Development, reliability, and piloting of a wheelchair caster failure inspection tool (C-FIT)

Anand A. Mhatre, Stephanie Lachell & Jonathan L. Pearlman

To cite this article: Anand A. Mhatre, Stephanie Lachell & Jonathan L. Pearlman (2019): Development, reliability, and piloting of a wheelchair caster failure inspection tool (C-FIT), Disability and Rehabilitation: Assistive Technology, DOI: 10.1080/17483107.2018.1554714

To link to this article: https://doi.org/10.1080/17483107.2018.1554714

© 2019 Informa UK Limited, trading as Taylor & Francis Group

Published online: 07 Feb 2019.

Submit your article to this journal

Article views: 67

View Crossmark data
Development, reliability, and piloting of a wheelchair caster failure inspection tool (C-FIT)

Anand A. Mhatre, Stephanie Lachell, and Jonathan L. Pearlman

ABSTRACT

Introduction: Wheelchair casters fail frequently in the field causing multiple user consequences and wheelchair breakdowns. To inform caster design improvement, there exists no validated tools that can collect caster failures. This need motivated the development of a user-reported, caster failure inspection tool (C-FIT).

Methods: To develop C-FIT, a multistep design and testing approach was used which included face validity testing, test-retest reliability testing and expert review. Reliability testing was conducted with two independent cohorts of wheelchair professionals who inspected caster failures physically and online through pictures. The tool was revised based on testing outcomes and expert feedback. For preliminary data collection and evaluating usability, C-FIT was piloted at wheelchair service centers in Scotland, Indonesia and Mexico.

Results: Caster failure items reported in the literature were screened to develop the initial list of C-FIT items. Face validity testing conducted through surveys with wheelchair experts (n = 6) provided 14 items for C-FIT inclusion. The test-retest reliability was found to be high for 10 items with physical failure inspections (n = 12). For each of these items, 75% or more participants had substantial to almost perfect agreement scores (κ = 0.6–1.0). Lower reliability scores were found with online failure inspections (n = 11). C-FIT received positive usability feedback from study participants and data collectors in the field. Pilot field data (n = 31) included comprehensive details about failures useful for manufacturers, designers and researchers to improve caster designs.

Conclusions: The C-FIT tool developed in this study has substantial reliability and can be used for documenting caster failures at wheelchair service centers.

IMPLICATIONS FOR REHABILITATION

- Collecting data on caster failures is an important first step to inform design improvements and caster quality testing methods.
- The caster failure inspection tool is a reliable tool that can be used during wheelchair repair and servicing to collect caster failures in a standardized way.
- The failure data can be used by wheelchair manufacturers, designers, technicians and researchers to develop reliable caster designs. Wheelchair providers can select caster designs based on context of use.

Introduction

The global unmet need for wheelchairs is around 75 million and several international organizations including the World Health Organization (WHO), United Nations (UN), United States Agency for International Development (USAID) and the International Society of Wheelchair Professionals (ISWP) are promoting improved access to wheelchair products [1–5]. While several advances have been made in addressing the need, the provision of high-quality products remains one of the most significant challenges [1,4,6–10]. Wheelchairs are known to experience frequent repairs and breakdowns. Product evaluation studies from Kenya, India and Mexico have reported breakdowns and multiple part failures with casters, brakes, seats and tires within 2–3 months to two years of wheelchair use [11–15]. In a series of cross-sectional survey studies conducted in the United States, about 44–64% of wheelchair users reported at least one breakdown within 6-months of wheelchair use and nearly one-third of them suffered immediate consequences such as being stranded or injured and missing school, work or appointments [16–19]. Repairs and breakdowns can go unaddressed which can make the loss of mobility long term, cause secondary complications and increase the likelihood of device abandonment [6,9,20–25].

Among different wheelchair part repairs reported in field evaluation studies, casters have been found to fail frequently with diverse failure modes [11–14,16,26]. Locked bearings, damaged bolts, wheels and forks, worn-out tires and missing fasteners are
common caster failures (Figure 1) [26]. One study documenting wheelchair incidents in the United States found that about one-third of part failures are caster failures. This study reported that tips and falls out of chairs are significantly associated with small-size casters having solid tires. [27]. In Scotland, among different part failures, caster failures were responsible for 27% of user consequences [28]. Another study reported that casters and wheels together contributed to 42% of the recorded wheelchair failures and 44% of adverse user consequences [16]. Additionally, casters are known to undergo a variety of failures during laboratory-based testing, especially during caster testing developed by the ISWP’s Standards Working Group (ISWP-SWG) [26,29–41].

The ISWP-SWG consists of wheelchair manufacturers, researchers and field experts. The group is developing quality testing standards for manufacturers and service providers that will improve reliability and usability of wheelchairs in less-resourced settings. Wheelchair casters are known to fail frequently based on evidence and hence, the caster testing protocol has been prioritized for development [11,12,14,26–28,42]. For improving the external validity of the protocol, ISWP-SWG is seeking to validate testing to field performance. For this purpose, caster failures between the two settings need to be compared. However, this is difficult as no failure data collection systems exist in the field and hence, the ISWP-SWG needs to constantly rely on anecdotal feedback from manufacturers.

Very few assessment tools are available for collecting data on wheelchair and caster repairs [42–46]. These tools are useful for rating overall caster condition or its maintenance state, but lack detailed failure mode inspection, which severely limits their use to field data collection. This prompted the development of a new tool that can be used by wheelchair technicians, designers, manufacturers, providers and testing agencies to collect and report caster failures. The purpose of this study was two-fold:

1. To develop a caster failure inspection tool (C-FIT) through an iterative design and testing approach that includes face validity and test-retest reliability testing.
2. To pilot the tool at wheelchair service centres and collect field data.

Methods and materials

A flowchart shown in Figure 2 describes the methodology used to develop and test C-FIT.

Initial C-FIT development

Caster failure modes found during wheelchair testing studies [29–41] and field trials [11,13,14,47–49] were compiled. They were manually screened for inclusion in C-FIT by a wheelchair testing engineer with 7 years of testing experience and a technician with over 20 years of experience in wheelchair maintenance.

Face validity testing

To establish face validity of C-FIT, an online Qualtrics survey [50] was distributed to ISWP-SWG experts. They were asked to vote for caster failure modes for inclusion as a tool item and rate the risk of user injury and other wheelchair part failures associated with each mode. Since no clinical or identifiable private information was collected in the survey, approval from the University of Pittsburgh’s Institutional Review Board (IRB) was not considered necessary. Based on the survey results, C-FIT was revised to improve face validity.

Test-retest reliability testing

A two cohort, repeated measures design was used to conduct the test-retest reliability testing. One cohort rated failures with the physical inspection of the casters and the other performed online inspections through photographs. A proposal for conducting this study was submitted to the IRB which determined that the study could proceed without approval as clinical data was not collected. A retest interval of two weeks was selected based on recall-related experiences from the test-retest study conducted by the research group earlier [43].

Convenience sampling was followed for recruiting participants in both cohorts. Individuals older than 18 years and having more than 1 year of field or research experience with wheelchairs were qualified to participate in the study. For the physical inspection group, individuals from university settings were approached in person or via email. The study purpose and procedures were communicated. Those who agreed to participate were provided participant codes through email. For the two study sessions, the participant was escorted to a quiet study room with a computer having a study survey on screen and a utility cart containing bagged casters with number tags placed in sequence [50]. A researcher working on this study accompanied the participant to answer and note questions during the study session.

For online participants, the survey link for the first survey was sent through a study invitation email to wheelchair experts in ISWP-SWG and technicians, wheelchair and assistive technology providers and clinicians. They were informed that completing the survey indicated their participation in the study. The second survey link was sent two weeks following the completion date of the first survey.

Caster samples

Twenty-eight casters were inspected by each study participant in a randomized order. Examples of samples used in the study are shown in Figure 3. Each failure item was represented at least two times among the 28 samples. The bent fork failure was an exception; only one sample had the failure. Half of the casters were field failures and the other half were failures from laboratory-based testing [26,51]. Casters inspected in the study had significant variability in their designs and failure modes.
To assess the reliability of 14 items, at least 8 participant raters are required in each group to be 95% certain that the reliability is >= 0.8 ± 0.1. Assuming dropout of 20%, a sample size of 10 individuals was required in both cohorts.

**Description of Failures**

Caster part | Failure Mode |
--- | --- |
Axle bearing | 1. Corrosion*  
2. Obstruction to rolling*  
3. Fracture*  
4. Play between bearings and axle bolt  
5. Loss of contact with wheel  
6. Loss of trueness of the bearing  
7. Dirt contamination |
Axle bolt | 1. Fracture |
Caster Wheel | 1. Fracture*  
2. Corrosion* |
Tire | 1. Tire worn-out*  
2. Tread worn-out*  
3. Cracking*  
4. Deflated*  
5. Tire etching on sides  
6. Pitting |
Stem Bearing | 1. Corrosion*  
2. Play between bearings and stem bolt  
3. Dirt contamination |
Stem Bolt | 1. Fracture* |
Fork | 1. Bent*  
2. Fracture*  
3. Corrosion* |
Fasteners | 1. Corrosion  
2. Loose fasteners |

*Failure selected for C-FIT inclusion.

**Power analysis for test-retest reliability testing**

Participant sample size was calculated using the procedure for the standard error of the reliability coefficient [52]. To assess the reliability of 14 items, at least 8 participant raters are required in each group to be 95% certain that the reliability is >0.8 ± 0.1. Assuming dropout of 20%, a sample size of 10 individuals was required in both cohorts.

**Survey design**

All surveys collected demographic data, introduced the participant to the study procedures, provided standardized nomenclature for caster components, listed possible failure modes and provided inspection instructions with photographic illustrations. A sample C-FIT with scoring options was included as well to familiarize the participant with the structure and content. The failure modes in C-FIT were hyperlinks to instructions for caster failure inspections. Failures were to be rated as present, not present, or unable to evaluate. A separate comment box was provided at the end of each caster assessment to note down any issues encountered during the inspection. For the online group, photographs of failed casters to be rated were embedded in the survey questions. The retest session surveys requested participants to rate items related to the usability. All items were scored individually on a 5-point Likert scale in which 1-do not agree and 5-fully agree. Items related to the use of C-FIT in the field were not included in retest surveys for the physical inspection group as all of them were laboratory researchers by profession. Feedback on improving the tool and suggestions for additional failure modes for inclusion in C-FIT were requested.

**Data analysis for test-retest reliability testing**

Descriptive statistics were calculated for the participant’s demographic information. Three response choices were available for rating a failure item – (1) Failure present, (2) Failure not present and (3) Unable to evaluate which were scored as 1, 2 and 3 respectively for data analysis. Missing responses were scored as 0. Test-retest reliability was calculated using Cohen’s Kappa [53] and percentage agreement. Fleiss’ kappa was used for inter-rater reliability estimate as raters were greater than two [54]. Using the algorithm of Landis and Koch [55], kappa values of 0.81 and above represented almost perfect agreement, values between 0.61 and 0.80 represented substantial agreement, 0.41–0.60 represented moderate agreement, 0.21–0.4 is fair agreement and values below 0.20 suggested slight to poor agreement. For evaluating the accuracy of the responses for each failure mode, participant responses were compared with caster failure ratings (true scores) scored by the primary author who has experience in wheelchair and caster testing. Data analysis was conducted using the statistical package IBM SPSS 24 [56].

**C-FIT revision**

Participant ratings and comments for each caster sample, and overall feedback on C-FIT were analyzed to evaluate participant performance and C-FIT’s reliability and usability. This analysis
informed the revisions to the C-FIT items and instructional materials.

**Expert review & Translation**

Following revision, another round of review from a wheelchair testing engineer, technician, clinician, an assistive technology provider and two wheelchair manufacturers in ISWP-SWG was carried out. Feedback on improving the usability of the tool and instructional materials was requested. Revisions were made based on the feedback. C-FIT was translated to Spanish. Two rehabilitation engineers who are native Spanish speakers reviewed and revised the translation.

**Field data collection**

The revised tool was disseminated to ISWP partners in Indonesia, Scotland and Mexico for pilot data collection through physical failure inspections. Data collectors were requested for feedback on the usability of C-FIT.

**Results**

**Initial C-FIT development results**

Fourteen failure modes were screened for initial inclusion in C-FIT (Table 1). Failures typically seen in the field and during testing were selected by the experts.

**Face validity testing results**

Table 2 shows the percentage scores received by each C-FIT item for inclusion. Failures of caster wheel corrosion and tire tread worn-out scored less than or equal to 60% for inclusion and were rated as having little to no risk of user injury or other wheelchair damage. Therefore, these two items were removed. Experts recommended four failure modes for inclusion – (1) Stem Bearing Fracture, (2) Tire Roll-off, (3) Stem and axle bolt not set to specified torque and (4) Caster Shimmy. The first two items were included and the latter two were left out because their inspections are complex and subjective.

**Reliability testing results**

Demographic characteristics of the participants are shown in Table 3. Table 4 shows the sample C-FIT from the surveys. The range of test-retest reliability scores and percentage of participants falling within each agreement interval are shown in Tables 5 and 6 for the two cohorts. Tables 7 and 8 show the inter-rater reliability and accuracy scores. Table 9 shows participants that agreed (Likert scale score > 3) with statements related to the C-FIT’s usability. Two participants in the physical inspection group commented that C-FIT is intuitive and easy to follow. One participant noted that the tool is useful with correct training experience.

**Participant feedback and revision**

Participants provided constructive feedback for C-FIT’s improvement. Their feedback coupled with the analysis of individual
rings, provided directions for tool revisions (Table 10). Failure modes of axle bearing contamination, bent stem bolt, locked stem bearings, tire tread loss and loose fork were added based on suggestions and failures evident with study samples. The tread loss failure eliminated after face validity review was added back as experts in resourced settings considered it to be a common failure with risk of consequences on smooth inclined floors.

**Expert feedback and revision**

Experts provided valuable inputs to improve the usability of the tool. They recommended restructuring the C-FIT for inspecting caster parts in a sequence which can streamline the data collection and consequently reduce the inspection and data collection time. Revisions were done based on suggestions and C-FIT was published online in both English and Spanish [57,58]. A website including instructional materials in both English and Spanish was also published through the WordPress platform [59]. These materials and the tool can be accessed through digital devices using Android and iOS platforms.

**Field data collection**

Three data collectors at repair facilities in Indonesia, Scotland and Mexico submitted failure data for a total of 31 casters (Figures 4–6) from 13 wheelchair models. Casters from Indonesia and Mexico had greater wear and tear. On the other hand, failures from Scotland were fracture and deformation failures. For each caster, along with failure ratings, comments were provided which included the causes of failures, user characteristics and design suggestions to prevent field failures. With frequent use of C-FIT, users were able to complete caster inspections within 5 min. Data collectors from Scotland and Indonesia noted that C-FIT was easy to use during wheelchair repairs and they would like to continue using it.

**Discussion**

Wheelchair casters fail frequently in the community and field trials conducted around the world report the high frequency and wide variety of caster failures seen with caster models [12–15,17].
Unfortunately, the available research evidence does not provide detailed information to inform caster design improvements and testing protocol developments. Furthermore, there is a lack of standard tools for data collection on caster failures. This need motivated the development of the C-FIT tool and this study describes the iterative approach used to develop and validate the tool. Data collected from face validity and test-retest reliability testing studies and reviews from wheelchair experts provided at various stages of the study were used to improve the usability and reliability of C-FIT.

Reliability and accuracy scores for the C-FIT items were favorable in the physical inspection group. More than 75% of our participants had substantial to almost perfect test-retest reliability for 10 failure modes. Additionally, 50–66.67% participants had substantial to almost perfect reliability for the other four failures of axle bearing fracture, axle bearing’s obstruction to rolling, tire cracking and stem bearing fracture. Nine failure modes had substantial to almost perfect inter-rater reliability scores. Accuracy was found to be greater than 75% except for two of the axle bearing failures. These results indicate that the C-FIT is substantially reliable for collecting data on failures when casters are inspected physically.

On the other hand, the online inspection group received lower reliability scores. Only 2 out of 14 failures had substantial to almost perfect test-retest and inter-rater reliability scores. The online participants commented that they were restricted to the visual analysis of failures. Most failures needed physical inspections which led to inconsistent ratings. Because of this, improvements to online caster inspection materials were not pursued further.

Feedback regarding the usability of C-FIT was positive from both study cohorts. Participants appreciated the training provided prior to inspections. Seventy-five percent of the participants from the physical group rated that C-FIT was easy to use for rating caster failures. The online participants commented that they were restricted to the visual analysis of failures. Most failures needed physical inspections which led to inconsistent ratings. Because of this, improvements to online caster inspection materials were not pursued further.

### Table 9. Feedback by study participants regarding usability.

<table>
<thead>
<tr>
<th>Feedback</th>
<th>Physical inspection group (n = 12)</th>
<th>Online inspection group (n = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The caster anatomy information was redundant in this survey</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>The instructions for evaluating casters were helpful</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>I need more training materials before using C-FIT</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C-FIT is easy to use for rating caster failures</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Evaluating the casters through photos was difficult</td>
<td>NA</td>
<td>5</td>
</tr>
<tr>
<td>I would like to use C-FIT for collecting failure data on casters</td>
<td>NA</td>
<td>4</td>
</tr>
<tr>
<td>I prefer using C-FIT online through a laptop, phone or tablet for collecting failures</td>
<td>NA</td>
<td>7</td>
</tr>
<tr>
<td>I prefer using a paper version of C-FIT for collecting failures</td>
<td>NA</td>
<td>3</td>
</tr>
</tbody>
</table>
### Table 10. Individual feedback and related revisions to C-FIT.

<table>
<thead>
<tr>
<th>No.</th>
<th>Inspection issues</th>
<th>Revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Participants were not sure which option to choose when parts were missing.</td>
<td>Missing parts was added as an option.</td>
</tr>
<tr>
<td>2.</td>
<td>In many cases, bearings were not visible and so, the participants guessed the failure based on condition of the caster. Broken bearing seals were not scored as failures.</td>
<td>Instructions were updated to evaluate bearings and bearing seals in detail and it was noted that participant should check the 'Unable to evaluate' option if the detailed inspection is not carried out. The instructional materials note that the participant should go through the informational materials prior to caster inspections and check each part.</td>
</tr>
<tr>
<td>3.</td>
<td>One participant considered loose washers are part of the bearings and scored it as a fracture failure. In some cases, failed bearings were left in the bags and not evaluated at all.</td>
<td>Instructions for tire failures were described in detail to explain differences in tire worn-out, tread worn-out and cracking. Instructions that dirt on tire do not account for wear were added. Instructions were added for checking the pneumatic tire by looking for a valve prior to rating the failure.</td>
</tr>
<tr>
<td>4.</td>
<td>Failures of tire worn-out and cracking were confusing to participants. Some participants rated tread wear as tire worn-out.</td>
<td>Corrosion was divided into mild and high corrosion as the failure modes. Both failure modes were distinguished based on the outcomes of the failure inspection. The obstruction to rolling failure was changed to sticky bearing failure for axle bearings.</td>
</tr>
<tr>
<td>5.</td>
<td>Any dirt on the tire was rated as tire wear</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>When inspecting for tire deflated failure, several participants did not check if the tire was pneumatic.</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Corrosion and obstruction to rolling failures were found to be subjective. Participants were ambiguous about what degree of corrosion and obstruction should be rated as failure.</td>
<td></td>
</tr>
</tbody>
</table>

![Field failures collected using C-FIT](image-url)
preferred platform for data collection compared to manual entries on paper. These results show that the tool can be integrated into practice and prove valuable for community data collection.

Assessment of individual ratings and caster comments highlighted the reasons for certain failure modes related to bearings, tires and corrosion for having low-moderate reliability and accuracy scores. Overall, there were discrepancies with C-FIT scoring options, participant inspections and training materials which led to inconsistent responses over the two study sessions. These findings and the series of feedback from participants and experts were valuable as they provided directions for improving the reliability and usability of C-FIT.

Following revisions, C-FIT was distributed to ISWP partners around the world and three of them volunteered for trialing it during wheelchair repairs. Preliminary field data indicates that C-FIT users were able to report comprehensive information on caster failures. Many casters failures in less-resourced settings may be due to the adverse environmental conditions and rough terrains witnessed there. Tire, bearing and some fracture failures are notable with certain models which may not be appropriate for less-resourced settings. Failures from Scotland were due to shocks and impacts which can be prevented with proper user training and maintenance. C-FIT is usable in wheelchair service centers based on the favorable feedback received from data collectors. Further data collection on the reported caster models is anticipated from multiple locations around the world. It can help in characterizing caster failures better, making reliable comparison with the laboratory testing failures and guide caster design improvements to increase robustness and reliability.

**Limitations**

Experts and participants involved in the study were a convenience sample and not randomly selected. Participants from the physical inspection group were engineers and researchers from a university setting and may not be representative of the data collectors in the field although following graduation, are often being employed by wheelchair manufacturers and suppliers. There may have been some learning effect which is typical of test-retest studies. Fourteen C-FIT items were tested as a part of the test-retest reliability study but the revised tool (includes 6 additional failure items) was not tested prior to piloting.

**Future work**

A database of field and laboratory failures collected using C-FIT will be developed for failure comparisons of caster models. The database shall allow technicians or manufacturers to view the failure data submitted by them in tabular and graphical summary format. To improve field usability and acceptance of C-FIT, the authors plan to conduct a field usability study with a suitable sample size of caster failure data collectors.

**Conclusions**

Wheelchair casters are a common point of failures, and failure evidence is necessary to drive design improvements and testing protocol development. Lack of validated tools to comprehensively collect caster failures motivated the development of a reliable caster failure inspection tool. This study developed the C-FIT tool through an iterative design and testing approach. Test-retest reliability testing of the tool demonstrated that rating failures with physical caster inspection is a reliable method. Field data collected using C-FIT provides valuable information for developing suitable caster quality testing protocols and improving caster designs.

**Acknowledgements**

The authors would like to thank the members of International Society of Wheelchair Professionals – Standards Working Group, wheelchair experts and reliability testing study participants for their time and participation in this study. The authors acknowledge the contributions by Karen Rispin, Stephanie Vasquez, Mendel Marcus and Mauricio Arredondo to this study.
Disclosure statement
The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Funding
This project was supported by the United States Agency for International Development (USAID) (Subaward No. APC-GM-0068 under USAID Cooperative Agreement No. AID-OAA-A-12–00047).

ORCID
Anand A. Mhatre http://orcid.org/0000-0001-6179-3727
Jonathan L. Pearlman http://orcid.org/0000-0003-0830-9136

References


