2 CONTENT ANALYSIS

BACKGROUND

What do people think? How many people think the same way? Is there a relationship between the prevalence of cultural information and individual psychology? These are the kinds of ethnographic questions that form the basis of cultural inquiry. This chapter focuses on the content of what people list, the ways in which that content is structured in individual minds, and how content is distributed across those individuals' minds.

Recall from the previous chapter that free-lists are especially useful for eliciting cultural data. Free-list data can also tell us about important properties of human cognition. In the following chapters, when I refer to **cognitive** or **representational models**, I refer to the content and structure of individuals' mental phenomena (D'Andrade, 1981; Hruschka, Sibley, Kalim, & Edmonds, 2008). Free-listing tasks, data, and analyses capture these knowledge sets and their structures. If we take the view that "culture" is shared, socially transmitted information, then **cultural models** are the pools of shared, socially transmitted units composed of individual cognitive models. Defined in this fashion, the prevalence of particular units of socially transmitted information indicates just how "cultural" those units are. These units are the individual items provided in free-lists.

In this view, then, directly assessing the way cultural models correspond to—or even induce—individual behavior requires (a) detailing the content of representational models, (b) assessing how widespread the content of those models is, and (c) examining the relationships between individuals' behavior and cognitive and cultural models. The first two requirements are descriptive, ethnographic accounts of cultural domains. The third disambiguates the relative impacts of individual and cultural models of a given domain on corollary behavior. If culture predicts behavior, then the prevalence of cultural models' constituent units in a group should covary with the target behavior. In principle, we should be able to thus link such methods to others in order to see a richer portrait of what we're trying to explain (D'Andrade, 1999; Weller & Romney, 1988).

We'll revisit these issues later in the book. In the meantime, this chapter will probe the unique aspects of free-list data that attend to their content. The first few sections walk through some basic descriptive approaches to free-list data. These methods examine the frequency, distribution, and salience of cultural content. Almost immediately, we will see important facets of the nature of lists, the structure of domains, and the distribution of cultural information. Later sections offer other alternatives and downstream methods that complement those of the earlier section.

It is important to note here that the methods in this and the next few chapters are predominantly *exact* methods; the values and graphs that they present often provide **point estimate** solutions rather than generate estimates with some range of likelihood and/or uncertainty around them. In other words, they describe the data as they pertain to the sample really well, and the statistics we examine are really precise, but they offer little in the way of anything indicating how confident we might be in their reproduction, likelihood, or generalizability. These descriptive methods are nevertheless extremely useful, especially while in the field and in need of some quick summaries of a domain. If all you want to do is describe the content and structure of the individual and cultural models of your sample, these methods are excellent tools. If you want to use aspects of these cognitive and cultural models to relate to other measurable aspects of ethnographic reality, we'll need to link them in more appropriate frameworks (e.g., regression). Later chapters will build on what we do in the present chapter and use inferential statistics with formal indices of uncertainty to bridge content analysis and other forms of prediction.

FREQUENCY ANALYSIS

Frequency analysis describes a distribution of counts related to the lists we collect. List length, for example, is one such count that can tell us a lot. We'll see that in cases where free-list lengths are not capped at a specific limit, the frequency of items that individuals list can indicate their knowledge of the domain; the more things they list, the more they know. As we'll see in Chapter 6, with this knowledge often comes other important traits such as accuracy and behavioral aptitude in that domain (Brewer, 1995; Gatewood, 1984; Koster, Bruno, & Burns, 2016).

In addition to list length, we can also look at the frequency of specific items that a sample lists. Just as the distribution of list lengths might indicate the distribution of knowledge that a sample has about a particular domain, the general shape of the frequency distribution of listed items gives you some insight into the relative heterogeneity of thought in a given sample and shows us which items are more culturally ubiquitous in a sample. If we want to understand what people are likely to think, know, feel, and remember about a particular domain, examining the frequency distribution of listed items is very useful.

R CODE BOX 2.1: FREQUENCY DISTRIBUTION PLOT

```
FL <- read.delim("FL.txt") # load data
FL.bin <- FreeListTable(FL, CODE = "CODE",
       Order = "ORDER",
       Subj = "Subj",
       tableType = "PRESENCE")
SUM <- colSums(FL.bin[,2:ncol(FL.bin)]) # sums
FREQ <- data.frame(SUM) # turn into data.frame
newdata <- FREQ[order(-FREQ$SUM)„ drop = F] # sort
barplot(newdata$SUM, names.arg = rownames(newdata),
       las = 2) # plot
```

This code will make a horizontally oriented version of the histogram in Figure 2.1 (see accompanying code for the vertically oriented version). The first line calls up the data and creates an object called `FL` that contains the data set. The second line creates a participant-by-item presence matrix indicating whether (= 1) or not (= 0) each participant listed each item type. After that, we take the sums of each column (i.e., the total frequency of participants who listed each item), followed by a line of code that sorts these sums from most to least frequent. The final line plots this distribution with the listed items as the *x*-axis labels. To instead plot the proportions of items listed, this should suffice:

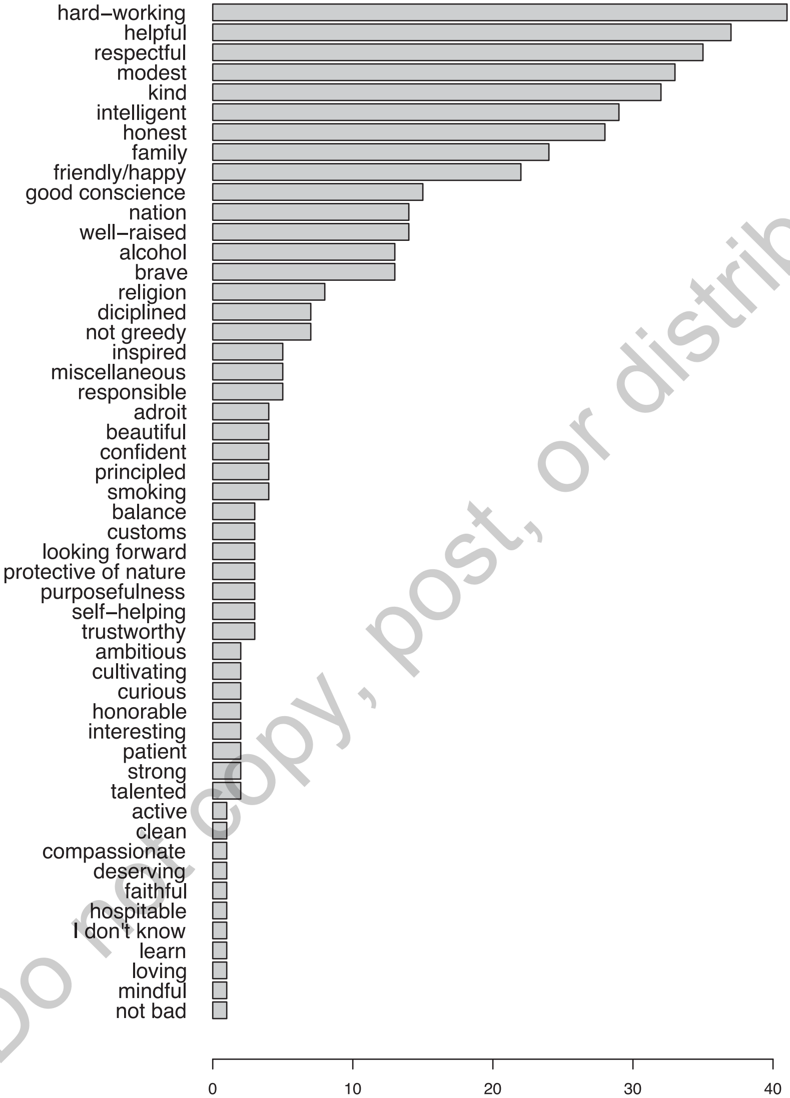
```
n <- length(unique(FL.bin$Subject)) # if no empty rows
barplot(newdata$SUM/n, names.arg = rownames(newdata),
     las = 2, ylim = c(0, 0.5))
```

The supplementary code provides a few different ways of plottingfrequency distributions (though, of course, these styles aren't limited to frequency data).

To illustrate, let us look at some data collected among a grand sample[1] of 89 indigenous southern Siberians from the Tyva Republic (Purzycki & Bendixen, 2020). In this study, we needed a better sense of what it means to be a "good Tyvan person," so we asked participants to free-list what they thought. Figure 2.1 shows the frequency distribution of the listed items

[1] Here, "grand sample" refers to the number of participants in an entire study rather than the number of individuals whose data appear in a particular analysis. As some folks don't answer or if we decide to drop "I don't know's," we'll get a smaller sample for analyses.

FIGURE 2.1 ■ Frequency Distribution of What Tyvans ($n = 89$) Think Constitutes a Good Tyvan.



Note: Tyvans listed 51 different items. Note that item frequencies are the total number of individuals listing items. (data from Purzycki & Bendixen, 2020).

(R Code Box 2.1 shows how to make this kind of plot in R). As you can see, among the most frequently listed characteristics are "hard-working," "helpful," "respectful," "modest," "kind," "intelligent," and so forth. The distribution is highly skewed to the right; the frequency of items tapers off down to a long, thin string of singly listed items, including "active," "clean," "compassionate," "deserving," "faithful," and "hospitable."

There are a few practical things to note here. First, as members of a "soft" domain (i.e., one filled with items that are relatively difficult to pin down), many of these items beg for even deeper investigation. What it means to be "respectful," for example, is left to our imagination; we could do another study to learn what Tyvans mean by this, just as we could do a full-scale study of what "hard-working" means. The possibilities are endless with soft domains. Second, despite this, these items are all coded quite specifically. Things like "kind" and "compassionate" are treated as different items. They certainly don't need to be. Similarly, one individual might list both "telling the truth" and "not lying," even though both items might be coded as "honesty" instead. In a focused and principled study, the decision to lump or split such items should be up to a well-defined rubric, ideally drawn from some theory of interest (see Chapter 1). In practical or purely exploratory contexts, however, one is left to their own devices. This is especially so when analyzing data in softer domains.

Third, when graphing frequency distributions, it is important to carefully attend to *how* you present items and their distribution. The 13th most-frequent item, for example, is labeled "alcohol"; this actually means *abstaining* from or not abusing alcohol. It pays to know what your data actually say! If someone else is coding the data, it is imperative to check closely if anything is ambiguous (see Chapter 1). Fourth, sometimes people just don't want to answer or can't provide anything much beyond "I don't know." Here, one did. Such responses can be important information, particularly in domains where there is considerable loss of knowledge or ambiguity in the domain. For salience analyses (see below), such answers are typically the only thing an individual says. Including "don't know's" in analyses can alter the overall appearance of responses.[2]

In open-ended free-lists tasks, how many items individuals list might tell you about expertise and how knowledge is distributed in a community. Knowing who domain experts are is quite useful for follow-up studies. As we'll

[2]If it makes sense to, you can easily recode things in R. Good practice dictates that you first create a new column. Simply replicate the original column, say, `data$FL2 <- data$FL1`, where `data` is the data set, `FL1` is the original column, and `FL2` is a new column that's exactly the same as `FL1`. In the case of converting "don't know's" into `NA`s in R, assuming your "don't know's" are coded as "D/K," this line of code will replace all of the "D/K" values in `FL2` with `NA`: `d$FL2[d$FL2 == "D/K"] <- NA`.

see in later chapters, this knowledge can be associated with *behavioral* mastery of a domain. The supplementary code shows how to plot the distribution of list lengths in a few different ways. By the same token, we can also gain insight into variation in cross-group knowledge. If one were interested in examining how much more a sample knows about, say, animals than plants, one could calculate the sum of listed items across domains and apply standard statistical mean comparisons with regression (see Chapter 5).

One could also assess the number of items listed across variables such as age, sex, occupation, and political affiliation. This often matters in making predictions. For instance, some find that age differences in free-lists are domain-specific (Purzycki & Bendixen, 2020; Schrauf & Sanchez, 2010). Presumably, other factors will also matter differently across domains. One specific case study in Dominica found that the more years of education Dominicans had, the fewer medicinal plants they listed, but having a commercial occupation predicted longer lists, as did age and being female (Quinlan & Quinlan, 2007). We'll cover prediction in Chapters 5 and 6.

In and of themselves, frequency distributions really only scratch the surface of what can be done with free-list data, but they can be very useful, especially when coming to terms with group-level properties of individuals' lists. As we'll see, the frequency of listed items tends to correspond to the order in which individuals list them. In other words, how culturally particular (or popular) items are affects the structure and accessibility of knowledge in individuals' minds. Cultural salience analysis examines this relationship.

SALIENCE ANALYSIS

Item Salience (s_i)

Essential to free-list methods is recording (1) who says (2) what and (3) in what order they say it. The order in which items are listed gives us a sense of how accessible or cognitively *salient* items are for individuals (Smith & Borgatti, 1997). We can formally define **item salience**, s_i, as

$$s_i = \frac{n + 1 - k_i}{n} \tag{2.1}$$

where n is the number of items listed and k_i is the order in which an item was listed. The value of s_i, then, is scaled so that $0 < s_i \leq 1$. Table 2.1 illustrates a hypothetical spreadsheet, D, of freely listed animals from two individuals. Each listed item has a corresponding s_i score. Code for how to do this rapidly on a real data set is in R Code Box 2.2.

TABLE 2.1 ■ Example Spreadsheet (*D*) of Two Individuals in Hypothetical Sample Free-Listing Animals.

ID	Order	Item	s_i
TVA001	1	sheep	1.00
TVA001	2	goat	0.67
TVA001	3	yak	0.33
TVA002	1	goat	1.00
TVA002	2	sheep	0.83
TVA002	3	cow	0.67
TVA002	4	camel	0.50
TVA002	5	crane	0.33
TVA002	6	yak	0.17

Note: ID refers to unique participant number, and s_i refers to value from Equation 2.2.

R CODE BOX 2.2: ITEM SALIENCE

```
FL.S <- CalculateSalience(FL,
        CODE = "CODE", Salience = "Salience",
        Subj = "Subj", Order = "Order")
```

This code creates a new object `FL.S` from the main free-list data `FL`. `FL.S`has the added column of item salience scores to the original data (stored as a new object so we do not alter `FL`). It identifies which items to analyze (`CODE = ""`), what to call the salience variable (`Salience = ""`), participant ID (`Subj = ""`), and the order of listed items (`Order = ""`). You can conduct item salience analysis on the Tyvan data illustrated in Figure 2.1 using the supplementary code.

Importantly, with this simple metric, we can see how cultural transmission relates to the structure of thought. More specifically, s_i tends to be positively correlated with how many individuals in a sample list the same item. In other words, the sooner individuals list items, the more likely they are to be listed by other individuals in a sample. This relationship lies precisely at the nexus of interaction between culture and individual cognition. We can see this in Figure 2.2.[3] If we take the Tyvan virtue data we saw earlier (Figure 2.1), calculate the number of Tyvans who listed specific items and each item's average item salience, there is a strong positive relationship.[4] This fact shows us an important aspect of culture, namely, that the prevalence of a concept in human communities corresponds to individual conceptual structure; *items' mentally represented relationships are often predicted by how many other minds share those items*.[5] One method condenses this insight into a single metric, namely, **cultural salience**.

R CODE BOX 2.3: SUBSETTING DATA AND REMOVING MISSING CASES

```
labs <- c("Subj", "Order", "CODE", "Salience")
FL.sub <- FL.S[labs] # subset
FL.c <- FL.sub[complete.cases(FL.sub), ] # axe NAs
```

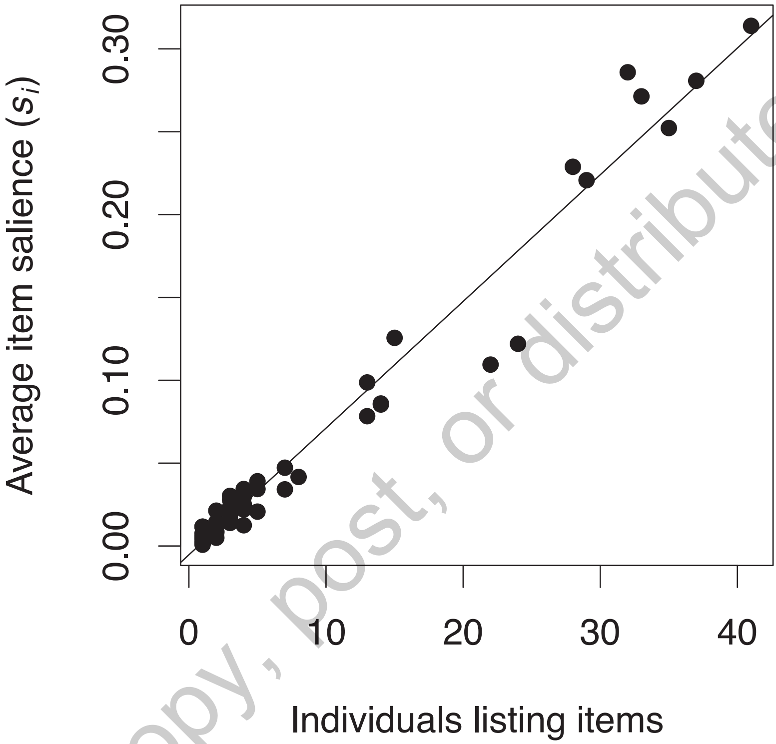
In the event that you have individuals who listed *nothing* in your data set (i.e., a row with `NA`), calculating all of the other item saliences will work, but be advised that the next few steps will require you to remove those individuals. One way to avoid that is to use the following to remove all rows with missing values. The first line creates a vector of variable names in order to subset the `FL.S` object (see above). The second line subsets those variables. The third line takes only *rows* without missing data and calls that `FL.c`. The brackets indicate that we are subsetting the data frame rows, hence the comma *after* the `complete.cases()` command. If you ever want complete cases of columns, you would put the comma before the `complete.cases()` command.

[3]Turning the data into a form that allows us to plot things takes some data wrangling. Often, we need to remove missing data from data sets to do other operations. R Code Box 2.3 offers one way to subset data and remove missing values.

[4]Note, however, that this relationship is not always so sharply linear (Bimler & Uusküla, 2021; Uusküla & Bimler, 2016).

[5]Of course, other contextual factors are likely to alter the ubiquity or recall of items. For example, temporal context can be very important in particular tasks; some research shows that freely listing information about medicinal plants is predicted by their recent use (Sousa et al., 2016).

FIGURE 2.2 ■ Correlation between how many individuals listed items and how high they are on individual lists on average.



Note: The x-axis is number of Tyvans who listed specific items and the y-axis is the average salience of that item in Tyvans' lists (data from Purzycki & Bendixen, 2020).

Cultural Salience (S_i)

Item salience is represented as an index of where an item is on an individual's list, weighted by the number of items listed. In contrast, *cultural* salience values indicate the aforementioned association between item placement and sample ubiquity. One common way of calculating the cultural salience of a concept is Smith's S (Smith & Borgatti, 1997). Calculating this entails taking the sum of all specific items' s_i scores and dividing that by the total number of participants who completed the task, N:

$$S = \frac{\Sigma \frac{(n+1-k_i)}{n}}{N} = \frac{\Sigma s_i}{N} \quad (2.2)$$

Defined in this fashion, Smith's *S* increases with position and frequency, assuming a constant sample size. R Code Box 2.4 shows you how to implement this in R.

R CODE BOX 2.4: CULTURAL SALIENCE (SMITH'S *S*)

```
SAL.tab <- SalienceByCode(FL.S,
      CODE = "CODE", Salience = "Salience",
      Subj = "Subj",
      GROUPING = "Grouping",
      dealwithdoubles = "MAX")
```

This code creates a cultural salience table (object `SAL.tab` using thep-reviously made `FL.S` object). It specifies all the same arguments as the previous code, with two additional components. First, with the `GROUPING` command, it allows us to calculate salience across multiple groups. If you have a variable that categorizes a group of interest (e.g., participant sex, cultural group, or any other category), it will chunk Smith's *S* scores by categories in this variable. Second, it has various ways to handle items that were listed multiple times (or items that were coded as the same type during the coding process). These options are `"MAX"`, `"AVERAGE"`, and `"MIN"`. You can conduct cultural salience analysis on the Tyvan data illustrated in Figure 2.1 using the supplementary code. If you want to sort the resulting table by Smith's *S*, the following will do so:

```
SAL.tab[order(-SAL.tab\$SmithsS)„ drop = F]
```

To make a flower plot like that in Figure 2.3, you input the object created by the `SalienceByCode()` function and give it a domain name. For the left flower plot in the figure, the code would look something like this: `FlowerPlot(SAL.tab, "Illness")`. You won't need to sort the data first, as the `FlowerPlot()` function will do it automatically. The supplementary code also includes options to hard-code flower plots for more customization.

As noted earlier, one recurring practical issue is that individuals can list the same item multiple times. This can happen by accident, or participants might use synonymous or nested items using different lexical tags that we the researchers want to treat similarly. One might, for example, list "ram" and "ewe" as two different animals. At the coding stage, however, we might be tempted to recode these items as "sheep." This therefore introduces multiple instances of the same code. At the analytical stage, there are a few different ways of addressing this, and the `AnthroTools` package automates these options. You might opt for the maximum s_i for repeated items and ignore the less salient equivalents. This option prioritizes the more salient-listed instance of the item type. This is quite reasonable, since any repeats shouldn't necessarily negate earlier-listed items. If you use this option in `AnthroTools`, the list length for individuals is corrected as a result. Another option to handle repeated items is to take the mean s_i of multiple items, but this is likely to reduce the apparent salience of items listed early and repeated much later in lists. You might also wish to make a more conservative estimate for repeat items by using the lowest s_i value instead.

Once we get our S values, we're now in a position to examine cultural models. One way to do this is to sort the output table and eyeball the salience values (see `R` Code Box 2.4). Another option is to plot these values. Figure 2.3 illustrates one way to convey cultural salience. These "flower plots" include the eight most[6] culturally salient items from a free-list task and present them in an intuitive way. The domain name is in the center circle, and the most salient item is in the top circle. Salience scores descend in a clockwise fashion, with bars of varying connection weights between the domain name and each item adjusted for S scores. The wider the connection weight, the more likely an individual will list it *and* it will be earlier in lists.

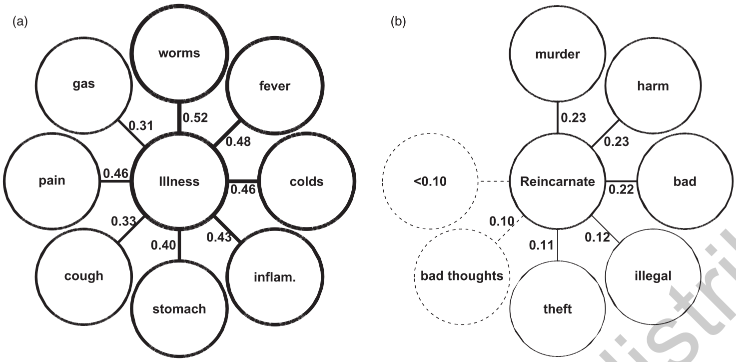
The left panel (a) illustrates the eight most salient items from a task that asked rural residents of Dominica ($n = 30$) to list the kinds of illnesses they treat with bush medicine (Quinlan & Quinlan, 2007). The most culturally salient illness to treat with bush medicine was "worms" and each subsequent item has a slightly lower Smith's S score. The right panel (b) illustrates the cultural salience of items listed from Chinese Buddhists, who were asked to list behaviors that negatively impact reincarnation (Willard, Baimel, Turpin, Jong, & Whitehouse, 2020). Here, "murder," "harm," and "bad behaviors" are the most culturally salient items. In this case, anything with a $S \leq 0.10$ is illustrated by the dotted line, thus implying a relatively "weak" connection between the domain and any other items.

As suggested by its formal definition, "salience" is a relative index. Indeed, this threshold of 0.10 illustrated in the flower plot—along with what constitutes "highly salient" for that matter—is not only arbitrary but relative to the other items' salience scores. Note too the relative differences in salience across these two tasks in Figure 2.3. The "things you treat with remedies found

---

[6]The number 8 is purely for aesthetic reasons. Truth be told, I designed them this way because the circles are large enough to capture most text and small enough to fill in the visual space along with the other graphical elements in a balanced way. The accompanying code includes code to hard-code similar plots and could be easily adjusted to adjust the number of circles in your plot.

FIGURE 2.3 ■ Examples of Flower Plots of Smith's *S* Scores Across Two Studies.



Note: Panel (a) illustrates cultural salience of illnesses treated by bush remedies in Dominica (data from Quinlan & Quinlan, 2007) and Panel (b) illustrates cultural salience of items listed by Chinese Buddhists about the kinds of behaviors that have a negative impact on the outcome of reincarnation (data from Willard et al., 2020).

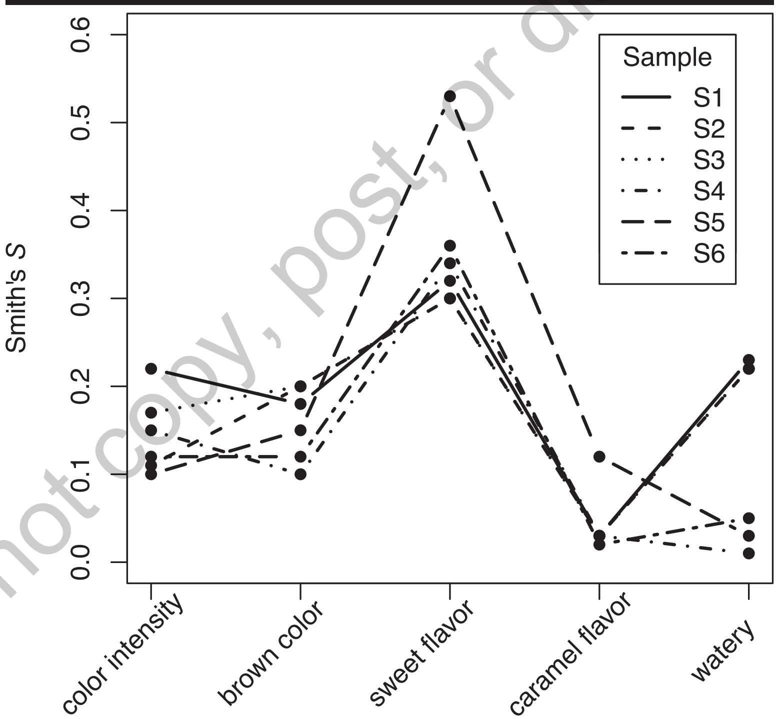
in the forest" contains a much more precisely defined set of units. Its most salient item is $S_{\text{worms}} = 0.52$, followed by seven items with fairly high cultural salience scores. However, the "things that have a negative impact on the outcome of getting reincarnated" includes vague items such as "bad thoughts" and behaviors. The cultural salience values quickly diminish to values < 0.10. Higher-inference domains elicit softer responses that have more overlap and therefore more coding ambiguity. Relatively softer domains are also likely to yield lower salience scores by virtue of their fluidity. Relatively harder domains are likely to have fairly concrete items and therefore less variation in coding and consistency in response patterns.

Having free-list data of the same domain collected among different groups is quite helpful for understanding the cultural variation that exists out there. Should the salience of an item or set of items in one group be different from another group's? If you expect or need to calculate salience across different groups (e.g., sex, cultural group, etc.), you can easily calculate cultural salience by applying Equation 2.2 to the different groups, adjusting the denominator N appropriately. The code in `R` Code Box 2.4 shows you how to rapidly calculate salience across groups using `AnthroTools`. Let's work through an illustration.

One study (Yoon, Kwak, Heo, & Lee, 2023) had participants try six different samples of coffee and freely list attributes of each sample. Here, "groups" were the samples the participants tried. As shown in Figure 2.4, there was some variation in Smith's S values across coffee samples (S1 through S6).

In particular, it is quite clear that "sweet flavor" was a more salient response to S2, even though it was a fairly salient attribute across samples. If you were trying to market a bitter coffee and this is all we had to go on, we might disqualify S2 for further consideration. Also, in response to Samples S1 and S4, "watery" came to mind sooner and more frequently than with the other samples. This example shows us that comparisons of Smith's *S* need not be limited to cultural groups. It also shows us that in addition to item and cultural salience, Smith's *S* can be a useful metric for **perceptual salience** as well. While opinions are clearly culturally mediated, it's unlikely that the participants were sharing their views with each other before reporting them. We'll assess group differences more formally in Chapter 4.

FIGURE 2.4 ■ Smith's *S* Scores of Five Attributes Freely Listed About Coffee Samples.



Note: Data from Yon, Kwak, Heo, and Lee (2023).

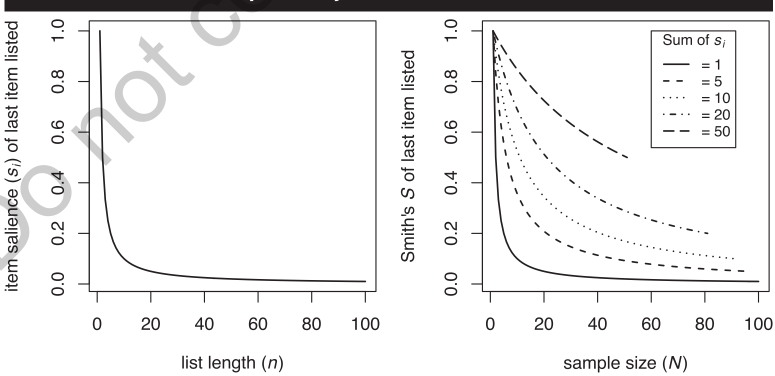
Notes on Cultural Salience

While very useful, there are some important things to keep in mind when using Smith's *S*. First, it is ideal for lists of similar or—even better—the same list lengths. Sometimes, participants trickle only a few things while others are geysers of knowledge. A consequence of this is that the last-listed item will have a different s_i value across lists of different lengths. In fact, s_i values become effectively indistinguishable from each other after list lengths > 20. Whenever list lengths are shorter than that, the last-listed item can have wildly different values.

You can see this in the left panel of Figure 2.5; item salience of last-listed items bottoms out quite quickly. We see how last-listed items' *S* values also decrease as a function of sample size. On the right panel of Figure 2.5, we see that at varying sums of s_i, Smith's *S* decreases at corresponding rates as the sample size increases. Overall, we can see that Smith's *S* decreases more dramatically the smaller the sample size is and the smaller the sum of s_i values are. Many thus subset the most salient items in a list for downstream analyses. The decision to truncate data like this will depend on the task at hand. Just be transparent and document your decision and motivation.

Unique singletons and cases of when participants answer with a mere "I don't know" can complicate analyses. Such instances pose the problem of whether or not to retain them in the sample. Does this count as an item listed? If so, including them might be interpreted as inflating *N*. Furthermore, such items' s_i values would always be 1. Understanding why "I don't know" happens and anticipating what you want to do in the event that it does is a good idea. Again, transparency and documentation help others to see why and how you made and executed your decisions.

FIGURE 2.5 ■ Item Salience (left) and Smith's *S* (right) Values Across List Lengths, *n*, and Sample Sizes, *N*, Respectively.



SALIENCE REVISITED

Other methods for calculating cultural salience emphasize the individual over cultural ubiquity. For example, one method (Sutrop, 2001) calculates cultural salience with

$$S_s = \frac{L}{N\bar{k}_i}$$

where L is the number of times an item appears across lists, N is the number of participants in the sample, and $\bar{k}_i$ is the average order number of the target item. The accompanying `R` code includes a function and application of Sutrop's S.

Note that this method can yield counterintuitive results. Take, for example, a hypothetical item that was listed first by one person in a sample of 100. Here, $S_s = 1/(100 * 1) = 0.01$. Smith's S would be the same. Yet, if two people in the same sample list an item and it has an average order of 4, then $S_s = 2/(100 * 4) = 0.005$. So, even though the second item is listed twice, its cultural salience score is half that of an item listed only once. If both individuals listed them at $k_i = 4$ with lists 10 items long, then Smith's S would be 0.014, roughly the same value. Furthermore, an issue that affects both Smith's and Sutrop's S is that we can arrive at the same value for different reasons. If, for example, a quarter of the sample listed something with an average order number of 18, $S_s \approx 0.01$. Yet, we get the same value if only one person from the same sample listed a particular item first. Of course, s_i will be 1 for all items listed first, but also if participants only list one item.

Item Salience (s_i')

One option (Robbins, Nolan, & Chen, 2017) that avoids such issues requires that we first calculate item salience, s_i', using

$$s_i' = \frac{n - k_i}{n - 1} \tag{2.3}$$

where, again, k_i is the order an item was listed and n is the number of items listed by the individual. This is slightly different from the s_i calculation in Equation 2.1. This equation rescales s_i so that the last-listed item of each individual's list is always 0, thus $0 \leq s_i' \leq 1$. If we used this method, then, the s_i values in Table 2.1 would be 1, 0.50, and 0 for participant TVA001 and 1, 0.80, 0.60, 0.40, 0.20, and 0 for participant TVA002. There are a few things to note with this particular metric.

First, it is not ideal in cases when you need to keep track of what people *don't* list. In other words, using a value for 0 for something that was listed last should not necessarily be treated as something that wasn't listed at all. This might matter for downstream analyses. Second, if a participant only lists a single item (it happens!) or replies "I don't know" and you wish to include such responses in a salience analysis, the equation is faced with an impossibility since we can't divide by zero (i.e., $(1 - 1)(1 - 1)$). Third, when all participants list more than one item, s_i' still has a perfectly linear relationship with the original s_i calculation in Equation 2.1. In other words, while the specific values of s_i' will be slightly different, their perfect correlation demonstrates that—aside from the consequences pointed out above—one is not an inherently better calculation of item salience than the other. You can see this demonstration in the accompanying code. In summary, if (a) you are *not* interested in what people *don't* say, (b) you're fine with discarding "I don't know" responses, and (c) no one in your sample lists single items, then s_i' can be a useful metric.

Cultural Salience (S')

This particular item salience calculation corresponds to another cultural salience calculation (Robbins et al., 2017). Using the s_i' calculation from Equation 2.3, then, we can use

$$S' = \frac{\sum s_i' + L - 1}{2N - 1}$$

Here, L is the number of individuals who list the item and N is the number of total participants. This particular calculation of cultural salience has the advantage of incorporating the number of lists in which items appear and thus rescales the cultural salience score accordingly. If, for example, out of a sample of 100 individuals and one person listed one particular item, then $S = 0.01$, but $S' = 0.005$.

We can exploit the perfect correlation between s_i and s' and rely on it to compare the values of Smith's S and S'. If we take the free-list data of "what makes a good Tyvan person," we can once again rapidly calculate Smith's S and, with a little coding effort, extract values of L and then S' to make the comparison.[7] Table 2.2 shows the updated output (see code for steps).[8]

In summary, the analyses detailed in this chapter all revolve around the content, accessibility, and "culturalness" of listed items. Frequency analysis of items assesses their relative ubiquity; the more common items are, the more representative they are in a sample. In its various metrics, salience analysis

[7]Let's say you've created an object called SBC, an output of the `SalienceByCode()` function. To calculate L and add it to SBC, run the following: `SBC$L <- SBC$SumSalience/SBC$MeanSalience`. To then calculate corresponding S' values, first define N, the number of participants who engaged in this task (e.g., `N <- 86`),

TABLE 2.2 ■ Truncated ($S > 0.09$) Table Comparing Smith's S and S' on Tyvan "Good Person" Data.

Code	L	$M_{Sal.}$	$\sum_{Sal.}$	S	S'
hard-working	41	0.66	26.99	0.31	0.39
kind	32	0.77	24.58	0.29	0.33
helpful	37	0.65	24.14	0.28	0.35
modest	33	0.71	23.33	0.27	0.32
respectful	35	0.62	21.69	0.25	0.33
honest	28	0.70	19.68	0.23	0.27
intelligent	29	0.65	18.98	0.22	0.27
good conscience	15	0.72	10.80	0.13	0.15
family	24	0.44	10.49	0.12	0.20
friendly/happy	22	0.43	9.41	0.11	0.18
alcohol	13	0.65	8.49	0.10	0.12

Note: L = number of times an item was listed. See supplementary code for full table and subsequent formal comparison (data from Purzycki & Bendixen, 2020).

captures an important cognitive aspect of cultural information, namely, its structure and how it relates to how widespread specific items are. If our questions revolve around what people think, know, and value; how popular items might be; and how easily remembered or accessible particular items are, content analyses are quite useful. In my experience, basic analyses of the content of beliefs can be quite surprising; our own intuitions or the claims of a handful of study participants often don't resemble cultural models drawn from free-lists. These methods give us a much better sense of *what* people actually think and a little glimpse into *how* they think. The next chapter explores the

then run SBC$Sprime <- (SBC$SumSalience + SBC$L - 1)/(2 * N - 1). See the accompanying code for a walk-through.

[8]As it turns out, a standard linear regression shows a very tight correlation between the two indices of the full data set. Here, $S' = -0 + 1.24 * S$. So, if S' were ever to be 1, Smith's S is predicted to be 1.24, an impossibility.

structure of cognitive and cultural models a bit more. In the meantime, let's build a little on the relationships between both frequency and content as well as culture and cognition. The following methods can be of further utility with companion questions and/or for downstream projects.

FURTHER METHODS IN CONTENT ANALYSIS

As mentioned, free-list data can be coupled with other methods or provide information that helps develop more precise, culturally relevant instruments. For instance, once participants have listed their items, we might ask them to list further things about each listed item. We'll revisit these "successive free-lists" (Ryan, Nolan, & Yoder, 2000) in the next chapter. If you wish to design a multi-item scale that requires culturally relevant items for downstream projects, doing a free-list task will help with the generation of appropriate items. This section walks through some follow-up methods enriched by free-list data.

Categorical Bias and Salience

Do people list certain classes of items sooner than others? Is there a relationship between the *order* in which people list items and some biasing process? If individuals were to list, say, foods that people typically eat, are they more likely to list things they like earlier? If they were to list automobiles, are they more likely to list fancier models earlier than more common types? The advertising-marketing fields have long recognized that the more people are exposed to particular brands, the more salient these brands and products are. While cross-culturally variable, this "top-of-mind" effect can predict how favorable individuals find products and the likelihood that they will purchase those items (Hakala, Svensson, & Vincze, 2012; Woodside & Wilson, 1985).

One method allows us to examine list order and bias between binary categories such as foods that one likes versus dislikes or cars deemed fancy versus mundane. Following the mock example from Table 1.3 where we asked people to list animals, we can ask whether or not participants regularly *eat* the animals they listed (the EAT column). If we look at the EAT column data, we see that there are two categories for "yes" and "no" (we could just as easily have coded them as 0s for "no" and 1s for "yes"). We can see whether or not dichotomous categories like this are associated with salience. In other words, are people more likely to list one category sooner? In this example, the question would be if edible items tend to be more salient.

We can use the following equation to obtain individual indices of such a bias, B, (Robbins & Nolan, 1997):

$$B = \frac{Y(Y + 2N + 1) - 2\sum k_Y}{2YN}$$

where Y is the total number of items listed that were subsequently coded with one category (in this example "yes"), N is the total of items that were coded with the other category ("no" in this case), and $\sum k_Y$ is the sum of all order values for the items coded with a Y. In keeping with the animal-eating example, then, participant EX001 has a $B = (2(2 + 2 * 1 + 1) - 2 * 3)/2 * 2 * 1 = 1$. That makes intuitive sense since all of their preferences to eat the listed foods are at the top. Participant EX002 has a $B = (2(2+2 * 2+1)-2 * 6)/2 * 2 * 2 = 0.25$ and EX003's $B = 0$ on the grounds that the only item coded "yes" rests at the end of the list. In summary, then, the larger B is, the more likely the category plays a role in item salience.

Some caution is warranted here. For instance, the equation isn't useful when an individual lists all of one category, as you would have to divide by zero. The authors suggest replacing the zero with a 1. This has some purchase but is far from ideal. If you opt for that, be transparent and clear about what you did. Another thing to keep in mind is that some arbitrary factors can greatly affect an individual's score. If, for example, someone listed 10 items and completely alternated Y and N categorization and started with a Y, then $B = 0.6$. However, if someone listed 10 alternating items and instead started with N, then $B = 0.4$. So, just by virtue of one's starting point, there can be a 0.2 difference in such patterns.

Individuals with values of $B > 0.5$ bias responses toward the Y-type items while those with scores < 0.5 are biasing responses toward the N-type items. Those with $B = 0.5$ are effectively answering by chance; it is effectively a coin flip as to whether or not they bias responses toward one category or another. Of course, answers to follow-up questions like these don't have to be "yes" or "no," but they *do* have to be any dichotomous category. In the logic of cultural salience analysis, summing all B values and dividing by the total sample size will give us a group-level value for B. The supplementary code walks through an example and provides a function to quickly calculate B and analyze it using regression.

A more informative alternative to identifying categorical bias would be using the binary category to predict order number (or vice versa) in a regression framework. These methods would not give us the equivalent of individual-level bias scores like B, but they would allow us to get an estimate of the degree to which a category type corresponds to order number. In some contexts such as comparing group biases in free-listing, we might also use B as a dependent variable in a regression. We'll review regression techniques that could do this in Chapters 5 and 6.

Using Free-Lists to Develop Scales

While the term refers to many things, we'll use **scales** here to refer to multi-item instruments that quantitatively measure one or more aspects of a particular domain. Scales are multiquestion constructs designed to measure some abstract entity. These questions either can have dichotomous responses or can be ordered categories. If you want to measure general happiness, for example, you'd probably design a scale that included multiple items that indicated happiness (e.g., "On a scale from 1 to 7, how happy are you with your material wealth?" and "… how happy are you are your job?"). If a scale has high **reliability** (also known as construct or composite reliability), each item will be highly correlated with each other item (Netemeyer, Bearden, & Sharma, 2003). So in our example, someone who rates low on one question is likely to rate low on the other questions, an individual of middling happiness might have relatively middle-ranged scores for each item, and an exuberantly pleased individual will score highly on all of the questions. Low scale reliability would mean that all of your items are measuring different things or different aspects of the same construct. Some scales are multidimensional, and there are other techniques available to examine those, such as principal components analysis and factor analysis (Bryant & Yarnold, 1995; Dunteman, 1989; Roos & Bauldry, 2021). Here, we'll just focus on unidimensional scales designed to measure the same underlying abstract construct.

In addition to maximizing the **construct validity** of scales, free-list data are also extremely useful for maximizing their **ecological validity**, or relevance to the real world. If you aren't quite sure what kinds of items are locally important or salient, but you want a battery of questions that measure it, it pays to do preliminary ethnographic work to find out, and free-listing is invaluable for this. If, for example, you're interested in understanding health and well-being in a population, you could ask individuals to list the kinds of objects that, if owned, would indicate a successful person (Dressler, Borges, Balieiro, & Dos Santos, 2005). Taking a range of the items listed and then asking other participants to, for instance, rank their importance or perhaps tell you whether or not they own one of the objects could be quite useful in understanding the distribution of wealth and/or prestige in a community. With the free-list data in hand, you can also more confidently demonstrate that you have a more formal sense of what the community members think indicates prestige or wealth, and any scale designed from such data will be immediately relevant. Once your questions are selected and interviews conducted, you can examine scale reliability using a variety of metrics.

To illustrate, we'll use Cronbach's α, a ubiquitously-if-under-critically-used method to assess how multiple items intercorrelate (Cronbach, 1951). To calculate a standardized version of Cronbach's α, you need the number of questions in your scale and the average of all of the correlations between each question. To calculate these correlations, we'll use the common Pearson's *r* correlation coefficient.[9] The equation for calculating Pearson's *r* for two variables, *a* and *b*, is

[9]There are a few ways to calculate correlation coefficients (Indahl, Næs, & Liland, 2018), and Pearson's *r* is just one.

$$r_{ab} = \frac{\Sigma_{i=1}^{n}(a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\Sigma_{i=1}^{n}(a_i - \bar{a})^2}\sqrt{\Sigma_{i=1}^{n}(b_i - \bar{b})^2}}$$

where n is the sample size (i.e., vector length); $\bar{a}$ and $\bar{b}$ are the means of a and b, respectively; and a_i and b_i are corresponding data points of a and b. In `R`, this is all easily done using the `cor.test()` function. The function included in the supplementary code shows the steps and adds some additional statistics (e.g., 95% confidence intervals).

Again, Cronbach's α requires the average Pearson's r between each question, $\bar{r}$, and the number of questions in the scale, q, and uses them to generate an index of internal consistency:

$$\alpha = \frac{q\bar{r}}{1 + (q - 1)\bar{r}}$$

Cronbach's α scores range from 0 to 1, where lower values mean there is less internal consistency across items, and higher values mean that the questions correlate quite well. The values tend to be interpreted much like some exam mark systems; $\alpha \geq 0.9$ are extremely good, $0.89 \geq \alpha \geq 0.80$ are good, $0.79 \geq \alpha \geq 0.70$ are decent, $0.69 \geq \alpha \geq 0.60$ are fair, $0.59 \geq \alpha \geq 0.50$ need some attention, and anything ≤ 0.59 requires an intervention. These thresholds are arbitrary. By itself, Cronbach's α is just a **point estimate**; to get a sense of how likely our calculated α scores are, we'll have to generate some uncertainty around that value (see Chapter 5).

As we'll see, randomly generated data can also produce reasonably high values of α, but their uncertainty intervals might have considerable range, thus suggesting high uncertainty. To capture some uncertainty, we can also calculate confidence intervals with the standard-but-completely-arbitrary range of 95% around this value:

$$\text{lower} = 1 - (1 - \alpha)F(\tfrac{p}{2}, \text{df}_1, \text{df}_2)$$

$$\text{upper} = 1 - (1 - \alpha)F(1 - \tfrac{p}{2}, \text{df}_1, \text{df}_2)$$

We'll discuss uncertainty intervals in more detail in Chapter 5, but such intervals link our data to corresponding probabilities and probability distributions. There are many ways to calculate uncertainty intervals around Cronbach's α (Trinchera, Marie, & Marcoulides, 2018), but one standard way entails calculating an F value, a statistic that specifies the area under an F distribution curve (Feldt, 1965). Among other features, the F distribution only has positive values, which is useful here because Cronbach's α can't be negative.[10]

The inputs we need to get F values are two different degree of freedom indices that give shape to the F curve: df_1 and df_2, where $df_1 = n - 1$ and $df_2 = df_1(q - 1)$. Here too, n is the number of individuals in the sample, and q is the number of questions in your scale. The value p corresponds to the breadth of the interval we choose (or "alpha" level). So, when $p = 0.05$, we're looking for a 95% confidence interval; when $p = 0.10$, we're after a 90% confidence interval; and when we want a 53.8% confidence interval, $p = 0.462$. `R` can quickly find the value of F using the `qf()` function, which is integrated into a step-by-step function for calculating α in `R` Code Box 2.5. The interval is to be interpreted in the standard way; at the 95% level, the range suggests that if the same study were conducted many times, 95% of those intervals would capture the true value of α. Again, see Chapter 5 for further discussion on the importance of quantifying uncertainty.

We can see how Cronbach's α works with the fake dichotomous data in Table 2.3. If you eyeball variables L1 through L5, you can see that they are probably more intercorrelated than the variables R1 through R5. For the former, I made them to closely resemble each other, whereas with the latter, I sampled 0s and 1s randomly, and each had a 50% chance of turning up. Running the function from `R` Code Box 2.5 yields the following: $\alpha_{L1:L5} = 0.93, 95\%CI = [0.84, 0.98]$ and $\alpha_{R1:R5} = 0.71, 95\%CI = [0.27, 0.92]$. So, even the randomly generated data had a Cronbach's α of 0.71, an otherwise "decent" score for intercorrelation. If we inspect the confidence intervals, however, we see that this "decency" is decidedly *indecent* and unacceptable; if we were to rerun the same study again multiple times, the calculated α would hop around quite a bit. The former had an "extremely good" Cronbach's alpha, and the intervals suggest that if we repeated the study 100 times, they might bounce around but in a far higher and narrower range. This is one example of how point estimates can be misleading without accompanying uncertainty intervals.

[10] More technically, the F distribution is a probability distribution of variance ratios (Iversen & Norpoth, 1976). Since Cronbach's α is a ratio of average item covariance to total variance of scores, the F distribution is useful to model its uncertainty. To simulate data using the F distribution in `R`, you need a sample size, n, and the two degrees of freedom: `rf(n, df1, df2)`. See accompanying code.

TABLE 2.3 ■ Simulated Data for Cronbach's α Example.

L1	L1	L3	L4	L5	R1	R2	R3	R4	R5
0	0	0	1	0	1	1	1	1	0
1	0	1	1	0	1	0	0	0	0
1	1	1	1	1	0	1	0	0	1
1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	1	0	0	0	0
1	1	0	1	1	1	0	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	1	1

Note: Variables L1 through L5 are designed to be intercorrelated, and variables R1 through R5 are randomly distributed (Cronbach, 1951).

R CODE BOX 2.5: FUNCTION FOR CRONBACH'S α AND INTERVALS

```
alpha.fun <- function(data, interval){
      k <- ncol(data)
      n <- nrow(data)
      p <- 1 - interval
      cormat <- cor(data)[lower.tri(cor(data))]
      rhat <- mean(cormat)
      alpha <- (k * rhat)/(1 + (k - 1) * rhat)
      df1 <- n - 1
      df2 <- df1 * (k - 1)
      ffl <- qf(p/2, df1 = df1, df2 = df2)
      ffu <- qf(1 - (p/2), df1 = df1, df2 = df2)
      upper <- 1 - (1 - alpha) * ffl
      lower <- 1 - (1 - alpha) * ffu
      return(data.frame(alpha = alpha,
            lower = lower,
            upper = upper))
}
```

The inputs of this function are the set of questions you wish tointercorrelate and the breadth of confidence interval you wish to use (e.g., 0.95 or 0.77). The only lines of code in this function that aren't self-explanatory are the lines that define `cormat` and the lines that define `ffl` and `ffu`. The `cormat` line creates an omnibus correlation coefficient for the lower half of a correlation matrix of your questions. The other two lines use the `qf()` or quantile function in `R` that retrieves the appropriate critical value in an *F* distribution. See the accompanying code for a walk-through example.

As you can use Cronbach's α on either dichotomous or ordered categorical data, you can certainly use it on a presence/absence or order data of free-lists. Given the skew, however, only a few items are likely to intercorrelate strongly, so it pays to work through this carefully if trying to examine aspects of free-list data directly with Cronbach's α. The `psych` package (Revelle, 2022) for `R` rapidly calculates this and a host of other useful statistics related to Cronbach's α with the `alpha()` function. It also performs test-items analysis, which will point to particular questions that might not be ideal in the scale (i.e., ones that don't strongly correlate). Test-items analysis specifically tells you what happens to α if you drop a specific variable.

Cronbach's α has the benefit of being fairly intuitive, and using it certainly beats ignoring the issue of internal reliability. However, there are a few critical assumptions of Cronbach's α (McNeish, 2018; Roos & Bauldry, 2021), and it is not the only metric that speaks to scale reliability (Hattie, 1985; Revelle & Condon, 2019). Some transcend Cronbach's limitations. For example, one important assumption of Cronbach's α is that the answers need to be unidimensional; participants need to answer all questions on the same dimension. Among other features, McDonald's ω can handle multidimensional structures in scale data (McDonald, 2013). The `psych` package will also perform ω.

SUMMARY

This chapter has focused primarily on assessing the content of free-list data. To some degree, it also addressed some aspects of structure; item and cultural salience indices, for instance, provide some information about specific items' positions in lists and the distribution of those items across individual minds. The following chapter digs a bit deeper into conceptual structures and focuses on interitem relationships. Structure analysis involves formal methods that give us metrics to examine just how central or essential things are to the way people think. But it also sheds light on important dimensions of human thought. In other words, structurally analyzing free-list data can give us a sense of *how* and potentially *why* people organize their knowledge the ways they do.