with ISWP Basic Test sample analyzed, in which Asia and Africa were the most common regions. Two-thirds of participants (65%) have taken the ISWP Basic test which reflects familiarity with the testing process.

The item level analysis reported that out of the 173 multiple choice questions tested, 61 questions (35.26%), met the *p*-values and *IDI* criteria (Set A); 35 (20.2%) did not meet the *p*-value criteria and 27 (15.6%) did not meet the *IDI* criteria (Set B); and 50 (28.9%) did not meet the *p*-values nor *IDIs* thresholds (Set C). Important factors related to the test revision design and the sample analyzed need to be considered when interpreting these results.

Testing design. We tested the total pool of questions arranging them based on the content domain. There were many questions similarly developed due to the random distribution of questions, a setting established in the ISWP Basic Test (34) to reduce retesting bias (61). The pilot test design included an automatic feedback per question indicating the correct and incorrect answer with an open-ended question afterward where participants can submit comments. This feedback was provided immediately after answering each question. We selected this method to capture possible comments arise from the answer options, the selected correct answer, and the item itself.

The fact that participants knew the correct answer and received similar questions potentially threatens the credibility of the item level analysis results. It is possible that a subject may have selected the correct answer after receiving feedback from a previous answer that inquired about the same topic. In those cases, *p*-values are not capturing the items' true difficulty level and consequently, *IDIs* may not be discriminating between test-takers. Due to this testing design and the fact that all items were included, we consider inappropriate to obtain the total test score of the test.

Sample analyzed. As previously described, the characteristics of the sample indicate that the majority of participants were familiar with the ISWP Basic Test and the testing procedure, were highly educated with more than 80% of the test-takers holding at least a bachelor's degree compared to the 66% presented in the ISWP Basic Test sample, mostly with professional degrees in PT and OT. Although research has shown that some rehabilitation programs reported limited time allocated to wheelchair service provision (18), the knowledge and skills acquired in undergraduate and graduate rehabilitation programs, the testwiseness skills, and the familiarity with regular testing procedures are factors that could contribute to higher test performance (62). Having a sample size of a highly educated professional is a limitation important to note when planning for the future steps of the test. Many places around the globe have limited to no rehabilitation professionals available in their communities or closest health services. In 2011, Gupta et al. conducted a study to describe the global situation of human resources for health-related rehabilitation services and concluded that many LMICs where the disease burden related to disability is higher, showed the lowest supply of rehabilitation health professionals (16). The WHO, aware of this situation, has developed international guidelines to build and strengthen community-based rehabilitation programs and strategies for involving people with disabilities and

personnel without health-related degrees (63). Along with those efforts, the WHO WSTP-B was created to reach all personnel who is expected to carry out basic wheelchair service delivery at their place of work, considering this may include people with and without a professional health-related degree such as physical therapist, occupational therapist, community-based rehabilitation workers, lay health workers, local craftsmen, technicians, wheelchair users, etc (33). It is extremely important that current efforts of improving the ISWP Basic test considered the wide range background of people who may be trained as a basic wheelchair service provider to ensure an accessible assessment tool for all personnel.

2.3.3.1 Proposed next steps

This section provides a list of suggestions developed by the author of this dissertation that may assist in planning for future work and finalizing the development of the new version of the ISWP Basic Test. The following suggestions contemplate the ATC as the leading group to complete this process.

<u>Results from the pilot set of questions</u>

- Retain the questions from Set A and B to form a pool of questions of 123 items. Due to the testing design and the characteristics of the sample analyzed previously described, it is likely that questions from Set B, which met one of the item cutoffs, could perform within expected *p*-values and *IDIs* cutoffs in real testing scenarios with a heterogeneous sample of test-takers. The two sets include questions from both subdomains, 15 (12.2%) from Core Knowledge and 108 (87.8%) from Wheelchair Service Steps, include questions from all first level components, and all second level components except for 'Fitting: general information", 'User training: pressure injury prevention', and 'Follow-up and repairs: repairs' (Table 11). This information could advise on the number of questions and their

respective subdomain that will need to be developed to reach the future weighting of the test.

- Review the questions from Set C following the revision process described in section 2.1.2. This Set consists of 50 questions, 14 (28%) from Core Knowledge and 36 (72%) from the Wheelchair Service Steps subdomain. As listed in Table 11, the Set reported the highest frequency of questions allocated in Core Knowledge and the lowest frequency of questions in the Wheelchair Service Steps subdomain when compared with the other sets. A possible explanation could be related to the highly educated sample of participants from the pilot test who may have found no difficulty in the Core Knowledge section. Once the revision process is concluded, the reviewed and retained questions could be included in the test with no value since they will be considered in the testing phase.
- Review the comments received from all Sets of questions. ATC members could work with the same group and domain to review the comments arise from the developed questions. This process could be a learning opportunity for members to evaluate the results from their efforts. The revision should look for aspects that may suggest a better word selection, paraphrasing, better distractors, and/or any issues with the questions.

Determine domains' weighting

The ATC in conjunction with a test development consultancy agency will determine the weighting of the test domain. Based on the decision, the current pool of questions (Set A and B) will be analyzed to determine if more questions need to be developed and if a second round of testing is necessary; if that is the case, the questions from Set C could be included.

Establish a periodic test revision.

The next complete test revision will occur when either of the following events happens first: 1) one year since the distribution of the new test or 2) when a sample size of at least 300 test-takers is reached.

3.0 Specific Aim 2. To develop and establish psychometric evidence of a wheelchair service provision self-assessment capacity skills questionnaire – basic level (WSP CAPS-B) based on the WHO WSTP-B

3.1 Methods

3.1.1 Development of the survey

This dissertation author (AYBM) and her primary advisor (MG) independently listed all skills related to wheelchair service provision included in the WHO WSTP-B and developed the survey's questions. The authors' professional background is in PT and Education; both have experience delivering wheelchair service provision training according to the WHO WSTPs and developing educational programs for high- and low- resourced settings.

The survey consists of 31 questions grouped by the authors on six theoretical constructs: 1) basic mobility skills, 2) transfers, 3) wheelchair service steps, 4) wheelchair safe and ready, 5) cushion fabrication and wheelchair adjustment, and 6) wheelchair users' training. The authors followed best practices for the development of self-efficacy scales which included: A) The use of a 100-point scale from 0 ("Cannot do"); through middle degrees of assurance, 50 ("Moderately certain can do"); to complete assurance, 100 ("Highly certain can do") (64, 65); B) the use of words in the survey questions focused on perceived capability (64); C) the implementation of unipolar gradations (64); and; D) the inclusion of instructions that asked participants to rate their current operative capabilities and not their potential capabilities or their expected future capabilities (64). In addition, authors attempted to minimize response biases by informing all

participants that their responses will remain confidential and protected by the research staff (64). Due to the bilingualism of the primary developer, the survey was created simultaneously in English and Spanish, but it was decided to be tested in Spanish with a sample of native Spanish speakers. The questionnaire was hosted in Qualtrics®, a survey software system that the University of Pittsburgh has a license to use.

3.1.2 Study population

A convenience sampling method was used to invite Mexican Universities with PT and OT programs and public clinics to participate in this study. The research team held meetings with the chairs of the Departments of PT and OT and with clinical coordinators from public clinics to share the scope of the project and to invite them to participate. Three universities, two public and one private, and one public clinic located across 4 states agreed to participate in this project. The clinic was staffed by students from Universities of other states. This a common practice in Mexico, where students completing their social service (one year on supervised practice) are relocated across the country to cover less-resourced regions and to service vulnerable populations.

PT and OT students who completed at least 70% of their curriculum and were currently enrolled in their programs were considered eligible to participate in this study. Professors and chairs of each university and the clinical coordinator of the public clinic shared the details of the project and invited their students to participate voluntarily.

3.1.3 Data collection

Eligible students received an email with the link to the Spanish version of the questionnaire. The questionnaire included a confidentiality statement, the description of its purpose, instructions, and the research staff's contact information for questions or technical problems. Upon agreement to participate, students were asked to complete the questionnaire in two-time points with a minimum of 21 days in between data collection. The first survey was distributed September 18, 2018 and participants were instructed to complete the survey within 10 days (time 1). Three weeks later, on October 9, 2018 participants were asked to retake the survey and complete it within 10 days (time 2). Only the data from participants in time 2 who completed time 1 was retained for analysis.

3.1.4 Statistical analysis

3.1.4.1 Internal Validity Evidence

Initially, the factorability of the 31 items was examined. To test the suitability of the data for factor analysis, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity were calculated (66, 67). After confirming data suitability, construct validity was calculated by Exploratory Factor Analysis (EFA) using a Principal Component Analysis (PCA) with varimax rotation. The number of factors to retain was determined based on the Cattell's Scree test and the Kaiser criterion. The Catell's Scree test is a visual exploration of a graphical representation of the eigenvalues that are plotted in descending order, and the graph is constructed following this way (68-70). The Kaiser criterion sets a threshold between large and small values at an eigenvalue of 1, the arithmetic mean of eigenvalues. Each eigenvalue greater than 1 is

interpreted as representing factor, and each value below 1 is discarded (67, 69). To determine factors' items, the highest factor load per item was considered. We applied varimax rotation to the analysis because the first factors were not clear based on technical criteria.

3.1.4.2 Internal Consistency Reliability

The Internal Consistency (IC) and reliability of the survey were assessed by the Cronbach's alpha (α). The IC was assessed by item-total correlation (ITC) using Pearson's correlation coefficient (*r*) which involved measuring the strength of the association between an individual item and the total score of the remainder of its scale. The correlation between each individual item and the domain and/or global scores omitting the item was assessed. In addition, the IC was calculated in each subscale created after the EFA. The reliability indicates the homogeneity of the responses of the participants in all the items of the scale. This coefficient allowed the authors to determine the correlation among survey items in order to measure the construct. An α value greater than 0.70 was considered satisfactory (71, 72).

3.1.4.3 Test-retest

A descriptive analysis was performed for each time; the Spearman coefficient was calculated to evaluate the correlation between time 1 and time 2 (73). For agreement, the intraclass correlation coefficient was estimated using a single-measurement, absolute agreement, two-way mixed-effects model and Bland-Altman plot was analyzed (73, 74). All statistical analyses were carried out using Stata 14®.

3.1.5 Ethics

This study was approved by the University of Pittsburgh Institutional Review Board and by the Ethics Committee of each local University and the public clinic. All participants agreed to participate in the study voluntarily and consented to use their deidentified data for research purposes.

3.2 Results

3.2.1 Sample analyzed

A total of 163 students were surveyed, 123 (75.5%) female and 40 (24.5%), male with a mean age of 22.26 years old (SD=4.17), in time 1. In time 2, 113 students completed the survey, 94 (83.2%) female and 19 (16.8%) male with a mean age of 22.9 years old (SD=4) responded the survey (Table 13).

Table 13 Characteristics of the study population

Characteristic	**T1 (N=163)**	**T2 (N=113)**
Age, mean (SD)	22.26 (4.17)	22.29 (4)
Sex, Female, n (%)	123 (75.5)	94 (83.2)
Field of study and university or clinic, n (%)		
Occupational Therapy	48 (29.4)	36 (31.9)
Private	25 (15.3)	17 (15)
Public	23 (14.1)	19 (16.8)

Table 13 (continued)

Physical Therapy	115 (70.6)*	77 (68.1)*
Private	105 (64.4)	67 (59.3)
Public	9 (5.5)	9 (8)

SD: Standard deviation

*One participant did not include University

3.2.2 Internal Validity Evidence

The KMO measure of sampling adequacy was 0.94, considered as an "excellent" factor-analytic data (66, 75). Bartlett's test of sphericity was also significant ($\chi 2$ (465) = 5234, $p < .05$) (76). Given these overall indicators, factor analysis was considered suitable for all 31 items.

The scree plot indicated that five factors explained 75.7% of the total variability of the scale were retained (Figure 3). The factor loading matrix is presented in Table 14.

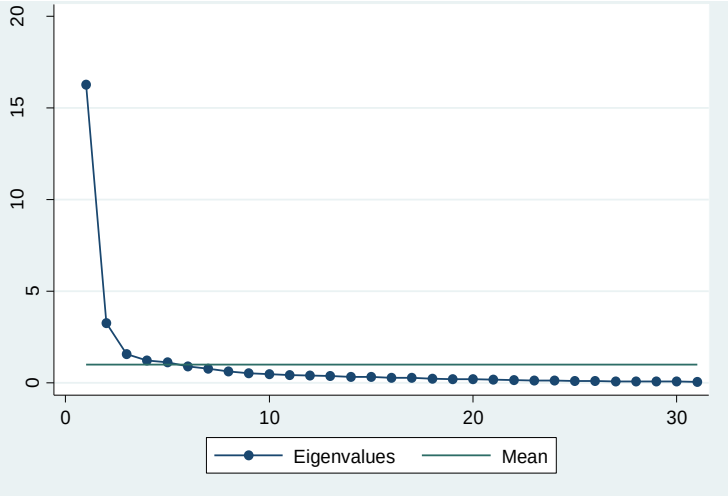


Figure 3 Cattell's scree test and the Kaiser criterion

Table 14 Factor loading matrix of principal components analysis with varimax rotation

Item	Factors				
	1	**2**	**3**	**4**	**5**
1				0.5207	
2				0.5077	
3				0.3549	
4				0.2276	
5				0.4480	
6			0.3467		
7			0.3033		
8			0.4159		
9			0.4823		
10			0.4911		
11		0.2511			
12	0.1954				
13					0.2931
14					0.3235
15					0.2103
16	0.3382				
17	0.2856				
18	0.2453				
19	0.2598				
20	0.2368				
21	0.3710				
22	0.3367				
23	0.2575				
24	0.3322				
25		0.4825			
26		0.4062			
27		0.3294			

Table 14 (continued)

28					0.4412
29					0.4794
30					0.3176
31		0.3644			

3.2.3 Internal Consistency Reliability

Table 15 presents the factors' regrouping considering the factor loading matrix and the results from the internal consistency analysis. We obtained values above 0.8, showing a good correlation between items. Factor 4, related to the construct of Basic mobility skills, showed the lowest internal consistency with $\alpha = 0.85$.

Table 15 Internal consistency of the empirical factors.

Factor	Construct	Item	A
1	Assessment interview and wheelchair safe and ready	12, 16, 17, 18, 19, 20, 21, 22, 23, 24,	0.964
2	Cushion – pressure check, fabrication and maintenance (including the wheelchair)	11, 25, 26, 27, 31	0.914
3	Transfers	6, 7, 8, 9, 10	0.922
4	Basic mobility skills	1, 2, 3, 4, 5	0.854
5	Measurements, adjustments, and wheelchair user's training	13, 14, 15, 28, 29, 30	0.884

Overall, the survey shows a Cronbach's alpha 0.97, suggesting a very good internal correlation. Two methods were conducted to evaluate items' discerning capacity: 1) A descriptive analysis looking for answers with high percentages, which would indicate the low capacity of the questions to differentiate answers with high and low score. In this case, the response given to all

items was homogeneous; no item was greater than 80%; and 2) we calculated the corrected ITC to evaluate the relationship of each item with the sum of the rest of the items. Table 16 presents the calculation of each item of the scale. Of the 31 items, all presented a good correlation, values > 0.3, indicated a well-discerning capacity. In addition, we examined the contribution of each item to the internal consistency of the scale. We removed each item, calculated the Cronbach's alpha and then compared it with the overall Cronbach's alpha to determine the changes when the item was absent. In this case, question 5_1 increased the general alpha in 0.0001. In spite of this, the increase was not considered significant to remove the item from the scale.

Table 16 WSP CAPS-B survey's items and internal consistency analysis

Code	Item	Median	Rank	Corrected item-total correlation	Cronbach's alpha if item deleted
	Basic mobility skills (Realizar habilidades básicas de movilidad en una silla de ruedas manual)				
Q5_1	1. I can push the wheelchair. (Puedo impulsar la silla de ruedas).	100	100	0.5275	0.9688
Q5_2	2. I can do turns in the wheelchair. (Puedo girar la silla de ruedas).	88	99	0.6075	0.9683
Q5_3	3. I can go up and down slopes. (Puedo subir y bajar pendientes en la silla de ruedas).	51	100	0.6099	0.9683
Q5_4	4. I can go up and down stairs with assistance. (Puedo subir y bajar escalones con asistencia).	32	100	0.6719	0.9679
Q5_5	5. I can do partial wheelies (lift the front wheels). (Puedo hacer equilibrio en las ruedas traseras (levantar las ruedas delanteras)).	30	100	0.4023	0.9696
	Transfers (Realizar transferencias)				
Q8_1	6. I can do an independent transfer from the wheelchair to the bed. (Puedo realizar una transferencia independiente de la silla de ruedas a la cama).	50	100	0.7288	0.9676
Q8_2	7. I can do an assisted transfer with a transfer board from the wheelchair to the bed. (Puedo realizar una transferencia asistida con una tabla de transferencia de la silla de ruedas a la cama).	49	100	0.6629	0.9680
Q8_3	8. I can do an assisted standing transfer from the bed to the wheelchair. (Puedo realizar una transferencia asistida de pie, de la cama a la silla de ruedas).	60	100	0.6569	0.9680
Q8_4	9. I can do an independent transfer from the wheelchair to the floor. (Puedo realizar una transferencia independiente de la silla de ruedas al piso).	50	100	0.7074	0.9677
Q8_5	10. I can do an independent transfer from floor to wheelchair. (Puedo realizar una transferencia independiente del piso a la silla de ruedas).	44	100	0.6680	0.9680

Table 16 (continued)

	Wheelchair service steps (Servicios de sillas de ruedas)				
Q9_1	11. I can do a pressure test to check if the pressure relief cushion is working (the test consists in placing your fingertips under the wheelchair user seat bones, ask the user to sit back down your fingers and check how much your fingers can move). (Puedo realizar una prueba de presión para comprobar si un cojín de alivio de presión funciona bien (la prueba consiste en colocar la punta de los dedos debajo de los isquiones del usuario, pedirle al usuario que se siente sobre los dedos, e identificar cuánto se mueven los dedos)).	47	100	0.7247	0.9676
Q9_2	12. I can carry out an assessment interview to obtain relevant information about the wheelchair user (for example, physical condition, lifestyle and environment, characteristics of the existing wheelchair, presence, risk of, or history of pressure sores, method of pushing, wheelchair user's objectives). (Puedo realizar una entrevista al usuario de sillas de ruedas para obtener información relevante (por ejemplo: historia clínica, estado físico, estilo de vida y entorno, características de la silla de ruedas en uso, presencia riesgo e historial de úlceras por presión, método de propulsión, objetivos del usuario, etc.)).	90	98	0.6585	0.9680
Q9_3	13. I can take basic body measurements to the wheelchair user (hip width, seat depth, calf length, backrest height). (Puedo realizarle mediciones básicas al usuario de sillas de ruedas (ancho de cadera, profundidad del asiento, longitud de la pantorrilla, y altura del espaldar)).	60	100	0.7045	0.9677
Q9_4	14. I can measure a wheelchair (seat width, seat depth, backrest height, footrest height, frame length, wheelbase). (Puedo tomarle medidas a la silla de ruedas (ancho, profundidad y altura del asiento, altura del espaldar, altura del apoyapiés, longitud del armazón, distancia entre ejes)).	60	100	0.7520	0.9674
Q9_5	15. I can adjust a wheelchair (for example, backrest height, armrest height, footrest height, push handles height, rear wheels and brakes position). (Puedo realizar ajustes a la silla de ruedas (por ejemplo: altura del espaldar, apoyabrazos, apoyapiés, manillas de empujar; cambiar la posición de las ruedas traseras y frenos)).	50	100	0.7544	0.9674

Table 16 (continued)

	Wheelchair safe and ready (Silla de ruedas lista y segura)				
Q10_1	16. I can check the whole wheelchair (that there are no sharp edges, no parts damaged or scratched, and that the wheelchair travels in a straight line). (Puedo revisar toda la silla (que no tenga bordes afilados, partes dañadas o rayadas, que se mueve en línea recta)).	73	100	0.7566	0.9674
Q10_2	17. I can check the front castor wheels (that they spin freely, spin without touching the fork, and that bolts are tight). (Puedo revisar las ruedas orientables (que giren libremente y sin tocar la horquilla, que los pernos estén ajustados)).	52	100	0.8114	0.9670
Q10_3	18. I can check the front castor barrels, that the castor fork spins freely (Puedo revisar que los ejes verticales de las ruedas orientales giren libremente).	50	100	0.8378	0.9669
Q10_4	19. I can check the rear wheels (that they spin freely, the axle bolts are tight, the tires are inflated correctly, the push rims are secure). (Puedo revisar las ruedas traseras (que giren libremente, que los pernos de los ejes estén firmes, los neumáticos estén inflados correctamente, los aros propulsores estén ajustados)).	53	100	0.8440	0.9668
Q10_5	20. I can check adjustments made to the wheelchair (for example, backrest height, armrest height, footrest height, push handles height, rear wheels and brakes position). (Puedo revisar ajustes hechos a la silla de ruedas (por ejemplo: altura del espaldar, apoyabrazos, apoyapiés, manillas de empujar; la posición de las ruedas traseras y frenos)).	50	100	0.8012	0.9671
Q10_6	21. I can check that the brakes function properly. (Puedo revisar que los frenos funcionen adecuadamente).	77	100	0.7432	0.9675
Q10_7	22. I can check that footrests are securely attached. (Puedo revisar que los apoyapiés estén bien ajustados).	75	100	0.7978	0.9671
Q10_8	23. I can check that the frame folds properly. (Puedo revisar que el armazón se pliegue adecuadamente).	60	100	0.8441	0.9668
Q10_9	24. I can check that the cushion and its cover have the proper size and that fully cover the seat. (Puedo revisar que el cojín y su funda tengan un tamaño adecuado y que el cojín cubra el asiento por completo).	75	100	0.7693	0.9673
	Cushion fabrication and wheelchair adjustment (Fabricación de cojín ajustes a la silla de ruedas)				

Table 16 (continued)

Q20_1	25. I can make a basic foam pressure relief cushion. (Puedo fabricar un cojín liberador de presión con espuma firme y suave).	31	100	0.6865	0.9678
Q20_2	26. I can check the fitting of a wheelchair of an individual wheelchair user (check the wheelchair size and the adjustments, the wheelchair user posture, that the cushion pressure is adequate, and the wheelchair fit while the wheelchair user is moving. (Puedo realizar una prueba de ajuste del usuario con su silla (verificar el tamaño y los ajustes hechos a la silla, la postura del usuario, que la presión del cojín sea adecuada, así como el ajuste de la silla mientras el usuario se mueve)).	40	100	0.7732	0.9673
Q20_3	27. I can solve common problems caused of fitting or adjustment (for example, seat depth too short or too long, footrests too high, legs tend to roll inwards/outwards, wheelchair is too wide, feet tend to slide off the footrests, etc.). (Puedo resolver problemas comunes de calce o ajuste (por ejemplo: el asiento es poco o demasiado profundo, los apoyapiés están demasiado altos, las piernas se deslizan hacia afuera/dentro, la silla es demasiado ancha, los pies se resbalan del apoyapiés, etc.)).	50	100	0.7894	0.9672
	Wheelchair users' training (Capacitación del usuario de silla de ruedas)				
Q12_1	28. I can train wheelchair mobility to the wheelchair user. (Puedo capacitar al usuario de silla de ruedas en cómo movilizarse en su silla de ruedas).	52	100	0.7362	0.9675
Q12_2	29. I can train the wheelchair user how to do transfers. (Puedo capacitar al usuario de silla de ruedas en cómo realizar transferencias).	50	100	0.7461	0.9675
Q12_3	30. I can train the wheelchair user on how to prevent pressure sores and what to do if a pressure sore develops. (Puedo capacitar al usuario de silla de ruedas en cómo prevenir úlceras por presión y qué hacer cuando se forman).	53	100	0.7401	0.9675
Q12_4	31. I can train the wheelchair user on how to care for a wheelchair and cushion at home (for example, how to clean the wheelchair and cushion, oil moving parts, pump up the tires, tighten nuts and bolts, tighten spokes, check the cushion, etc.). (Puedo capacitar al usuario de silla de ruedas en cómo cuidar su silla de ruedas y el cojín en casa (por ejemplo: lavar la silla de ruedas y el cojín, lubricar las piezas móviles, inflar los neumáticos, ajustar pernos y tuercas, ajustar los rayos, examinar el cojín, etc.)).	43	100	0.7412	0.9675

3.2.4 Test-retest

Time 1 and 2 medians were 1700 and 1720, interquartile range (IQR) was 1144 and 1263 points, respectively. Results of the Spearman correlation indicated a strong, statistically significant positive association between time 1 and 2, r_s (113) = 0.71, p <0.05. (Figure 4) (73). Likewise, data was analyzed using a single-measurement, absolute agreement, two-way mixed effects models obtaining and intraclass correlation index (ICC) of 0.73 with 95% confidence interval of 0.63-0.80 which represents "moderate" to "good" reliability (74). The Bland Altman plot illustrates the global agreement between the two measurements, time 1 and time 2. The mean differences between the two times are close to zero. Few values are outside the 95% confidence bands when the average between time 1 and time 2 is around 1000 and 2500, and those values do not show a particular behavior pattern (Figure 5) (77).

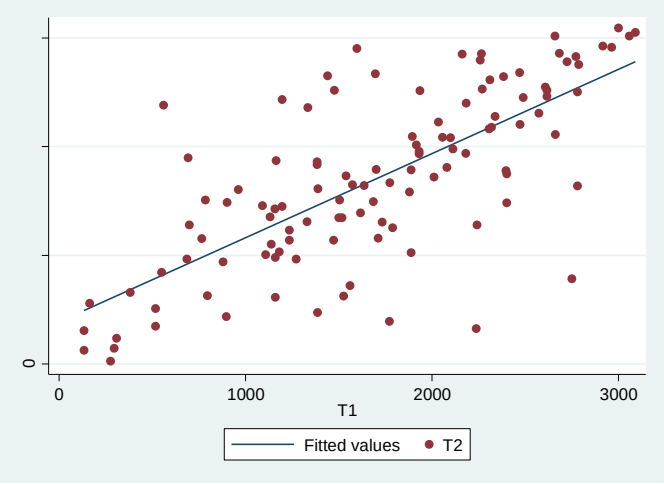


Figure 4 Scatter plot between time 1 and time 2

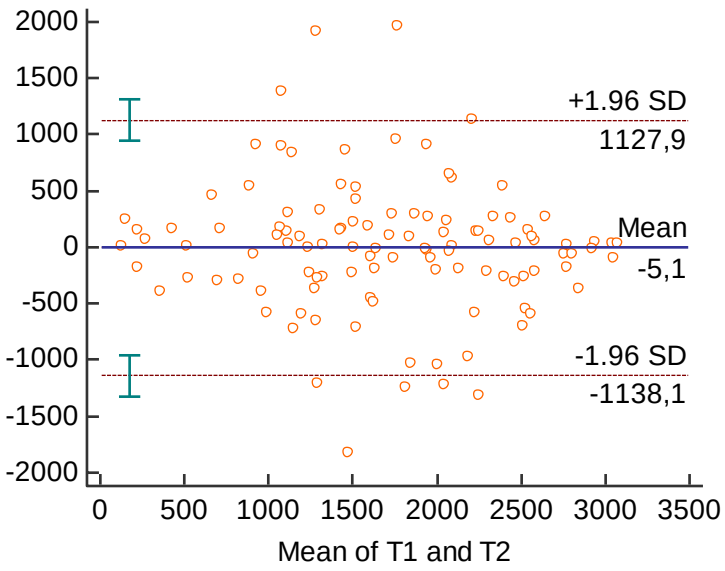


Figure 5 Bland Altman chart

3.3 Discussion

To our knowledge, this is the only published study to date that developed and psychometrically evaluated a self-report capacity questionnaire of all basic wheelchair service provision skills recognized by international guidelines. Previous studies have developed self-reported questionnaires of specific wheelchair service provision steps, such as wheelchair skills (44, 78, 79), but have not included other aspects of the wheelchair service delivery process.

The development of a new instrument requires the assessment of its psychometric properties such as construct, content, and criterion validity (80). Even though the survey development process did not include a subject matter expert revision, the developers' experience in survey design and expertise in the field resulted in the development of a tool with good reliability and agreement, in which all items were retained because none showed significant changes in the Cronbach's alpha when they were eliminated. The construct validity of the scale was assessed by

EFA using a PCA with varimax rotation. The subscales obtained showed consistency empirically and conceptually and each subscale presented high internal consistency (greater than 85%), resulting in good reliability evidence of each construct. The stability of the scale over time showed a good agreement between the two measurements made. Similarly, the reliability between measurements, reported by the Spearman correlation coefficient, reported a strong and positive relationship.

We used the content of the WHO WSTP-B to guide the development of the tool. This training package was developed by an international group of subject matter experts in wheelchair service provision and stakeholders. The package is open-source, available in multiple languages, including in Spanish, and includes training materials. Some research teams have identified limited content and training time for various wheelchair service provision steps. For instance, the WSP in its most recent version included 32 mobility skills for manual wheelchair users as opposed to the 7 skills encompassed in the WHO WSTP-B (23, 42). In addition, the WSP research team has conducted substantial studies where different training modalities have been tested among groups with diverse characteristics. All of their studies suggest more training time to acquire wheelchair skills than the 70 minutes allocated in the WHO WSTP-B (23, 81). Another training package, the WMTP, considered a more extensive list of maintenance activities to be taught to wheelchair users and caregivers than the ones included in the WHO WSTP-B.

Although we recognized the limitations of the WHO WSTP-B, we selected the package due to its presence worldwide, its availability in multiple languages, and most importantly, because it encompasses all wheelchair service delivery steps. Nevertheless, we encourage the WHO to consider the findings from these and other related studies to inform the development of future versions of the WHO WSTPs. As future versions of WHO WSTPs become available and more

studies are published that suggest the importance of particular wheelchair provision skills, the 31 skills represented on this tool will need to be reviewed, and additional skills may need to be added to reflect the research findings.

We decided to test the questionnaire among a group of PT and OT Mexican students that have completed at least 70% of their curriculum and were currently enrolled in their programs. Depending on the University, some students were undertaking their social service while others were about to start it. The social service in Mexico is a professional requirement in which students from health-related fields complete one year of supervised practice in less-resourced regions where vulnerable populations are concentrated. We selected this group of participants due to the strong network we have in the country and because at that time, the research team was conducting another study in Mexico. We had a high dropout rate of 50 (30.7%) participants between time 1 and time 2. A possible explanation is that participants did not perceive any benefit nor time retribution to participate in the study. We wanted to ensure that all students voluntarily participate in the study and we asked professors not to link the invitation to any activity related to their curriculum. Reflecting on this decision, we could have provided more details about the scope of the project and offered other types of incentives such as access to continuing education modules hosted at the ISWP website.

To continue to determine the utility of this tool for measuring basic wheelchair provision skills, future studies may investigate how self-ratings compare to experts' ratings of that person's performance on demonstrating the various skills and if increased providers' skills are linked with patient satisfaction and better outcomes. This may suggest the extent to which test-takers over- or underestimate their ability to perform the particular tasks and infer how the tools' results reflect fidelity to practice. Additionally, these results may also indicate whether this tool alone can

accurately represent skill performance or whether additional skill assessments will generally be required. Furthermore, if increased providers' wheelchair skills are linked with patient satisfaction and better outcomes, these findings could advocate for more training time in wheelchair skills provision skills in PT and OT programs.

3.3.1 Limitations

Important limitations need to be considered when interpreting the results from this study. This work established psychometric evidence of a self-reported basic skills questionnaire that should not be interpreted as validity evidence. This tool is in its development phase and further confirmatory analysis over repeated samples are needed to establish validity evidence.

Despite the survey was developed simultaneously in English and Spanish, it was exclusively tested in Spanish. A formal evaluation of the translation equivalence and testing of the English version survey are needed. Similarly, the test was developed by a native Mexican Spanish speaker and tested solely in Mexico, which may provide some bias towards the Mexican vocabulary and limit the generalizability to other Spanish regions including, but not limited to, those spoken across Central America, Caribbean, South America and those across Spain and around the world.

4.0 Conclusions

Evidence highlights the insufficient training time professional rehabilitation programs allocate to wheelchair service provision and the association of limited competency with inappropriate wheelchair distribution (6, 7, 14, 18, 19). This situation has adverse effects on the life of wheelchair users and the full and equal enjoyment of their human rights (17). The Wheelchair Sector has called to strengthen and support competency in wheelchair service personnel as part of the priority actions to achieve by 2023 (20).

This work developed and validated two assessment tools that can support the evaluation of wheelchair service providers in international settings. We acknowledge that testing competency in wheelchair service provision is a multidimensional construct that involves the interaction between knowledge and skills, and for that reason, efforts were focused on improving a current knowledge test and taking the first steps towards developing a self-report wheelchair skills questionnaire. Considering the global shortage of rehabilitation health professionals, particularly in LMICs, we selected the WHO WSTP-B as the primary training package to develop the assessment tools. The WHO WSTP-B has a target audience that includes personnel without professional degrees, it has been widely distributed around the world, and translated into several languages. The ISWP Basic Test, which has been used internationally for 5 years in its original form, was updated using another evidence-based training package, the WSP. We did not include the WSP in the development of the WSP CAPS-B since there are validated assessment tools that measure objective and subjective wheelchair skills capacity and because the purpose of our questionnaire was to measure all basic wheelchair provision skills.

Language barriers have been recognized as important limitations to scaling up interventions and reducing disparities between LMICs and high-income countries (82, 83). We were interested in developing assessment tools in English and Spanish because those languages represent the two most requested ISWP Basic Tests versions. ISWP has a broad group of native English and Spanish members across low- to high-income countries that could benefit from the immediate adoption of these assessment tools. Moreover, this large sample of members could help to enhance the tool and conduct further research before other translations are conducted.

The two proposed assessment tools in this work have the potential to be used in different ways. The tools can be incorporated as pre- and post-training intervention assessments. In this way, they could inform baseline knowledge and self-perceived capacity which may assist educators in adapting the training to respond to the groups' needs. In professional rehabilitation programs, the tools can be used at different time points across the curriculum to monitor students' performance and advise the need for program modifications. ISWP could integrate both assessment tools as part of the WSP certification and start a formal translation process focused on cross-cultural equivalence to provide the tools in other languages.

The ultimate goal of improving competency in wheelchair service providers is to support good practice and ensure that wheelchair users receive a product that meets their needs and promotes the full and equal enjoyment of all human rights. Future work could evaluate the possible associations between trained/untrained providers and client center outcomes.

I believe that developing appropriate tools for measuring knowledge and skills could inform the training needs of the personnel involved in wheelchair service provision, establish competency expectations, and consequently, improve the quality of care.

Appendix A Question's Guidelines

Table 17 Question's Guidelines

Guidelines	Score: Y/N*
Question Text	
Stem is meaningful by itself and presents a definite problem.	
Stem does not contain irrelevant material.	
Stem is a question or partial sentence.	
Stem is negatively stated only when significant learning outcomes require it. Negative statement is in italics or capitalization.	
Stem targets a specific cognitive process.	
Answers' options	
All alternatives are plausible.	
Alternatives are stated clearly and concisely.	
Alternatives are mutually exclusive.	
Alternatives are homogenous in content.	
Alternatives are free from clues.	

* Dichotomously scored, Y=yes, N=no.

Appendix B - E

If you are interested in accessing the following appendixes

- B – Basic Test Results
- C – Pilot's Total pool of Questions
- D – Pilot Tested Questions Results
- E – Pilot Questions' Comments

Please contact Dr. Mary Goldberg at mgoldberg@pitt.edu

Bibliography

1.World Health Organization. World Report on Disability. Geneva; 2011.

2.World Health Organization. Guidelines on the provision of manual wheelchairs in less resourced settings. Geneva: WHO; 2008.

3.World Bank. Disability Inclusion 2018 [Available from: http://www.worldbank.org/en/topic/disability.

4.United Nations. The Universal Declaration of Human Rights [Available from: http://www.un.org/en/universal-declaration-human-rights/.

5.Toro ML, Eke C, Pearlman J. The impact of the World Health Organization 8-steps in wheelchair service provision in wheelchair users in a less resourced setting: a cohort study in Indonesia. BMC Health Serv Res. 2016;16:26.

6.Borg J, Lindstrom A, Larsson S. Assistive technology in developing countries: national and international responsibilities to implement the Convention on the Rights of Persons with Disabilities. Lancet. 2009;374(9704):1863-5.

7.Visagie S, Scheffler E, Schneider M. Policy implementation in wheelchair service delivery in a rural South African setting. Afr J Disabil. 2013;2(1):63.

8.Carver J, Ganus A, Ivey JM, Plummer T, Eubank A. The impact of mobility assistive technology devices on participation for individuals with disabilities. Disabil Rehabil Assist Technol. 2016;11(6):468-77.

9.United Nations, Division for Social Policy and Development Disability. Convention on the Rights of Persons with Disabilities – Articles. Geneva: United Nations; 2006 [cited 2019 Feb 09]. Available from: https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/convention-on-the-rights-of-persons-with-disabilities-2.html.

10.Mukherjee G, Samanta A. Wheelchair charity: a useless benevolence in community-based rehabilitation. Disabil Rehabil. 2005;27(10):591-6.

11.Greer N, Brasure M, Wilt TJ. Wheeled mobility (wheelchair) service delivery: scope of the evidence. Ann Intern Med. 2012;156(2):141-6.

12.Economic Commission for Latin America and the Caribbean (ELAC). Social Panorama of Latin America. Briefing paper. 2012.

13.Banks LM, Kuper H, Polack S. Poverty and disability in low- and middle-income countries: A systematic review. PLoS One. 2017;12(12):e0189996.

14.McSweeney E, Gowran RJ. Wheelchair service provision education and training in low and lower middle income countries: a scoping review. Disabil Rehabil Assist Technol. 2017:1-13.

15.Elwan A. Poverty and disability: A survey of the literature. Washington, DC.: Social Protection Unit, Human Development

Network; 1999.

16.Gupta N, Castillo-Laborde C, Landry MD. Health-related rehabilitation services: assessing the global supply of and need for human resources. BMC Health Serv Res. 2011;11:276.

17.United Nations, Division for Social Policy and Development Disability. Convention on the Rights of Persons with Disabilities (CRPD) New York: United Nations; 2006 [cited 2018 Jan 3]. Available from: https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html.

18.Fung KH, Rushton PW, Gartz R, Goldberg M, Toro ML, Seymour N, et al. Wheelchair service provision education in academia. Afr J Disabil. 2017;6:340.

19.Williams E, Hurwitz E, Obaga I, Onguti B, Rivera A, Sy TRL, et al. Perspectives of basic wheelchair users on improving their access to wheelchair services in Kenya and Philippines: a qualitative study. BMC Int Health Hum Rights. 2017;17(1):22.

20.World Learning. Wheelchair Stakeholders' Meeting: Executive summary 2018 [Available from: https://wheelchairnetwork.org/wp-content/uploads/2019/08/01a-WL-Wheelchair-Two-Pager-FINAL.pdf.

21.Fullerton JT, Thompson JB, Johnson P. Competency-based education: the essential basis of pre-service education for the professional midwifery workforce. Midwifery. 2013;29(10):1129-36.

22.Goudreau J, Pepin J, Dubois S, Boyer L, Larue C, Legault A. A second generation of the competency-based approach to nursing education. Int J Nurs Educ Scholarsh. 2009;6:Article15.

23.World Health Organization. Wheelchair Service Training Package: Basic Level. Geneva: WHO; 2012.

24.World Health Organization. Wheelchair Service Training Package: Intermediate Level. Geneva: WHO; 2013.

25.World Health Organization. Wheelchair Service Training Package: Managers. Geneva: WHO; 2015.

26.World Health Organization. Wheelchair Service Training Package: Stakeholders Geneva: WHO; 2015.

27.Kirby RL, Doucette SP. Relationships between wheelchair services received and wheelchair user outcomes in less-resourced settings: a cross-sectional survey in Kenya and the Philippines. Archives of Physical Medicine and Rehabilitation. 2019;100(9):1648-54. e9.

28.Borg J, Larsson S, Östergren P-O, Rahman AA, Bari N, Khan AN. User involvement in service delivery predicts outcomes of assistive technology use: A cross-sectional study in Bangladesh. BMC health services research. 2012;12(1):330.

29.Visagie S, Mlambo T, Van der Veen J, Nhunzvi C, Tigere D, Scheffler E. Impact of structured wheelchair services on satisfaction and function of wheelchair users in Zimbabwe. African Journal of Disability. 2016;5(1).

30.Shore S. The long-term impact of wheelchair delivery on the lives of people with disabilities in three countries of the world. African Journal of Disability (Online). 2017;6:1-8.

31.Accelovate. Wheelchair use and services in Kenya and Philippines: a cross-sectional study. [Available from: http://reprolineplus.org/system/files/resources/wheelchair-study-report-2015Dec.pdf.

32.Bazant ES, Himelfarb Hurwitz EJ, Onguti BN, Williams EK, Noon JH, Xavier CA, et al. Wheelchair services and use outcomes: A cross-sectional survey in Kenya and the Philippines. African Journal of Disability (Online). 2017;6:1-14.

33.World Health Organization. Wheelchair Service Training Package. Trainer's Manual Basic Level. Geneva: WHO; 2012. 248 p.

34.Gartz R, Goldberg M, Miles A, Cooper R, Pearlman J, Schmeler M, et al. Development of a contextually appropriate, reliable and valid basic Wheelchair Service Provision Test. Disabil Rehabil Assist Technol. 2016;12(4):333-40.

35.International Society of Wheelchair Professionals (ISWP). ISWP Wheelchair Service Provision Basic Test by region 2018 [Available from: http://wheelchairnetwork.org/wp-content/uploads/2017/07/Basic-Test-takers-by-region_12June2018.pdf.

36.Burrola-Mendez Y, Goldberg M, Gartz R, Pearlman J. Development of a Hybrid Course on Wheelchair Service Provision for clinicians in international contexts. PLoS One. 2018;13(6):e0199251.

37.Burrola-Mendez Y, Toro-Hernandez ML, Goldberg M, Pearlman J. Implementation of the hybrid course on basic wheelchair service provision for Colombian wheelchair service providers. PLoS One. 2018;13(10):e0204769.

38.Burrola-Mendez Y, Bonilla-Escobar FJ, Goldberg M, Pearlman J. Comparing the effectiveness of a hybrid and in-person courses of wheelchair service provision knowledge: A controlled quasi-experimental study in India and Mexico. PLoS One. 2019;14(5):e0217872.

39.Toro-Hernández ML, Mondragón-Barrera MA, Torres-Narváez MR, Velasco-Forero SE, Goldberg M. Undergraduate physiotherapy students' basic wheelchair provision knowledge: a

pilot study in two universities in Colombia. Disability and Rehabilitation: Assistive Technology. 2019:1-6.

40.University of Pittsburgh Model Center on Spinal Cord Injury. Wheelchair Maintenance Training Program 2017 [Available from: http://www.upmc-sci.pitt.edu/node/927.

41.Dalhousie University. Wheelchair Skills Program 2019 [Available from: https://wheelchairskillsprogram.ca/en/skills-manual-forms/.

42.Kirby RL RP, Smith C, Routhier F, Axelson PW, Best KL, Betz K, Burrola-Mendez Y, Contepomi S, Cowan R, Giesbrecht E, Kenyon LK, Koontz A, MacKenzie D, Mortenson B, Parker K, Smith E, Sonenblum S, Tawashy A, Toro M, Worobey, L. Wheelchair Skills Program Manual Version 5.1 Published electronically at Dalhousie University, Halifax, Nova Scotia, Canada. 2020. Available from: www.wheelchairskillsprogram.ca/eng/manual.php

43.Dalhousie University. Wheelchair Skills Program. Publications and impact. 2012 [Available from: https://wheelchairskillsprogram.ca/en/publications-impact/.

44.Rushton PW, Kirby RL, Miller WC. Manual wheelchair skills: objective testing versus subjective questionnaire. Arch Phys Med Rehabil. 2012;93(12):2313-8.

45.Toro ML, Bird E, Oyster M, Worobey L, Lain M, Bucior S, et al. Development of a wheelchair maintenance training programme and questionnaire for clinicians and wheelchair users. Disabil Rehabil Assist Technol. 2017;12(8):843-51.

46.Ardianuari S, Goldberg M, Pearlman J, Schmeler M. Development, Validation and Feasibility Study of a Remote Basic Skills Assessment for Wheelchair Service Providers. Disability and Rehabilitation: Assistive Technology.

47.DeVellis RFJMc. Classical test theory. 2006:S50-S9.

48.Hambleton RK, Jones RW. Comparison of classical test theory and item response theory and their applications to test development. Educational Measurement: Issues and Practice. 1993;12(3):38-47.

49.De Champlain AF. A primer on classical test theory and item response theory for assessments in medical education. Med Educ. 2010;44(1):109-17.

50.Crocker L, Algina J. Introduction to classical and modern test theory: ERIC; 1986.

51.Lord FM. The relation of the reliability of multiple-choice tests to the distribution of item difficulties. Psychometrika. 1952;17(2):181-94.

52.Kelley TL. The selection of upper and lower groups for the validation of test items. Journal of educational psychology. 1939;30(1):17.

53.Ebel RL. Measuring educational achievement: Englewood Cliffs; 1965.

54.Tan S. Misuses of KR-20 and Cronbach's alpha reliability coefficients. Egitim ve Bilim. 2009;34(152):101.

55.Miller LA, Lovler RL. Foundations of psychological testing: A practical approach: Sage Publications; 2018.

56.Kuder GF, Richardson MW. The theory of the estimation of test reliability. Psychometrika. 1937;2(3):151-60.

57.Kolen MJ, Brennan RL. Test equating, scaling, and linking: Methods and practices: Springer Science & Business Media; 2014.

58.Moody C, Obear K, Gasser H, Cheah S, Fechter T. ACPA standards for demographic questions. 2013.

59.Hughes JL, Camden AA, Yangchen T. Rethinking and updating demographic questions: Guidance to improve descriptions of research samples. Psi Chi Journal of Psychological Research. 2016;21(3):138-51.

60.Park CW, Gardner MP, Thukral VK. Self-perceived knowledge: Some effects on information processing for a choice task. The American Journal of Psychology. 1988:401-24.

61.Lievens F, Reeve CL, Heggestad ED. An examination of psychometric bias due to retesting on cognitive ability tests in selection settings. J Appl Psychol. 2007;92(6):1672-82.

62.Geiger MA. An examination of the relationship between answer changing, testwiseness, and examination performance. The Journal of experimental education. 1997;66(1):49-60.

63.World Health Organization. Community-based rehabilitation: CBR guidelines. 2010. Available from: https://www.who.int/publications/i/item/community-based-rehabilitation-cbr-guidelines.

64.Bandura A. Guide for Constructing Self-Efficacy Scales. Self-Efficacy Beliefs of Adolescents2006. p. 307-37.

65.Pajares F, Hartley J, Valinate G. Response format in writing self-efficacy assessment: Greater discrimination increases prediction. Measurement and evaluation in counseling and development. 2001;33(4):214-.

66.Dziuban CD, Shirkey EC. When is a correlation matrix appropriate for factor analysis? Some decision rules. Psychological bulletin. 1974;81(6):358.

67.Kaiser HF. The application of electronic computers to factor analysis. Educational and psychological measurement. 1960;20(1):141-51.

68.Cattell RB. The scree test for the number of factors. Multivariate behavioral research. 1966;1(2):245-76.

69.Ruscio J, Roche B. Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. Psychological assessment. 2012;24(2):282.

70.Ledesma RD, Valero-Mora P. Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. Practical assessment, research & evaluation. 2007;12(2):1-11.

71.Cortina JM. What is coefficient alpha? An examination of theory and applications. Journal of applied psychology. 1993;78(1):98.

72.Nunnally JC. Psychometric theory 3E: Tata McGraw-Hill Education; 1994.

73.Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med. 2018;18(3):91-3.

74.Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-63.

75.Kaiser HF. A second generation little jiffy. Psychometrika. 1970;35(4):401-15.

76.Tobias S, Carlson JE. Brief Report: Bartlett's Test of Sphericity and Chance Findings in Factor Analysis. Multivariate Behav Res. 1969;4(3):375-7.

77.Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb). 2015;25(2):141-51.

78.Rushton PW, Smith EM, Miller WC, Kirby RL, Daoust G. Reliability and responsiveness of the Self-Efficacy in Assessing, Training and Spotting wheelchair skills (SEATS) outcome measure. Disabil Rehabil Assist Technol. 2019;14(3):250-4.

79.Kirby RL, Rushton PW, Smith C, Routhier F, Best KL, Boyce J, et al. Wheelchair Skills Program Manual Version 5.0 Published electronically at Dalhousie University, Halifax, Nova Scotia, Canada.2019 [Available from: www.wheelchairskillsprogram.ca/eng/manual.php.

80.Berchtold A. Test–retest: agreement or reliability? Methodological Innovations. 2016;9:2059799116672875.

81.Keeler L, Kirby RL, Parker K, McLean KD, Hayden JA. Effectiveness of the Wheelchair Skills Training Program: a systematic review and meta-analysis(.). Disabil Rehabil Assist Technol. 2019;14(4):391-409.

82.Weiss B, Pollack AA. Barriers to global health development: An international quantitative survey. PloS one. 2017;12(10):e0184846.

83.Yamey G. What are the barriers to scaling up health interventions in low and middle income countries? A qualitative study of academic leaders in implementation science. Globalization and health. 2012;8(1):11.