# Validation Scoring

*Kono Team*

*July 12, 2018*

## Contents

**Introduction**

The Validation exercise was aimed at assessing how well the CAHWs were performing their duties of animal health surveillance in the communities. It tested their level of technical expertise and effectiveness as a CAHW. Each CAHW was assigned a final score or grade at the end of the exercise.

The validation exercise consisted of two components: 1) Field Practical and 2) Field visits. The Field Practical evaluated technical expertise of CAHWs and their ability to fill up the CAHW forms accurately. Here, CAHWs had to undergo two main tasks: a) An Animal Station exam where CAHWs observed sick animals and b) Form Review process that checked their understanding of the CAHW forms.

The Field Visits captured how effectively CAHWs were surveilling sick animals in the communities. CAHWs were assessed on the basis of: c) Sick Animal Observation where facilitators verified if sick animals were reported by the CAHW and d) Monitoring Form Follow-Up, where select owners of sick animals were interviewed.

In the sections below, we go over each component, explain how the process took place, aggregate the data and break down the scoring system calculated on unweighted-sum basis.

# Field Practical

This was the first exercise of Validation which took place in February 2018. Teams of facilitators and RAs gathered in four different hubs to examine CAHWs' ability to identify symptoms and correctly report them (Animal Station) and CAHWs' understanding of filling up forms (Form Review).

CAHWs were distributed to different hubs, with each being invited to the closest hub near their communities. In these hubs, several stations were set-up with sick animals for them to observe and report. A team of professional vets from MAFFS diagnosed the animal's symptoms in advance, so that the CAHWs' responses can be checked. CAHWs went through these stations, undertook the medicine exams, sat down with RAs to review their previously submitted forms from their communities.

A total of 267 (out of 287) CAHWs participated in this exercise. Those who did not attend get the lowest possible score. From those 267 CAHWs that participated in the field practical 8 had not sent any form by the time of the evaluation and thus RAs didn't have any form to review. These 8 CAWHs get the same score that they obtained in the Animal Station exercise.

## Animal Stations

After observing sick animals at the animal stations, each CAHW had to fill out a CAHW form. These forms were then scored based on the correct responses provided by the vet.

Animal Station scoring is divided into four subsections as listed below. They were allotted points on the basis of the following:

1. Identifiers: Points given for each correct field in the top of the form. These fields allow us to identify CAHWs, community members and their place of residency.
2. Cases: Points given/substracted for fields correctly/wrongly marked or left blank. This subcomponent evaluates the CAHWs' ability to accuretly report the number and type of animals at a given station.
3. Symptoms: Points are given for correctly identifing symptoms and for not misreporting. Misreporting mistakes are classified from low to high based on the how critical the misreported symptom is.
4. Additional Information: Points related to *other_events* and *sub_info_points*. Points giving for correctly filling out the bottom of the form. The bottom of the form captures information related to animals that are not included in the form.

After computing points for each subsection, we normalize scores by dividing over the total possible scores that could have potentially obtained at a given station. Next, we aggreagate scores by community and standardize them. Lastly, we compute final scores as the unweighted sum of each of the subcomponents (see Section Consolidating Animal Station Scores).

### Identifiers (AS)

In this subsection CAHWs are rewarded (points shown below) for every field correctly filled on each of the forms at their designated AS.

Table 1: For 'name_wrong' and 'comm_wrong', columns 'name' and 'locattioncomm' are compared against 'Answer keys' using fuzzy join. Variables matching at a higher distance score than 0.75 are considered correct and thus not flagged. Top of the form variables are code as 1 whenever wrong

| Variables | Description | Scoring |
|---|---|---|
| mult_form | One form submitted? | 10 |
| name_wrong | CAHW name correct | 3 |
| comm_wrong | CAHW community correct | 3 |
| w_dateday | Day correct | 1 |
| w_datemonth | Month correct | 1 |
| w_dateyear | Year correct | 1 |
| w_locationdistrict | Location correct | 3 |
| w_locationchiefdom | Chiefdom correct | 3 |
| w_own_comm | Own community | 5 |
| w_id | CAHW id correct | 10 |
| w_type | observed/reported? | 5 |

CAWHs get zero points if a field was wrongly filled. Variables from the top of the form, or *identifiers* are coded as 1 when an error was made in a given field.

### Cases (AS)

In this subsection CAHWs recieve points (see Table below) for columns accurately left blank and for columns correctly marked.

| Top Variables | Description | Scoring |
|---|---|---|
| w_blank | Incorrect blank fields | -1 |
| c_blank | Correct blank fields | 3 |
| c_mark | Correct marked fields | 10 |
| w_animal | Correct animal type | 5 |

### Symptoms (AS)

CAHWs are rewarded for reporting symptoms that should have been reported, and for not reporting symptomss that should not have. Additionally, CAHWs get points when they informed about plausible symptoms that could have been observed in the animals. The scoring varies by hub, day, and station, as well as scoring differences for different symptoms and types of misreporting.

| Variables | Description | Scoring |
|---|---|---|
| correctly_marked | Correctly reported symptoms | 10 |
| low-level mistake | Low-level mistakes avoided | 1 |
| mid-level mistake | Mid-level mistakes avoided | 2 |
| high-level mistake | High-level mistakes avoided | 3 |

**Additional Information (AS)**

Here CAHWs are rewarded for not filling out parts of the forms that need to be filled out by their supervisors. Additionally, it gives point not filling out fields that represent other cases not present at the station, namely wild and domestic animals.

| Variables | Description | Scoring |
|---|---|---|
| other_events | Other events | 5 |
| sub_info_yn | submission field correctly left blank? | 3 |

**Consolidating Animal Station Scores**

The highest possible score that can be obtained in subsections cases and symptoms varies by hub, day and station; scores are therefore normalized: raw scores are divided by the total possible score at a given station. Next,form scores are averaged by CAHW. Those CAHWs who did not attend the excercise get 0 points (the minimum score). Lastly scores are standarized.

| | n | mean | sd | min | max |
|---|---|---|---|---|---|
| Identifiers | 287 | 0 | 1 | -3.44 | 0.65 |
| Cases | 287 | 0 | 1 | -3.50 | 0.47 |
| Symptoms | 287 | 0 | 1 | -3.62 | 0.52 |
| Additional Information | 287 | 0 | 1 | -3.38 | 0.34 |
| AS Score | 287 | 0 | 1 | -3.59 | 0.50 |

**Correlation Matrix (AS)**

| | identifiers | cases | symptoms | additional_info |
|---|---|---|---|---|
| identifiers | 1.00 | 0.93 | 0.94 | 0.88 |
| cases | 0.93 | 1.00 | 0.97 | 0.89 |
| symptoms | 0.94 | 0.97 | 1.00 | 0.92 |
| additional_info | 0.88 | 0.89 | 0.92 | 1.00 |

**Cronbach's alpha (AS)**

**Alpha Total (AS)**

| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | ase | mean | sd | median_r |
|---|---|---|---|---|---|---|---|---|---|
| | 0.98 | 0.98 | 0.98 | 0.92 | 46.84 | 0 | 0 | 0.97 | 0.92 |
| ##### | Alpha Drop | (AS) | | | | | | | |

| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | alpha se | var.r | med.r |
|---|---|---|---|---|---|---|---|---|
| identifiers | 0.97 | 0.97 | 0.97 | 0.93 | 37.76 | 0 | 0 | 0.92 |
| cases | 0.97 | 0.97 | 0.96 | 0.91 | 31.36 | 0 | 0 | 0.92 |
| symptoms | 0.96 | 0.96 | 0.95 | 0.90 | 26.87 | 0 | 0 | 0.89 |
| additional_info | 0.98 | 0.98 | 0.97 | 0.95 | 53.25 | 0 | 0 | 0.94 |

**Item Statistics (AS)**

Table 9: raw.r is the correlation of the item with the entire scale, std.r is the correlation of the item with the entire scale, if each item were standardized. r.drop is the correlation of the item with the scale composed of the remaining items.

| | n | raw.r | std.r | r.cor | r.drop | mean | sd |
|---|---|---|---|---|---|---|---|
| identifiers | 287 | 0.97 | 0.97 | 0.95 | 0.94 | 0 | 1 |
| cases | 287 | 0.98 | 0.98 | 0.97 | 0.96 | 0 | 1 |
| symptoms | 287 | 0.99 | 0.99 | 0.99 | 0.98 | 0 | 1 |
| additional_info | 287 | 0.95 | 0.95 | 0.92 | 0.91 | 0 | 1 |

## Form Review

For Form Review, two research assistants randomly selected up to 5 forms submitted by the CAHW since installation. The assistants went over the form in detail with the CAHWs and quizzed them about why the forms were filled out the way they were and what specific fields indicated. The idea was to check how well the CAHW understood how to report the symptoms using the form.

The Form Review scoring is dividided into sections, a) top of the form, where the RAs evaluated the CAHW's correct understanding of identifiers, and b) the bottom of the form, where RAs evaluated the CAHW's abiity to understand how to report cases, symptoms and additional information.

**Form Review Scores**

Scores are normalized based on the total possible points that a CAHW could have obtained in this exercise. The total possible points depend on whether CAHWs reported any cases and include any additional information.

We give 0 points (the minimum score before standardization) to CAHWs who didn't attend the field practical exercise. We then standardize scores.

From the 267 CAHWs that participated in the field practical 8 had not sent any form by the time of the evaluation and thus RAs didn't have any form to review. We impute the standardized score that they recieved in the Animal Station exercise.

**Summary Statistics**

Table 10: The top of the forms has information on identifiers, and the bottom includes points obtained for additional information and correctly identified symptoms and cases

| | n | mean | sd | min | max |
|---|---|---|---|---|---|
| Top of the form | 287 | 0.00 | 0.99 | -2.42 | 0.98 |

| | n | mean | sd | min | max |
|---|---|---|---|---|---|
| Bottom of the form | 287 | 0.01 | 0.99 | -2.42 | 0.98 |
| FR Score | 287 | 0.01 | 0.99 | -3.18 | 0.79 |

**Correlation Matrix (FR)**

| | fr_top_p | fr_bot_tot |
|---|---|---|
| fr_top_p | 1 | 1 |
| fr_bot_tot | 1 | 1 |

**Cronbach's alpha (FR)**

**Alpha Total (FR)**

| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | ase | mean | sd | median_r |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 1 | 1 | 1 | 530.67 | 0 | 0.01 | 0.99 | 1 |
| ##### | Alpha Drop | (FR) | | | | | | | |

| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | alpha se | var.r | med.r |
|---|---|---|---|---|---|---|---|---|
| fr_top_p | 1.00 | 1 | 0.99 | 1 | NA | NA | 1.00 | 1 |
| fr_bot_tot | 0.99 | 1 | NA | NA | NA | NA | 0.99 | 1 |

**Item Statistics (FR)**

Table 14: raw.r is the correlation of the item with the entire scale, std.r is the correlation of the item with the entire scale, if each item were standardized. r.drop is the correlation of the item with the scale composed of the remaining items.

| | n | raw.r | std.r | r.cor | r.drop | mean | sd |
|---|---|---|---|---|---|---|---|
| fr_top_p | 287 | 1 | 1 | 1 | 1 | 0.00 | 0.99 |
| fr_bot_tot | 287 | 1 | 1 | 1 | 1 | 0.01 | 0.99 |

# Field Visits

As part of the validation process, facilitators visited 287 CAHW communities to asses the CAHWs' performance in the field. Facilitators stayed two days in each of the communities where they observed and recorded sick animals, and also tried to follow-up on the filled forms sent by CAHWs. These visits aim at identifying animals that CAHWSs falsely reported sick, and sick animals that CAHWs failed to report.

By the time of our field visits (May 2018), 20 CAHWs had not submitted any filled form, and therefore facilitators didn't have any form to follow up or match against.

## Sick Animal Observation

If CAHWs are doing their job well, they should be actively walking around their communities looking for and reporting on sick animals. To determine if CAHWs have been doing this, facilitators inspected each community and recorded the number of sick animals that could be found, their symptoms, and the name of the owner. These surveys were then matched against the CAHW reports, to see if the CAHWs had also properly reported on these animals. The number of unreported sick animals found by the facilitator is then used to judge the CAHW's work. More unreported sick animals corresponds to worse performance.

CAHW gets a reward for properly reporting a case. If a CAHW report matches by our more restrictive process they get a higher reward than if it matches under our generous matching process. The main distinction between these processes is the inclusion of community member name as a matching criteria. Those that match by the restrictive process will be excluded from double counting. For those that have multiple matches of the same type (ex. several generous matches for the same case) the case resulting in the highest number of points is kept. In all cases, the rewards are discounted by the level of misreporting in symptoms. (Max and Luke, 2018)

We recorded two types of missing data for this part of the scoring: a) CAHWs for whom no animal observations were found by facilitators (53 obs) and b) CAHW communities where facilitators observed sick animals but CAHWs had not sent any form by the time of the visit (14 obs). For type a) we imputed the standardized mean, and for type b) the lowest possible score (0 before standardizing).

SAO scores are computed as follows:

$$sao\_score = 50\frac{C_g}{S_g}\frac{M_g}{e} + 100\frac{C_r}{S_g}\frac{M_r}{e}$$

Where,

$M_g$ : generously matched cases of reported sick animals.

$M_r$ restrictively matched cases of reported sick animals.

$S_g$ : total number of symptoms reported in all generously matched cases.

$S_r$ : total number of symptoms reported in all restrictively matched cases.

$C_g$ : total number of correctly reported symptoms in all generously matched cases.

$C_r$ : total number of correctly reported symptoms in all restrictively matched cases.

$e$ : our estimate of the # of animals in the community.

- We estimate the # using data from other surveys we have conducted
- Broadly, we multiply the number of HHs in the community by the average # of animals per HH

### Summary statistics (SAO)

| | n | mean | sd | min | max |
|---|---|---|---|---|---|
| Restrictively matched | 287 | 0.01 | 0.99 | -0.08 | 14.69 |
| Generously matched | 273 | 0.03 | 0.99 | -0.18 | 12.36 |
| SAO Score | 287 | 0.03 | 0.99 | -0.20 | 11.85 |

## Monitoring form follow-up

Under Monitoring Form Follow-Up, up to 5 reporting forms per CAHW were randomly selected. Then facilitators went to the communities and interviewed the owner of the sick animals reported in the forms. In the interviews, they were asked about how many animals they owned, if their animals were sick in the recent past, whether the CAHW treated their sick animal etc. The aim was to verify whether CAHW was correctly reporting sick animals in the community.

In this section, we take the average for all reports we have full information on (with a maximum of 20 points). Then, CAHWs get up to 10 points for not providing bad information (forms with unrecognized names, for example). MFU scores are computed following the formula below. Final Scores are standardized.

$$MFU_i = F + 10(\frac{(R - B)}{R})$$

Where,

$F$ : is the average report score.

$R$ : the total number of reports followed up on.

$B$ : the number of the unsuccessful reports in which the reason the person could not be found was bad

Lastly, scores are standardized. Communities with no monitoring forms to follow-up (20 obs.) get the standardized mean value as their final score.

| | n | mean | sd | min | max |
|---|---|---|---|---|---|
| F | 287 | 0 | 0.964 | -1.656 | 1.406 |
| 10((R – B) / R) | 287 | 0 | 0.964 | -2.577 | 0.490 |
| MFU Score | 287 | 0 | 0.964 | -2.185 | 1.225 |

### Consolidating Field Visits Scores

### Summary Statistics (Field Visits)

| | n | mean | sd | min | max |
|---|---|---|---|---|---|
| SAO | 287 | 0.032 | 0.986 | -0.197 | 11.847 |
| MFU | 287 | 0.000 | 0.964 | -2.185 | 1.225 |
| Field Visits Score | 287 | 0.000 | 1.000 | -1.695 | 8.444 |

### Correlation Matrix (Field Visits)

| | sao_final_score | mfu_final_score |
|---|---|---|
| sao_final_score | 1.000 | 0.067 |
| mfu_final_score | 0.067 | 1.000 |

### Cronbach's alpha (Field Visits)

### Alpha Total (Field Visits)

| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | ase | mean | sd | median_r |
|---|---|---|---|---|---|---|---|---|---|
| | 0.13 | 0.13 | 0.07 | 0.07 | 0.14 | 0.1 | 0.02 | 0.71 | 0.07 |
| ##### | Alpha Drop | (Field Visi | ts) | | | | | | |

| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | alpha se | var.r | med.r |
|---|---|---|---|---|---|---|---|---|
| sao_final_score | 0.07 | 0.07 | 0 | 0.07 | NA | NA | 0.07 | 0.07 |
| mfu_final_score | 0.00 | 0.07 | NA | NA | NA | NA | 0.00 | 0.07 |

**Item Statistics (Field Visits)**

Table 21: raw.r is the correlation of the item with the entire scale, std.r is the correlation of the item with the entire scale, if each item were standardized. r.drop is the correlation of the item with the scale composed of the remaining items.

| | n | raw.r | std.r | r.cor | r.drop | mean | sd |
|---|---|---|---|---|---|---|---|
| sao_final_score | 287 | 0.74 | 0.73 | 0.19 | 0.07 | 0.03 | 0.99 |
| mfu_final_score | 287 | 0.72 | 0.73 | 0.19 | 0.07 | 0.00 | 0.96 |

# Final Component Scores

In the sections below we show the final scores using two different approaches. In the first approach we present final scores derived from the 4 subsections (Animal stations, Form Review, Sick animal observation and Monitoring form follow-up), and in the second approach we aggregate sick animal observation and Monitoring form follow-up into one item, field visits. The correlation matrix and the Cronbach´s alpha are provided.

## Validation Scores (4 components)

### Summary Statistics

| | n | mean | sd | min | max |
|---|---|---|---|---|---|
| AS | 287 | 0.000 | 1.000 | -3.592 | 0.505 |
| FR | 287 | 0.000 | 0.964 | -2.185 | 1.225 |
| SAO | 287 | 0.007 | 0.987 | -3.181 | 0.792 |
| MFU | 287 | 0.032 | 0.986 | -0.197 | 11.847 |
| Validation Scoring | 287 | 0.000 | 1.000 | -2.818 | 4.877 |

### Correlation Matrix (4 components)

| | as_final_score | mfu_final_score | fr_final_score | sao_final_score |
|---|---|---|---|---|
| as_final_score | 1.00 | 0.02 | 0.91 | 0.05 |
| mfu_final_score | 0.02 | 1.00 | 0.10 | 0.07 |
| fr_final_score | 0.91 | 0.10 | 1.00 | 0.03 |
| sao_final_score | 0.05 | 0.07 | 0.03 | 1.00 |

### Cronbach's alpha (4 components)

### Alpha Total (4 components)

| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | ase | mean | sd | median_r |
|---|---|---|---|---|---|---|---|---|---|
| | 0.5 | 0.49 | 0.64 | 0.2 | 0.98 | 0.05 | 0.01 | 0.62 | 0.06 |
| ##### | Alpha Drop | (4 componen | ts) | | | | | | |

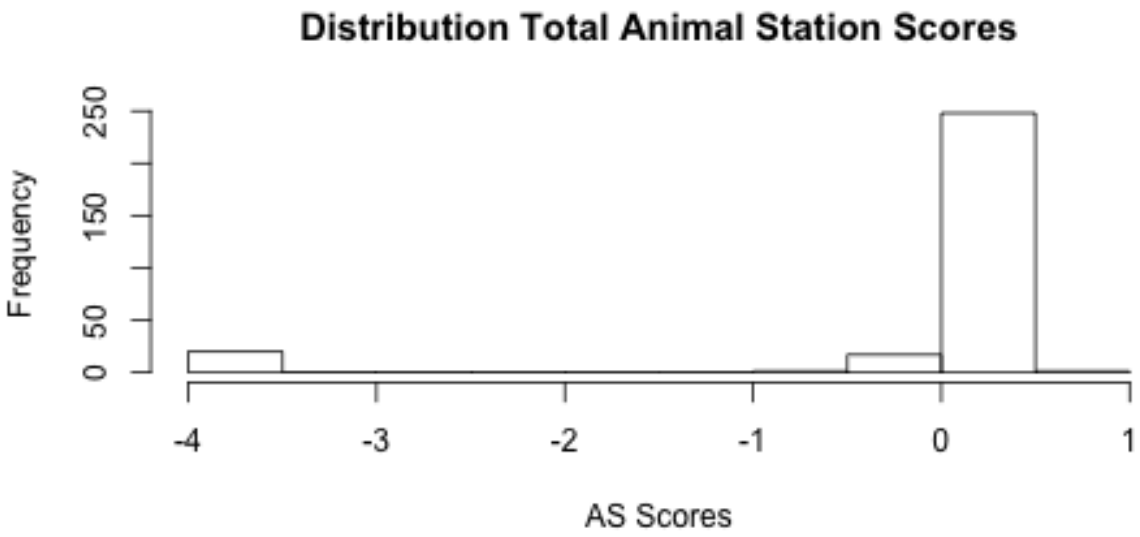
| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | alpha se | var.r | med.r |
|---|---|---|---|---|---|---|---|---|
| as_final_score | 0.17 | 0.17 | 0.12 | 0.07 | 0.21 | 0.08 | 0.00 | 0.07 |
| mfu_final_score | 0.60 | 0.60 | 0.73 | 0.33 | 1.47 | 0.04 | 0.25 | 0.05 |
| fr_final_score | 0.13 | 0.13 | 0.09 | 0.05 | 0.15 | 0.09 | 0.00 | 0.05 |
| sao_final_score | 0.62 | 0.61 | 0.74 | 0.34 | 1.57 | 0.04 | 0.24 | 0.10 |

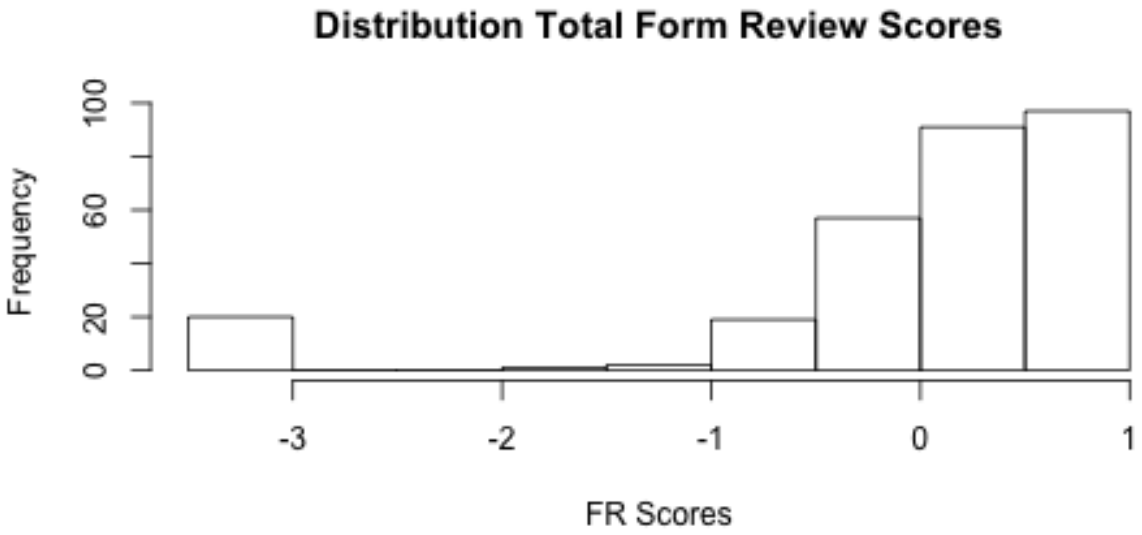
### Item Statistics (4 components)

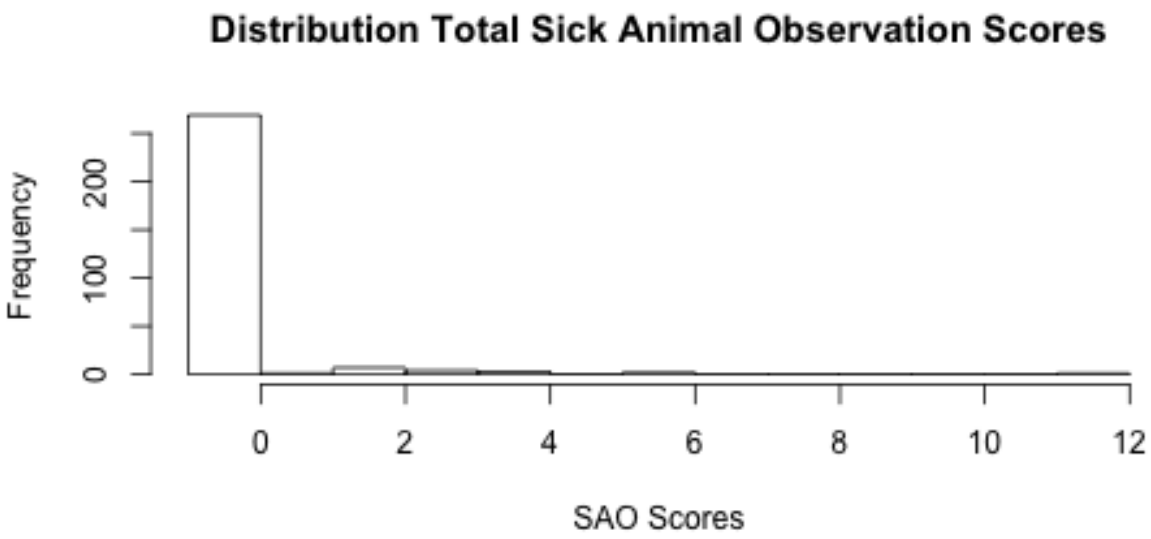
Table 26: raw.r is the correlation of the item with the entire scale, std.r is the correlation of the item with the entire scale, if each item were standardized. r.drop is the correlation of the item with the scale composed of the remaining items.

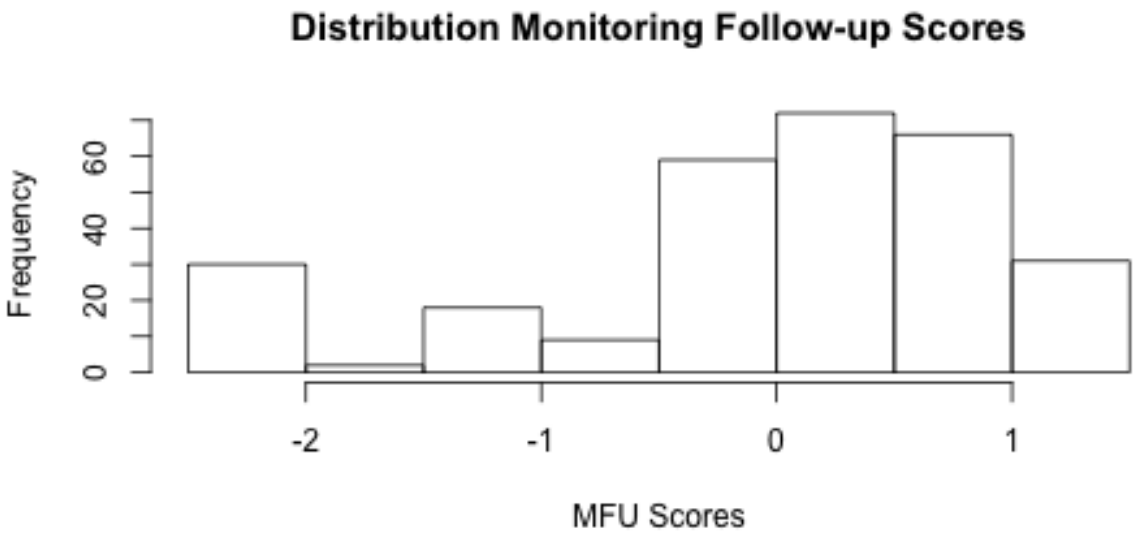
| | n | raw.r | std.r | r.cor | r.drop | mean | sd |
|---|---|---|---|---|---|---|---|
| as_final_score | 287 | 0.79 | 0.79 | 0.90 | 0.54 | 0.00 | 1.00 |
| mfu_final_score | 287 | 0.46 | 0.47 | 0.11 | 0.09 | 0.00 | 0.96 |
| fr_final_score | 287 | 0.81 | 0.81 | 0.93 | 0.58 | 0.01 | 0.99 |
| sao_final_score | 287 | 0.45 | 0.45 | 0.08 | 0.06 | 0.03 | 0.99 |

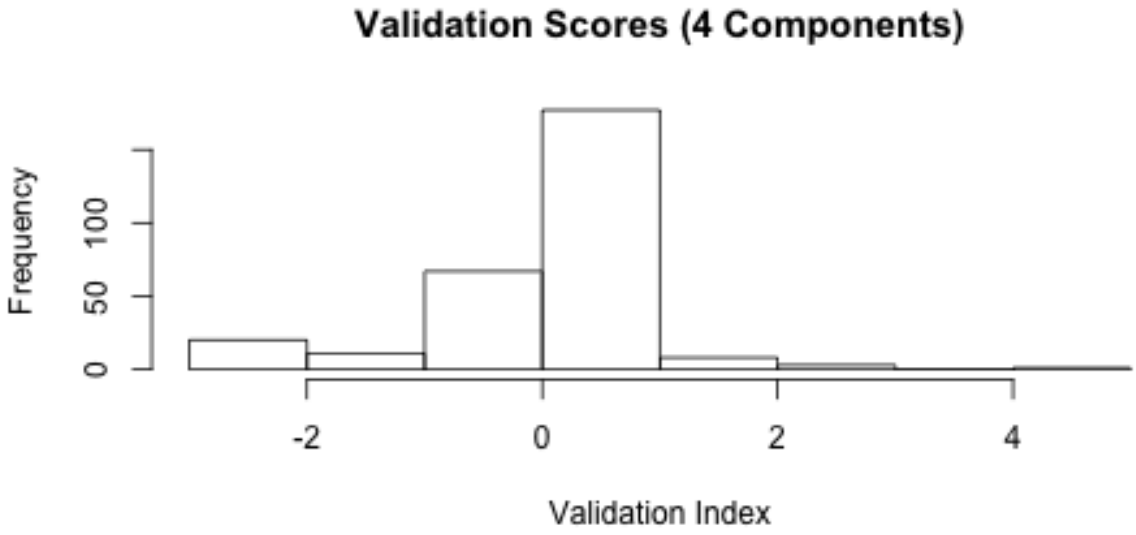
**Figures: Distribution of components scores**


Distribution Total Animal Station Scores
Frequency
0
50
150
250
-4
-3
-2
-1
0
1
AS Scores


Distribution Total Form Review Scores
Frequency
0
20
60
100
-3
-2
-1
0
1
FR Scores


Distribution Total Sick Animal Observation Scores
Frequency
0
100
200
0
2
4
6
8
10
12
SAO Scores


Distribution Monitoring Follow-up Scores
Frequency
0
20
40
60
-2
-1
0
1
MFU Scores


Validation Scores (4 Components)
Frequency
0
50
100
-2
0
2
4
Validation Index

## Validation Scores (3 components)

In this section we aggregate Sick Animal Observation and Monitoring form follow-up into one component, field visits. We again present summary statistics, correlation matrix and Cronbach's alpha.

### Summary Statistics

| | n | mean | sd | min | max |
|---|---|---|---|---|---|
| AS | 287 | 0 | 1.000 | -3.592 | 0.505 |
| FR | 287 | 0 | 0.964 | -2.185 | 1.225 |
| Field Visits | 287 | 0 | 1.000 | -1.695 | 8.444 |
| Validation Scoring | 287 | 0 | 1.000 | -3.091 | 3.808 |

### Correlation Matrix

| | as_final_score | mfu_final_score | field_visits_scores |
|---|---|---|---|
| as_final_score | 1.00 | 0.02 | 0.05 |
| mfu_final_score | 0.02 | 1.00 | 0.72 |
| field_visits_scores | 0.05 | 0.72 | 1.00 |

### Cronbach's alpha (3 components)

### Alpha Total (3 components)

| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | ase | mean | sd | median_r |
|---|---|---|---|---|---|---|---|---|---|
| | 0.52 | 0.52 | 0.58 | 0.27 | 1.09 | 0.05 | 0 | 0.71 | 0.05 |
| ##### | Alpha Drop | (3 componen | ts) | | | | | | |

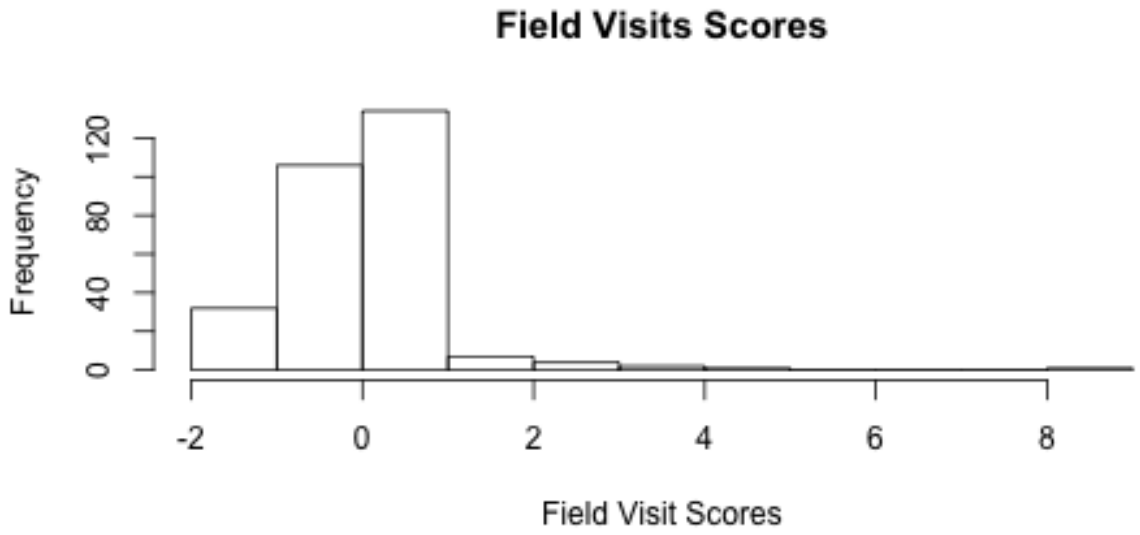
| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | alpha se | var.r | med.r |
|---|---|---|---|---|---|---|---|---|
| as_final_score | 0.84 | 0.84 | 0.72 | 0.72 | 5.23 | 0.02 | NA | 0.72 |
| mfu_final_score | 0.10 | 0.10 | 0.05 | 0.05 | 0.11 | 0.11 | NA | 0.05 |
| field_visits_scores | 0.05 | 0.05 | 0.02 | 0.02 | 0.05 | 0.11 | NA | 0.02 |

### Item Statistics (3 components)

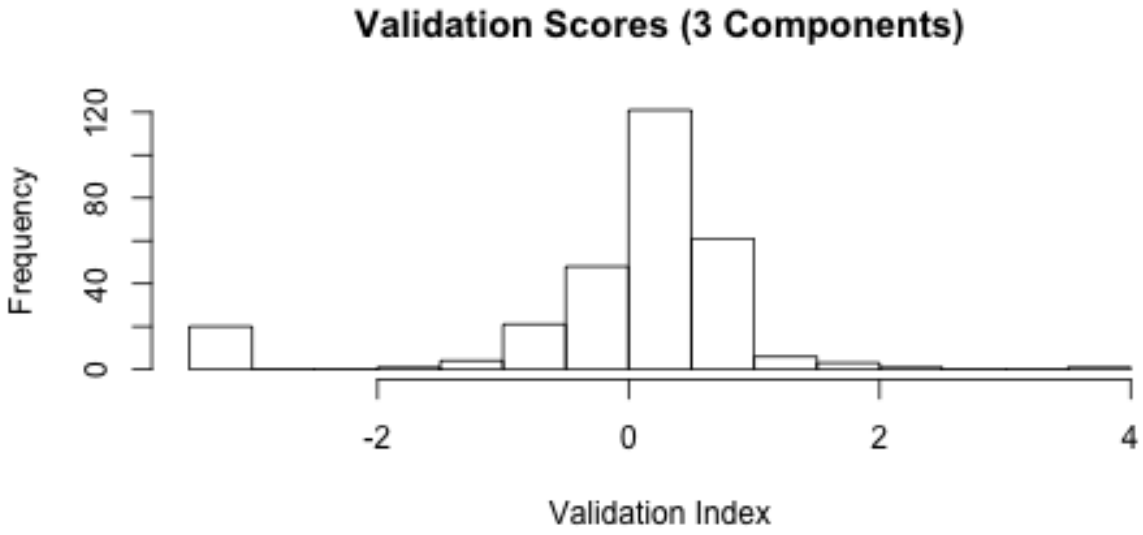
Table 31: raw.r is the correlation of the item with the entire scale, std.r is the correlation of the item with the entire scale, if each item were standardized. r.drop is the correlation of the item with the scale composed of the remaining items.

| | n | raw.r | std.r | r.cor | r.drop | mean | sd |
|---|---|---|---|---|---|---|---|
| as_final_score | 287 | 0.51 | 0.50 | 0.05 | 0.04 | 0 | 1.00 |
| mfu_final_score | 287 | 0.81 | 0.82 | 0.78 | 0.52 | 0 | 0.96 |
| field_visits_scores | 287 | 0.83 | 0.83 | 0.80 | 0.53 | 0 | 1.00 |

**Figures: Distribution of components scores**




```
# Scores with grouped field visits
p123 <- quantile(final_score$final_score, probs = c(1/3, 2/3) )
final_score %<>% mutate( tier = ifelse( final_score <= p123[1], 3,   ifelse(final_score < p123[2], 2, 1
```


Validation Scores (3 Components)
Frequency
0
40
80
120
-2
0
2
4
Validation Index