Collections Management and the Strawberry Island Collection: Making Corrections to the Collection in 2021

Andrew Gillreath-Brown

July 27, 2021

Introduction

The Strawberry Island (45FR5) collection is a massive, but an invaluable collection with approximately 36,770 artifacts in the inventory (excluding the 320 documents). Archaeological work on Strawberry Island took place from 1967 until about 1980, although most of the excavations were conducted from 1977–1979. The original inventory was done in just over a year, spanning approximately July 17, 1992 to November 10, 1993 (dates taken from the first and last original coding sheets). However, additional inventory coding was done in September 2004.

In summer 2021, Andrew Gillreath-Brown led a team of four to evaluate and fix errors within the collection, including artifact bags and tags, coding sheets, and database. We focused on rehabbing 30 boxes in the collection. All 325 boxes of the collection were brought from the WSU Museum of Anthropology warehouse to the Museum of Anthropology in College Hall in summer 2019. This report details the changes and methods used to work on the collection in summer 2021 (see Gillreath-Brown and Fairlane (2019) for information about the 2019 work). The main goals for the 2021 work included, but not limited to:

- Check identification on artifacts and revise as necessary,
- check for duplicate numbers and assign new numbers as necessary,
- identify artifacts and raw materials that were listed as "Not identified,"
- rebag as necessary (many older "red line" zip bags do not age well), and
- retag as necessary (paper tags and incomplete sharpie writing on outer surface of bag).

Collection Summary

The collection has a total of 37,090 inventory entries in the database. Chipped stone artifacts make up much of the collection (52.18%), which is followed by bone (23.14%), samples (10.33%), and shell (7.44%) (Figure @ref(fig:fig1)).

The number of artifacts per box varies greatly across the collection, ranging from 1–920 artifacts in 325 boxes. Figure @ref(fig:fig2) shows the distribution of the number of artifacts per box.

In addition, the number of unidentified chipped stone (i.e., CS-25-55) artifacts varies across each box ranging from 0–721 artifacts per box. However, most boxes (n = 288) have 0 to 25 unidentified chipped stone artifacts (not shown in Figure @ref(fig:fig3)). Specifically, 252 boxes have zero CS-25-55 designations, and 36 boxes have between 1–25 unidentified chipped stone artifacts. Overall, there are 5,502 CS-25-55 artifacts in the collection. Figure @ref(fig:fig3) shows that some boxes have a lot of CS-25-55.

2004 Work on the Collection

In 2004, additional work was undertaken on the collection. However, the exact purpose of the additional coding is unclear. Furthermore, several mistakes were made during the 2004 work, including the recoding (or reusing) of old numbers, which are randomly dispersed throughout the collection. One of the goals was
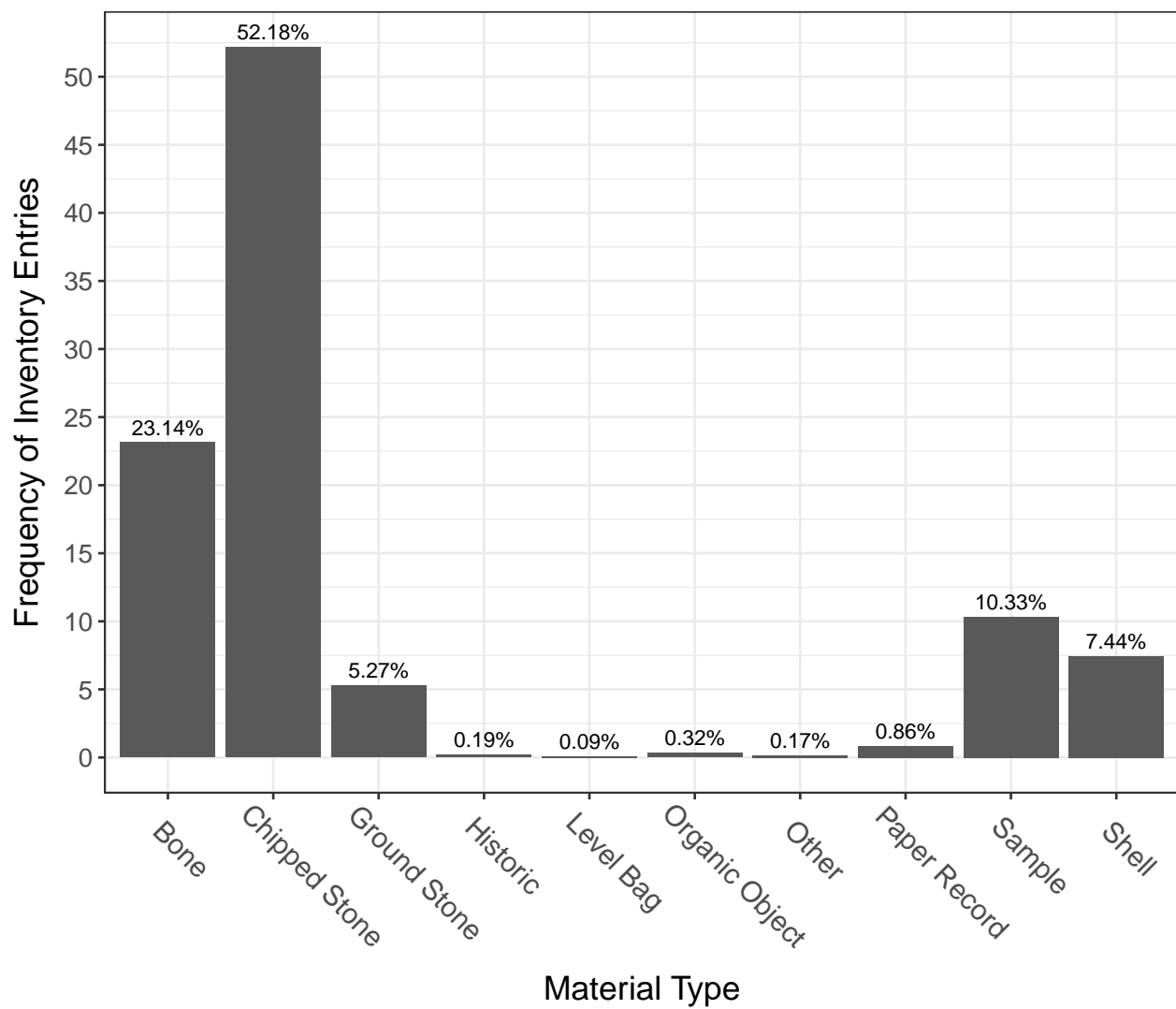
Figure 1: The frequency of material types in the Strawberry Island collection.
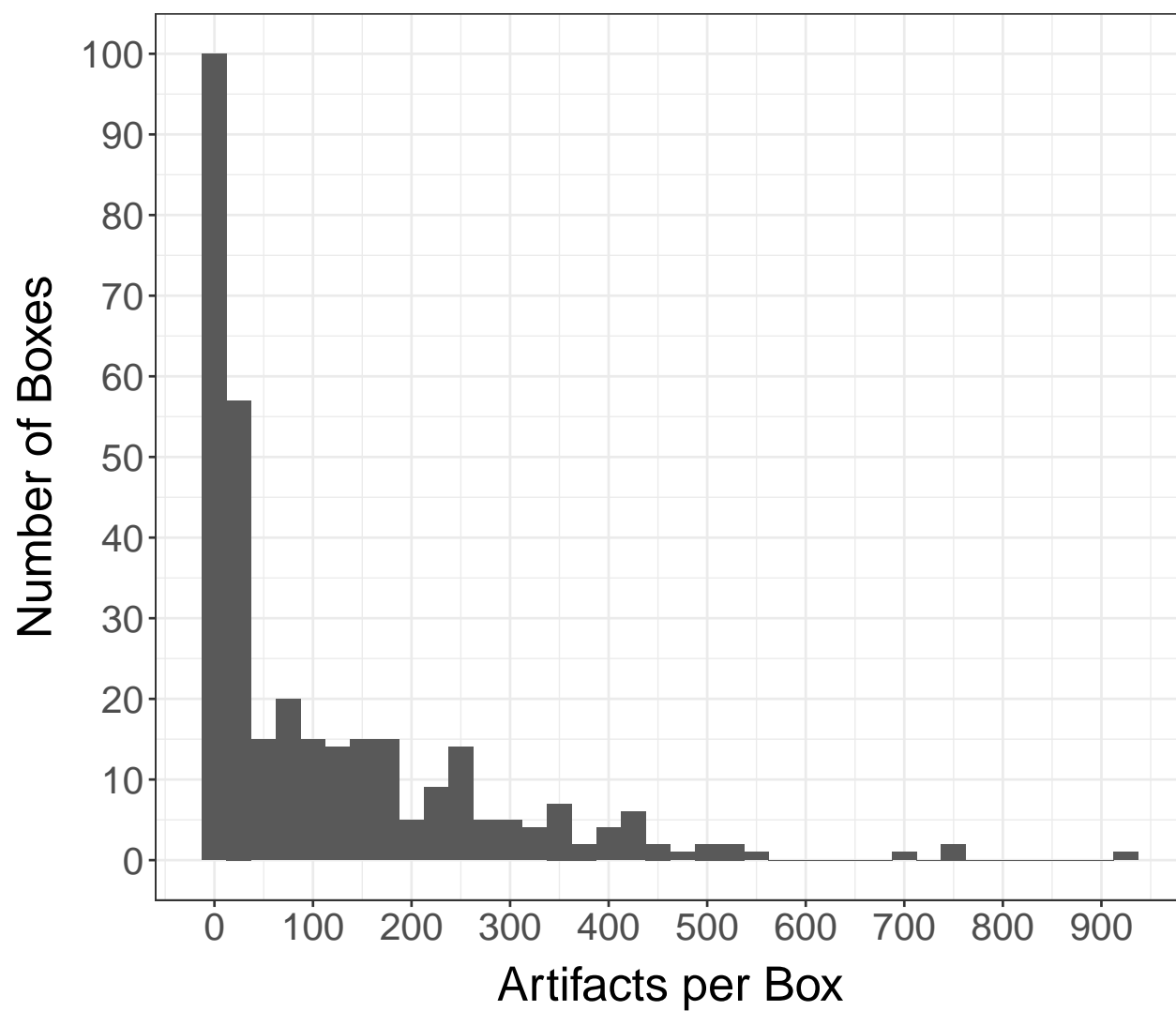
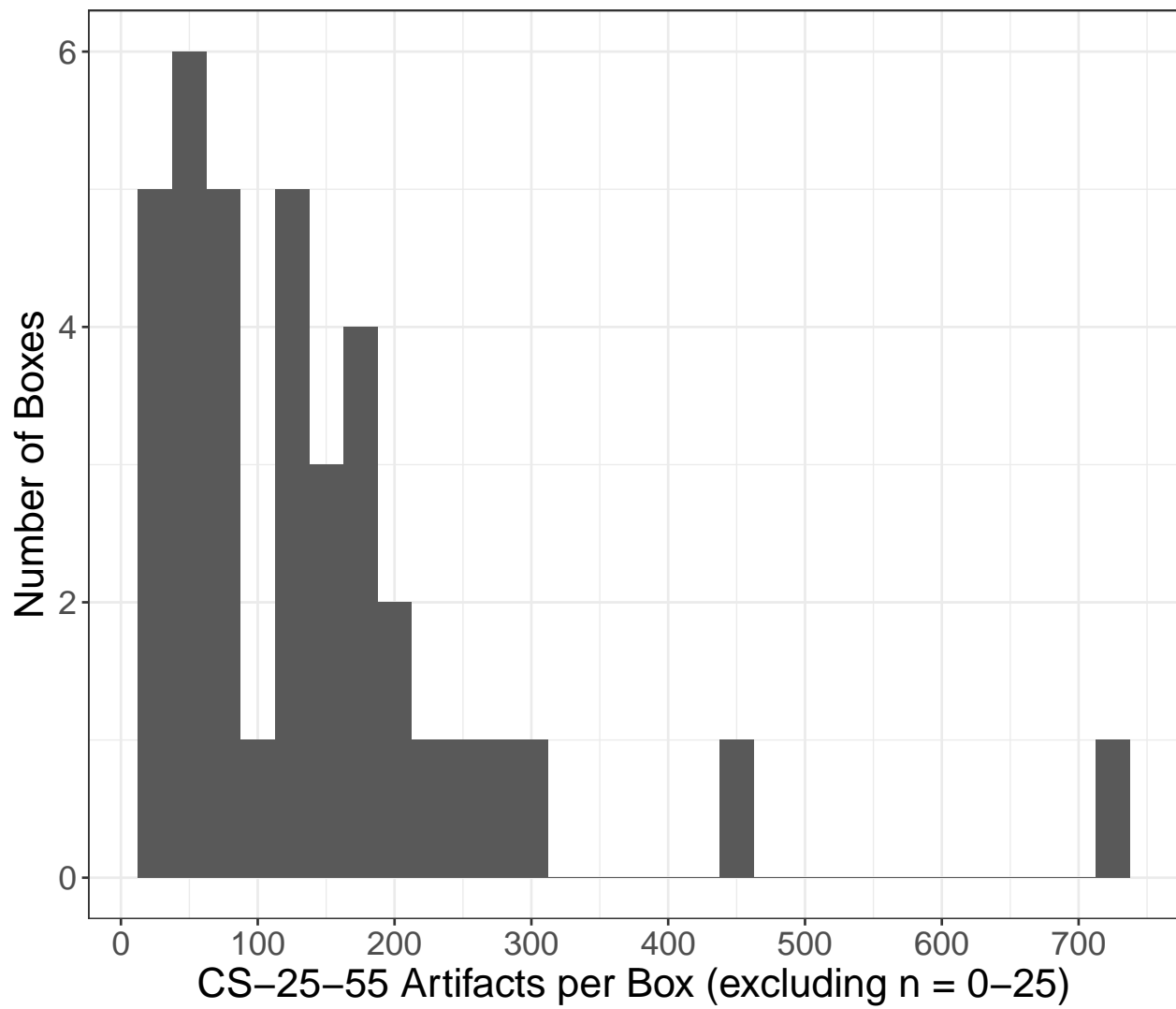Figure 2: The distribution of number of artifacts per box.

Figure 3: The distribution of number of unidentified chipped stone artifacts (CS-25-55) per box, excluding the 0 to 25 bin.

clearly to identify and code artifacts. However, they sometimes re-used inventory numbers that they thought were not being used rather than creating new inventory numbers. However, this led to several mistakes, including re-using inventory numbers that were already assigned to another artifact. It is possible that at some point someone checked some of the artifacts and decided that they were either not artifacts or that they should be bagged with other artifacts from the same provenience. Thus, in 2004, someone may have tried to use the non-used inventory numbers. Further, they may have tried to remedy the non-used inventory numbers, but ended up creating more issues in the process. Box 192 seems to hold a lot of the problems that were created in 2004.

Methods

For 2021, the main process was to go through each box and examine each artifact. We used Google Sheets to document any change that was made to the bag, tag, or database. This is the same document that I used to document changes in 2019; however, one column distinguishes whether the work was completed in 2019 or 2021. This will allow someone in the future to look at the one document to see what changes were made for 2019 and 2021 rather than having to examine two documents. We marked any change in red and added comments for specifically what was changed for an artifact in the Google Sheets document. We checked off inventory numbers for artifacts as they were examined, detailing that an inventory entry had been checked.

We examined the information on the bag, tag inside the bag, and the database to ensure that all information matched. We also checked off each item on the box list to ensure that the box inventory was correct. If there were any discrepancies, then we checked the comment field to see if there were any notes about the provenience. If the correction was not clear (particularly if the outside of the bag, the tag, coding sheet, and/or the database had different information), then we checked the original field catalog using the catalog number (if available), the original bag (if available), and/or the artifact notecards, which cover more excavation areas than the original field catalog. Sometimes the original bags were in one artifact bag, then the other bags specified which bag (by inventory number) that the original bag was in. We also evaluated the identification of the artifact, then made any necessary changes to the identification. If an artifact was missing a count or weight, then we added that information. Furthermore, if an artifact did have a count or weight, then we added a comment to the comments field to clarify the material in a way that could be helpful to a future researcher. For example, if there was a bag of mammal bone, which had been counted rather than weighed, then we wrote in the comments about whether there were large identifiable pieces or tiny bits of unidentifiable bone (or something in between). We wanted to give researchers an opportunity to target the best samples.

We also opted to use the new museum coding style, as we could make new identifications much quicker. For example, if an artifact was coded as not having cortex (CS-15), when it clearly did have cortex, then it was recoded as "CS-DB," rather than using the old system. Conversely, if an artifact was coded as "CS-15," but clearly had retouch. Then, it was re-coded as "CS-RE."

Each person working was required to complete a checklist as they worked on a box to ensure that each required step had been completed (Figure @ref(fig:checklist)).

Once the changes were documented in Google Sheets, then the information in the box list and the original coding sheets were updated. Further, the Corps of Engineers scanned all the archival copies of the coding sheets in 2019. Since there have been changes to many of the original coding sheets, these needed to be updated, so that there is a digital version of the coding sheets, which may have other changes since the archival copies had been made. Thus, we scanned the updated versions of the coding sheet and added the documents to the digital collection. Therefore, the scanned archival copies and the copies of the original coding sheets that are stored in the records room do not have the most up-to-date information. However, it might be necessary to consult these records if someone needs to see what has changed over time.

Results and Discussion

In 2019, I checked and made corrections to 13 boxes (i.e., Box 1–13), which were comprised of approximately 2,072 artifact bags, and to an artifact in Box 192, which was a duplicate. There were 2 new inventory numbers

Checklist

Name: _____   **Box Number: _____**

1. Were all changes recorded in the Google Sheets document?

    Y   N   Comments:

2. Were all changes made to the Box List?

    Y   N   Comments:

3. Were all changes made to the Original Coding Sheets?

    Y   N   Comments:

4. Were all changes recorded in the Microsoft Access Database?

    Y   N   Comments:

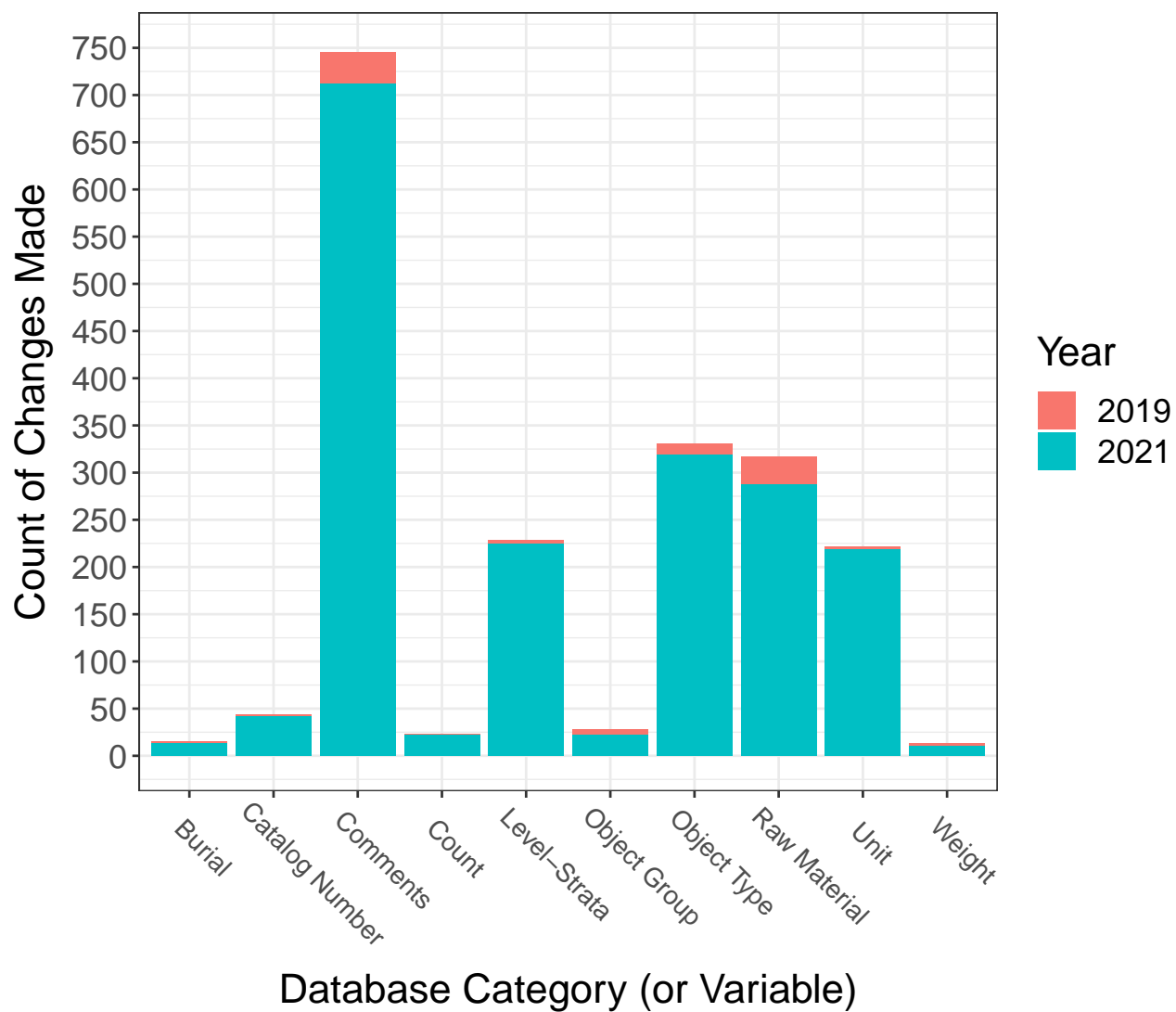Figure 4: Checklist that each person was required to fill out for each box.

Figure 5: The number of changes for each category across the different categories in the database for 2019 and 2021.

added in 2019. I made a total of 94 changes to the database for the 13 boxes across ten categories of the database (Figure @ref(fig:fig5)).

In 2021, we checked and made corrections to 30 boxes of artifacts (Table @ref(tab:tab1)), which were comprised of approximately 3,025 artifact bags. We also added 4 new inventory numbers to the collection. After completing the work in summer of 2019, we thoroughly evaluated the 2019 process, then worked on revising the plan for rehabilitation. Thus, in 2021, we worked towards scrutinizing each inventory entry. One major difference is that we consulted the original field catalog and artifact notecards to resolve many of the provenience issues. Unfortunately, some of the original field catalog and artifact notecards likely never made it into the collection; thus, for some artifacts, we were unable to resolve provenience issues.

Table 1: Box numbers and the number of inventory numbers in each box for the 30 boxes that were completed in 2021.

Box Number	Inventory Count
14	245
15	759
17	5
18	12
19	42
24	37
31	8
36	17
37	31
44	4
53	186
73	51
74	50
89	104
90	111
101	68
107	117
116	107
119	120
144	100
198	68
212	47
226	103
229	220
241	49
258	6
262	103
301	116
307	50
323	106

In 2021, we made a total of 1,871 changes to the database for the 30 boxes across ten categories (or variables) of the database (Figure @ref(fig:fig5)). We created a column in Google Sheets that detailed exactly what was changed for each inventory entry (if any changes were made for a given entry). For example, if the coding sheet and database had an incorrect unit number, then we specified exactly what was changed. So, if the coding sheet and database had S42/E91, but the correct number was S44/E94, then we would write, "We updated the unit in the coding sheet and database, which was changed from S42/E91 to S44/E94." We kept a very detailed record of what exactly was changed for each inventory entry. We made a lot of changes to the

comments field, where we clarified information about the artifact itself and/or the provenience/metadata. Generally, we provided more detailed information in the database comments, including, but not limited to, bone type and condition, use wear, raw material, clarification on provenience, and size of shell. The other main changes were adjusting the object type, raw material, excavation unit, and the level and strata.

Many of the changes detailed did not require alterations to the coding sheet or database, since some changes were only made to the bag, tag inside the bag, and/or the box list. The changes mostly consisted of the following, but not limited to, re-coding an artifact; adding additional comments, which could relate to the coding (e.g., raw material type); updating the exterior of a bag to match the tag inside (as well as double-checking that this matches the coding sheet and database); and coding new inventory entries. Appendix 1 has all inventory entries that were checked and any changes that were made to a given entry for 2019 and 2021.

In the original coding of the artifacts, artifacts were sometimes bagged in very inconsistent ways. For example, in Box 73, the database showed that there were about 22 individual inventory numbers that had shell. However, at some point, all the pieces of shell were placed in one bag with all 22 inventory numbers written on the outside of the bag and all the tags placed inside. These inconsistencies make it difficult to detail the counts and weights for an individual entry. In this instance, we weighed the entire bag with all 22 inventory numbers and put the weight in the first entry in the database, then had no weight for the remaining 21 inventory entries in the database, except for adding the weight to the comments. Furthermore, a comment was added to the database that inventory numbers 13,990 to 14,011 were all in the same bag, and that the weight listed for 13,990 includes all the inventory numbers. In some instances, larger pieces of shell that had been bagged together still had the original catalog number written on the artifact, which means theoretically the shell could be separated out into individual bags. However, we did not separate the shell back into individual bags, as it would be time consuming and is unnecessary with the corrections and comments that we made.

Provenience errors across the bags, tags, coding sheets, and database were abundant. In some instances, the bag, the tag, and database would have different numbers for the unit number. Thus, when possible, we would consult the original field catalog and/or the artifact notecards to resolve unit issues. We discovered that sometimes the database had been corrected, but nothing else had been updated. Ultimately, we figured out what was the most likely correct provenience, then updated bags, tags, coding sheets, and database so that they all matched. However, any numbers that were changed were recorded in the Google Sheet. We also used the lithic coding sheets to resolve provenience issues. However, the lithic coding sheets are difficult to use as the numbers are not in a sorted order by catalog number or area. For example, in Box 229, many artifacts (138 of 220) did not have any provenience (i.e., unit or level/strata) recorded in the database or coding sheets. Thus, we used the lithic coding sheets to find the provenience, as there were very few original catalog numbers available for the areas (i.e., A200 and A203). Ultimately, due to these inconsistencies, new tags and bag labels need to be printed for all artifacts.

We corrected many identifications on chipped stone and ground stone artifacts. For ground stone artifacts, the most misidentified type was abrader, which were typically moved to "groundstone other." We made 319 changes to object types and 287 changes to the raw material designations (excluding comments that were made about the raw material when the artifact is identified as RM = 50, which means "other stone"). We also updated the identifications of many chipped stone artifacts that had been labeled as CS-25-55, which means unsorted debitage and no assigned raw material. Many of these artifacts were easily identifiable, especially with using the new museum coding style (mentioned above).

We also took several steps and actions towards making the collection more accessible when researchers need to go through boxes. For example, we grouped inventory bags together with fewer inventory bags in each larger bag. For example, Box 53 only had 2 bags of grouped artifacts, although it had a total of 186 artifacts. This makes it difficult to quickly find an artifact when there are almost 93 artifacts in one bag. We also noticed that many of the older bags no longer closed, particularly the red line bags that hold large stone artifacts. Thus, for expediency's sake, we put the bags into newer bags (though you could still clearly see the provenience on the outside of the bag) that were able to seal. However, if new tags and bag labels are ever created, then the old bags could be discarded. We also had to add provenience and identification information

to the outside of bags, when the bag only had the inventory number.

We also found another duplicate, where inventory number 15,321 was duplicated. This was found because the database had different information than the coding sheets and there was another artifact that was labeled as inventory number 15,321 in Box 53. In Box 53, inventory numbers 15,321 and 15,322 were bagged together (though the catalog numbers were written on the artifacts). The bagging together likely led to inventory number 15,321 being excluded from the inventory and from the box list. To fix the situation, we corrected inventory number 15,321 in Box 53 in the database, then assigned the other artifact in Box 192 a new inventory number (i.e., 37,089). It is still highly likely that Box 192 holds more duplicates. However, we are unlikely to find the other artifacts without going through the rest of the collection. However, one quicker solution would be to check all artifacts in Box 192 against the coding sheets, as this could reveal which ones are likely duplicates if they have other information than a given artifact in Box 192.

Conclusion and Future Work

Overall, the collection needs a lot more work to make it a reliable collection to use for research. Additionally, the identifications in the collection are very inconsistent in terms of the level of identification, how much information was provided, and how accurate that information is.

Overall, the collection needs a lot of updates, edits, and corrections, as there are many errors in the collection. Additionally, many artifacts (i.e., n = 5,502) labeled as CS-25-55 are coded as unidentifiable; however, many of those are identifiable, which also includes a lot of identifiable raw material for the artifacts.

This collection, which is an extremely important collection, continues to need a major update. Some of the codes for artifacts are incorrect, such as a ground stone (particularly abraders, as well as others) that should have been labeled chipped stone; raw materials being misidentified/coded; and incorrect object types. Additionally, there were many people working on the collection that did not understand lithics. However, the new museum coding style makes this much easier by having fewer specific categories. Much of the raw material is identifiable, which is frequently either basalt or chert. Next, many of the items need to be rebagged, particularly the bags with the red zip line, as many of the bags do not close (many with chipped and/or ground stone). Next, several of the bags have incomplete information on the outside of the bag. So, people wrote certain fields, like feature, but then did not actually fill in the information. The tags inside of bags also have this issue as some of the tags only have the inventory number and the catalog numbers. The entire collection needs new tags and to have bag labels to ensure that the marker does not wear off many of the bags. Finally, unit numbers throughout the database, as sometimes the more precise numbers are used and sometimes the less precise are used.

The digital database also needs a lot of work, as new fields could be created, and some fields need to be separated into two columns. For example, strata and level are in the same column, when they should be in two columns. Much of the time, the strata are listed in the comments field, although sometimes both strata and level are in the strata/level column. The date is also frequently in the comments field, although there is a date column. The date being the comments is particularly a problem with inventory numbers greater than 20,000. Other information in the comments, such as elevation levels, also need to be in new columns. There are also inconsistencies with the precision on the units and inconsistent formatting across rows.

References

Gillreath-Brown, Andrew, and Cassady Fairlane. 2019. *Reassessment and Corrections of the Strawberry Island (45fr5) Site Collection in Summer 2019.* Museum of Anthropology, Washington State University.