

Mobile Toolbox Data Rescoring Guide

Last Updated: May 8, 2025

Table of Contents

Purpose	3
Core Cognitive Measures	3
Overview	3
Arranging Pictures	4
Fixed in MyCap Version 2.1.6	4
Error: Incorrectly Counted Live Steps; Change to itemCount Logic	4
Arrow Matching	5
Fixed in MyCap Version 2.1.6	5
Error: Incorrectly Counted Live Steps	5
Error: Missing Items from Set B Scoring	6
Error: Incorrect Time Units	7
Faces and Names Learning	8
There are no recorded issues with this task.	8
Faces and Names Test	9
Fixed in MyCap Version 2.1.6	9
Error: Incorrectly Counted Live Steps	9
Number Symbol Match	10
Fixed in MyCap Version 2.1.6	10
Error: Incorrectly Counted Live Steps	10
Sequences	11
Fixed in MyCap Version 2.1.6	11
Error: Incorrectly Counted Live Steps	11
Shape Color Sorting	12
Fixed in MyCap Version 1.9.4	12
Error: Test Duration Differing Between Devices	12
Fixed in MyCap Version 2.1.6	12
Error: Incorrectly Counted Live Steps	12
Error: Incorrect Time Units	13
Spelling	15
Fixed in MyCap Version 2.1.6	15
Error: Incorrectly Counted Live Steps	15

Word Meaning.....	16
Fixed in MyCap Version 2.1.6	16
Error: Incorrectly Counted Live Steps.....	16

Purpose

The purpose of this document is to systematically record any measurement-related fixes that have been identified and may impact data analysis and the calculation of scores. It serves as a comprehensive rescoring guide and tracking tool, capturing crucial information such as the nature of the error, its severity, how the error may influence the data produced, and any necessary steps to manual data manipulation that need to be considered for data analysis. Additionally, this document functions as a historical record, illustrating how measures and resolutions have evolved over time. By maintaining an accurate log of these developments, we can better understand the trajectory of our quality assurance efforts and improve our processes moving forward.

Core Cognitive Measures

Overview

This documentation outlines a series of bugs identified during the integration of Mobile Toolbox measures with MyCap (MTB2). Each entry provides insight into the nature of the errors, their impact on measure functionality, and any relevant resolutions or workarounds. This comprehensive overview aims to facilitate better understanding and management of integration challenges over the specified period, contributing to ongoing improvements in the measure's performance and reliability.

Arranging Pictures

Fixed in MyCap Version 2.1.6

Error: Incorrectly Counted Live Steps; Change to itemCount Logic

Summary: The item counting logic for Arranging Pictures changed since its first release, and live steps were incorrectly counted, resulting in the itemCount field potentially capturing both practice and live items. Additionally, this resulted in several nil scores relating to when the step times out, the measure times out, the step was skipped, or the measure was skipped. Thus, the itemCount and adjacentPairs score fields require rescoring if any time outs occurred.

Duration of the Bug: Data between MyCap Version 1.9.4 and MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- arp_item_count
- arp_adjacentpairsscore

Full Description: The algorithm for calculating itemCount changed since its first release. Instead of counting a parent step if any of its child steps have a response, only parent steps whose child steps all have responses are counted. Thus, itemCount was calculated incorrectly.

Additionally, this measure inferred which steps were live based on whether they were explicitly designated as (1) “practice=false” or (2) if the score value was nil. For (2), multiple conditions could result in nil scores, such as if a step times out, if the measure times out, if a step is skipped, or if the measure is skipped. Due to this, the calculation of the itemCount and the adjacentPairs score fields may be impacted.

Recommendations for Data Analysis: Researchers will need to export the participant JSON file from REDCap and manually recalculate scores by filtering data for only live steps.

Arrow Matching

Fixed in MyCap Version 2.1.6

Error: Incorrectly Counted Live Steps

Summary: Live steps were incorrectly counted. Moreover, this resulted in several nil scores relating to when the step times out, the measure times out, the step was skipped, or the measure was skipped. Thus, reaction time-dependent scores for Arrow Matching did not include items where the participant timed out. Other score fields may be affected as a result of

Duration of the Bug: Data between MyCap Version 1.9.4 and MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- arw_item_count
- arw_totalerrors
- arw_rawscore
- arw_totalerrors
- arw_ratescore
- arw_mediancorrectrt
- arw_medianincorrectrt

Full Description: The itemCount was incorrectly capturing the step type description that denominates which steps are live test items. This meant that the item count was calculated incorrectly and potentially counted both practice and live items. As a result, this may affect calculation of the rawScore, totalErrors, rateScore, and medianIncorrectRT scores.

Additionally, because of this error, scores for items that were timed out were recorded as nil rather than 0. For the rate score (arw_ratescore), this resulted in timed out items to be excluded from the denominator. For reaction time specific scores (arw_medianincorrectrt), response time was excluded from the calculation of the median reaction time score.

Field	Android value	iOS value	Intended value
response	nil	nil	nil
score	nil	0	0
responseTime	2000	2000	2000
rawScore	No effect	No effect	No effect / effect = 0
rateScore	Excluded from denominator	Included in denominator	Included in denominator

medianincorrectRT	Excluded	Included	Included
-------------------	----------	----------	----------

Recommendations for Data Analysis: Researchers will need to export the participant JSON file from REDCap and manually extract these scores. Because the calculation of arw_totalerrors is dependent on arw_item_count, this score will need to be manually recalculated using the following formula: (itemCount– rawScore) + nAnticipationLive

Researchers should work from the item-level data from the taskdata.JSON to make sure that scores are computed correctly. To do this, replace nil (or NA) item-level scores with 0 and replace nil (or NA) item-level reaction time with the max time of 2000ms.

For the rate score (arw_ratescore), sum item-level scores, sum the item-level reaction time (in ms), and divide the item-level reaction time by 60 to convert it to the unit of seconds. Then divide the summed item-level scores by the summed reaction time (in sec). For the median incorrect reaction time score (arw_medianincorrectrt), find the median reaction time of the imputed item-level reaction time data.

Error: Missing Items from Set B Scoring

Summary: Seven of the step identifiers in the “B” set of this task were repeated after step identifiers were renamed, resulting in only 43 scored steps instead of 50. Thus, an incorrect score calculation is provided for all runs of this set, which impacts users who complete this measure more than once on the same device.

Duration of the Bug: Data between MyCap Version 1.9.4 and MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- arw_rawscore
- arw_ratescore
- arw_medianincorrectrt
- arw_mediancorrectrt
- arw_totalerrors (all for set B only)

Full Description: Seven of the identifiers in the “B” set of this task were repeated after step identifiers were renamed in MyCap Version 1.9.4., i.e., duplicated step identifiers were omitted, resulting in only 43 scored steps instead of 50. Specifically, the following items were missing from the score calculations: ARW_B40, ARW_B41, ARW_B42, ARW_B43, ARW_B44,

ARW_B45, and ARW_B46. In their place, these were repeated: ARW_B30, ARW_B31, ARW_B32, ARW_B33, ARW_B34, ARW_B35, and ARW_B36. When calculating scores, items with duplicated item name identifiers are omitted. Therefore, every score would be calculated from only 43 items. Thus, an incorrect score calculation is provided for all runs of this set, which impacts users who complete this measure more than once on the same device.

Recommendations for Data Analysis: The raw stepHistory data within the taskdata.JSON includes the correct response and score for every item. For participants who took the “B set” of Arrow Matching between MyCap Version 1.9.4 and MyCap Version 2.1.6, researchers will need to export the participant JSON file from REDCap and manually extract these scores using the item-level data.

Error: Incorrect Time Units

Summary: Two of the reaction time dependent scores for Arrow Matching were being calculated in the wrong unit of time (seconds instead of milliseconds), thus resulting in participants appearing to have a score of 0.

Duration of the Bug: Data collected in versions earlier than MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- arw_medianincorrectrt
- arw_mediancorrectrt
- arw_ratescore

Full Description: The incorrect unit of time (seconds instead of milliseconds) were being used to calculate the median incorrect reaction time score (arw_medianincorrectrt) and median correct reaction time score (arw_mediancorrectrt).

Recommendations for Data Analysis: For participants who already took this test before MyCap Version 2.1.6, researchers will need to export the participant taskdata.JSON file from REDCap and manually extract these scores. Specifically, in the taskdata.JSON the scores are reported in seconds and can be converted to milliseconds to get the correct unit of time.

Faces and Names Learning

There are no recorded issues with this task.

Faces and Names Test

Fixed in MyCap Version 2.1.6

Error: Incorrectly Counted Live Steps

Summary: Live steps were incorrectly counted. Additionally, this resulted in several nil scores relating to when the step times out, the measure times out, the step was skipped, or the measure was skipped. Thus, the itemCount score field may require rescoring.

Duration of the Bug: Data between MyCap Version 1.9.4 and MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- fnt_item_count

Full Description: The itemCount was incorrectly capturing the step type description that denominates which steps are live test items. This measure inferred which steps were live based on whether they were explicitly designated as (1) “practice=false” or (2) if the score value was nil. For (2), multiple conditions could result in nil scores, such as if a step times out, if the measure times out, if a step is skipped, or if the measure is skipped. Due to this, the calculation of the itemCount score field may be impacted under the above conditions.

Recommendations for Data Analysis: Researchers will need to export the participant JSON file from REDCap and manually extract these scores.

Number Symbol Match

Fixed in MyCap Version 2.1.6

Error: Incorrectly Counted Live Steps

Summary: Live steps were incorrectly counted. In addition to live items, practice items may be included in the final item count. Moreover, this resulted in several nil scores relating to when the step times out, the measure times out, the step was skipped, or the measure was skipped. Thus, the itemCount score field may require rescoring.

Duration of the Bug: Data between MyCap Version 1.9.4 and MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- nsm_item_count

Full Description: The itemCount was incorrectly capturing the step type description that denominates which steps are live test items. This measure inferred which steps were live based on whether they were explicitly designated as (1) “practice=false” or (2) if the score value was nil. For (2), multiple conditions could result in nil scores, such as if a step times out, if the measure times out, if a step is skipped, or if the measure is skipped. Due to this, the calculation of the itemCount score field may be impacted under the above conditions.

Recommendations for Data Analysis: Researchers will need to export the participant JSON file from REDCap and manually extract these scores.

Sequences

Fixed in MyCap Version 2.1.6

Error: Incorrectly Counted Live Steps

Summary: Live steps were incorrectly counted. In addition to live items, practice items may be included in the final item count. Moreover, this resulted in several nil scores relating to when the step or the measure was skipped. Thus, the itemCount score field may require rescoring.

Duration of the Bug: Data between MyCap Version 1.9.4 and MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- mfs_item_count

Full Description: The itemCount was incorrectly capturing the step type description that denominates which steps are live test items. This measure inferred which steps were live based on whether they were explicitly designated as (1) “practice=false” or (2) if the score value was nil. For (2), multiple conditions could result in nil scores, such as if a step or the measure is skipped. Due to this, the calculation of the itemCount score field may be impacted under the above conditions.

Recommendations for Data Analysis: Researchers will need to export the participant JSON file from REDCap and manually extract these scores.

Shape Color Sorting

Fixed in MyCap Version 1.9.4

Error: Test Duration Differing Between Devices

Summary: Android users see the blank screen and the word “Shape” or “Color” longer than iOS users. However, the duration allowed for participant to record their response remains the same between operating systems.

Duration of the Bug: Data collected before MyCap Version 1.9.4 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android

Variable / Field Names (system name):

- scs_rawscore
- scs_ratescore

Full Description: Android users see the blank screen (Part A) for 2.1 seconds instead of 1.0 seconds. Likewise, the word “Shape” or “Color” (Part B) is presented for 2.0 seconds instead of 0.8 seconds. The timing on the part where a participant records their response is correct.

Recommendations for Data Analysis: Comparisons should not be made between devices for data collected before the official release of the MTB. A new version of the task has been made available in MyCap that should be used when starting new data collection.

Fixed in MyCap Version 2.1.6

Error: Incorrectly Counted Live Steps

Summary: Live steps were incorrectly counted. In addition to live items, practice items may be included in the final item count. Moreover, this resulted in several nil scores relating to when the step times out, the measure times out, the step was skipped, or the measure was skipped. Thus, the itemCount score field may require rescoring.

Duration of the Bug: Data between MyCap Version 1.9.4 and MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- scs_item_count
- scs_totalerrors
- scs_ratescore
- scs_medianincorrectrt
- scs_itemcount

Full Description: The itemCount was incorrectly capturing the step type description that denominates which steps are live test items. This measure inferred which steps were live based on whether they were explicitly designated as (1) “practice=false” or (2) if the score value was nil. For (2), multiple conditions could result in nil scores, such as if a step times out, if the measure times out, if a step is skipped, or if the measure is skipped. Due to this, the calculation of several score fields may be impacted under the above conditions.

Scores for items that were timed out were recorded as nil rather than 0. For the rate score (scs_ratescore), this resulted in timed out items to be excluded from the denominator. For reaction time specific scores (scs_medianincorrectrt), response time was excluded from the calculation of the median reaction time score.

Field	Android value	iOS value	Intended value
response	nil	nil	nil
score	nil	0	0
responseTime	2000	2000	2000
rawScore	No effect	No effect	No effect / effect = 0
rateScore	Excluded from denominator	Included in denominator	Included in denominator
medianincorrectRT	Excluded	included	included

Recommendations for Data Analysis: Researchers will need to export the participant JSON file from REDCap and manually extract these scores. Because the calculation of scs_totalerrors is dependent on scs_item_count, this score will need to be manually recalculated using the following formula: (itemcount– rawScore) + nAnticipationLive

Error: Incorrect Time Units

Summary: Two of the reaction time dependent scores for Shape-Color Sorting were being calculated in the wrong unit of time (seconds instead of milliseconds), thus resulting in participants appearing to have a score of 0.

Duration of the Bug: Data collected in versions earlier than MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- `scs_medianincorrectrt`
- `scs_mediancorrectrt`
- `scs_ratescore`

Full Description: The incorrect unit of time (seconds instead of milliseconds) were being used to calculate the median incorrect reaction time score (`scs_medianincorrectrt`) and median correct reaction time score (`scs_mediancorrectrt`).

Recommendations for Data Analysis: Researchers will need to export the participant JSON file from REDCap and re-export these scores. Researchers should work from the item-level data to make sure that scores are computed correctly. To do this, replace nil (or NA) item-level scores with 0 (`item_level_scores[is.na(item_level_scores)] <- 0`) and replace nil (or NA) item-level reaction time with the max time of 2000ms (`item_level_time[is.na(item_level_time)] <- 2000`).

For the rate score (`scs_ratescore`), sum item-level scores, sum the item-level reaction time (in ms), and divide the item-level reaction time by 60 to convert it to the unit of seconds. Then divide the summed item-level scores by the summed reaction time (in seconds). For the median incorrect reaction time score (`scs_medianincorrectrt`), find the median reaction time of the imputed item-level reaction time data.

Spelling

Fixed in MyCap Version 2.1.6

Error: Incorrectly Counted Live Steps

Summary: Live steps were incorrectly counted. In addition to live items, practice items may be included in the final item count. Moreover, this resulted in several nil scores relating to when the step or the measure was skipped. Thus, the itemCount score field may require rescoring.

Duration of the Bug: Data between MyCap Version 1.9.4 and MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- spl_item_count

Full Description: The itemCount was incorrectly capturing the step type description that denominates which steps are live test items. This measure inferred which steps were live based on whether they were explicitly designated as (1) “practice=false” or (2) if the score value was nil. For (2), multiple conditions could result in nil scores, such as if a step or the measure is skipped. Due to this, the calculation of the itemCount score field may be impacted under the above conditions.

Recommendations for Data Analysis: Researchers will need to export the participant JSON file from REDCap and manually extract these scores.

Word Meaning

Fixed in MyCap Version 2.1.6

Error: Incorrectly Counted Live Steps

Summary: Live steps were incorrectly counted. In addition to live items, practice items may be included in the final item count. Moreover, this resulted in several nil scores relating to when the step or the measure was skipped. Thus, the itemCount score field may require rescoring.

Duration of the Bug: Data between MyCap Version 1.9.4 and MyCap Version 2.1.6 are impacted by this error.

Version of Measure Impacted: English

OS Impacted: Android & iOS

Variable / Field Names Impacted:

- wm1_item_count
- wm2_item_count

Full Description: The itemCount was incorrectly capturing the step type description that denominates which steps are live test items. This measure inferred which steps were live based on whether they were explicitly designated as (1) “practice=false” or (2) if the score value was nil. For (2), multiple conditions could result in nil scores, such as if a step or the measure is skipped. Due to this, the calculation of the itemCount score field may be impacted under the above conditions.

Recommendations for Data Analysis: Researchers will need to export the participant JSON file from REDCap and manually extract these scores.