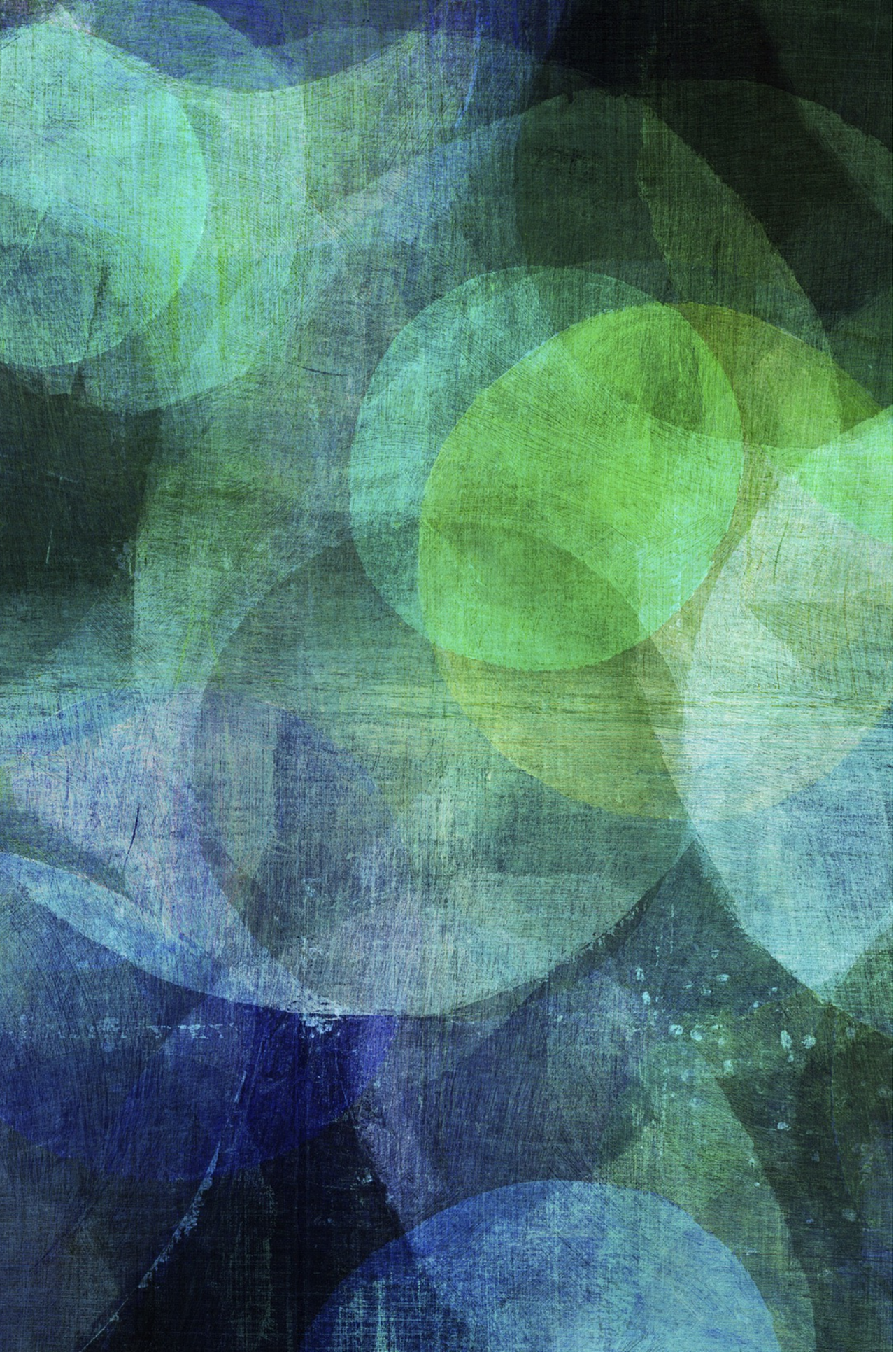
# IV. WHAT WILL YOU MEASURE?

# THIS CHAPTER COVERS:

- What it means to measure
- The need for validity and reliability
- Types of validity
- Types of measurement
- Creation of the code sheet
- The coding process
- Using precoded data: advantages and disadvantages



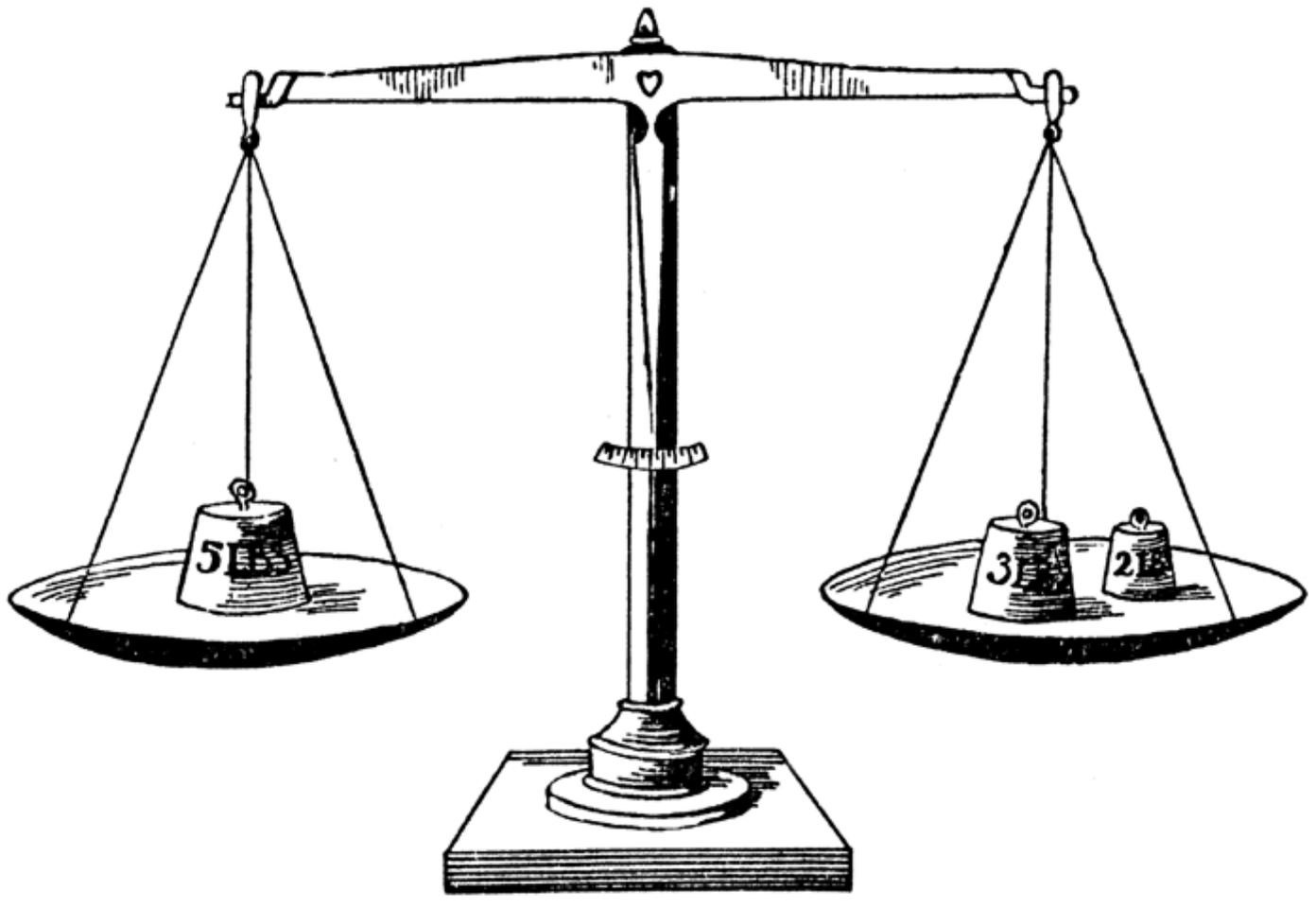
**AS PREVIOUSLY DISCUSSED, VARIABLES ARE QUANTITIES OR QUALITIES THAT VARY ALONG SPECIFIED DIMENSIONS (E.G. SUBSISTENCE TYPES OR PRESENCE OF CURRENCY AS A MEDIUM OF EXCHANGE).**

Cross-cultural researchers may be trying to test theories about the causes and effects of certain variables, how variables are related to one another, or how they are distributed across cultures.

However, before researchers are able to carry out any of these tests, they must develop measures for each of their variables that are both **valid** and **reliable**.

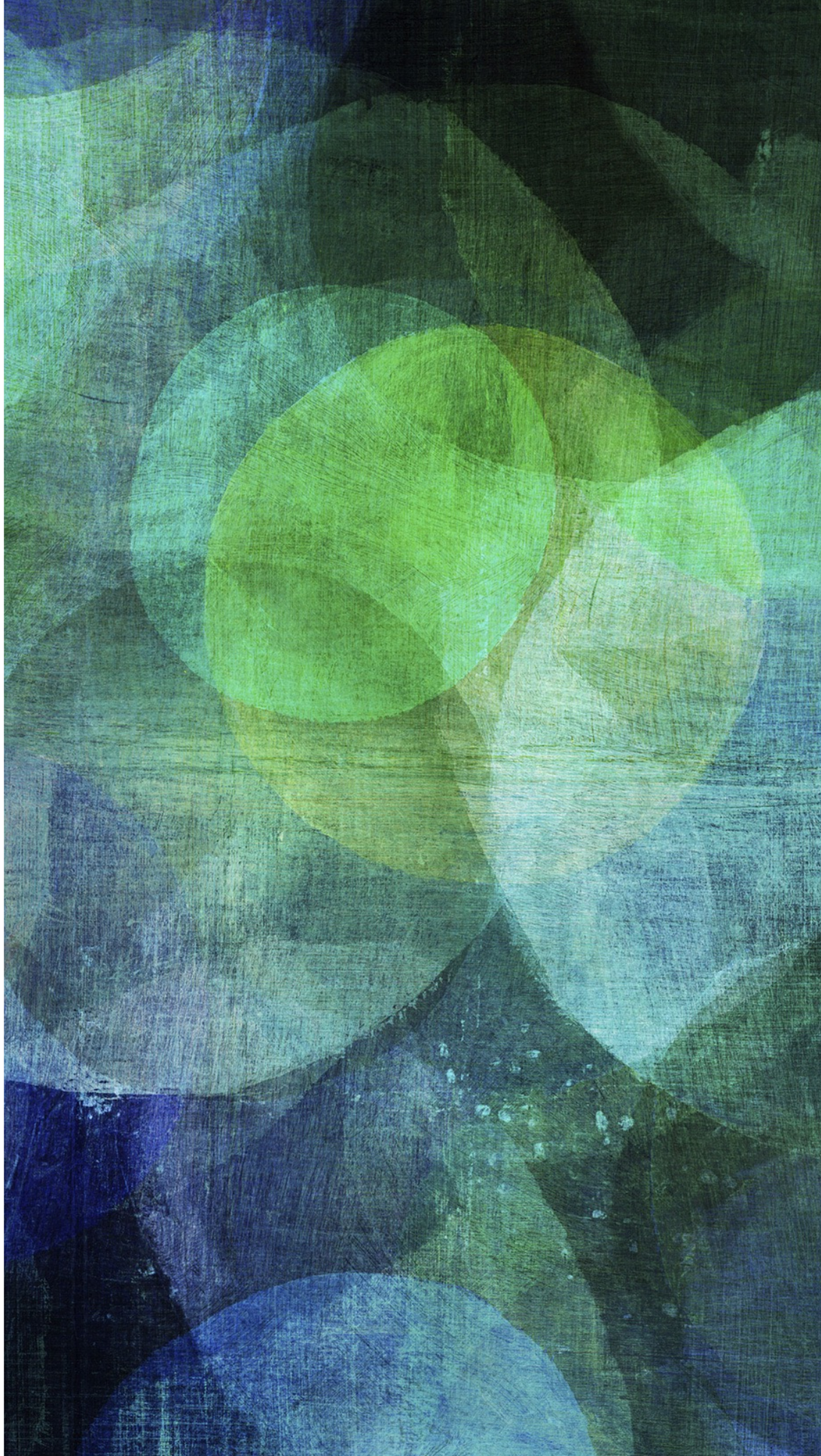
**Before we discuss** what makes a good measure, it is necessary to understand that nothing is measured directly. This is true of **all sciences**— social, physical, and biological.

**Think about measurement** tools that we use in everyday life. When we step on a scale, the scale does not tell us our actual weight; it is a measure of the force that our bodies exert on a spring.




By itself, the number on the scale is meaningless, and the measurement indirect. This number only gains meaning through the comparison of many people's numbers or the examination of how one person's number changes over time. This brings us to another important point about measurement: **measures allow us to compare**.

# WHAT MAKES A GOOD MEASURE?



ALL MEASURES ARE INDIRECT, BUT THIS DOES NOT MEAN THAT ALL MEASURES ARE EQUALLY GOOD. SOME MEASURES WILL HAVE MORE ERROR THAN OTHERS. IT MIGHT BE HELPFUL TO CONCEPTUALIZE THE CONCEPT OF ERROR THROUGH AN EQUATION:

$$m = T + e$$

WHERE "M" IS THE MEASURED SCORE, "T" IS THE "TRUE" SCORE AND "E" IS THE ERROR IN MEASUREMENT.

# ERROR CAUSES OUR MEASUREMENT SCORE TO DIFFER FROM THE "TRUE" SCORE. A GOOD MEASURE REDUCES ERROR BY MAXIMIZING THE FOLLOWING:

- **Validity**: It is essential that all measures measure what they purport to measure. Measures are designed to tap into the theoretical concepts being tested. If a measurement achieves this, we say it has validity. *The more a measure departs from the theoretical concept, the less valid the measure*. This sounds simple, but establishing validity is actually one of the hardest things to do. We will discuss validity in more detail later.

- **Reliability**: A measure is reliable if it has **consistency or stability**--if the measure yields the same results each time the same procedures are used (i.e. results are reproducible). Ways to assess reliability will be discussed in a later section.

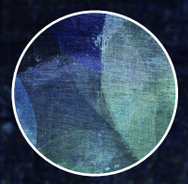
In this section, we will focus on how to minimize error through the *design* of our measures.

# DESIGNING MEASURES USING SECONDARY DATA

MEASURES HAVE TO BE SPECIFIED FOR EACH VARIABLE IN YOUR HYPOTHESIS. RECALL THAT WE HAVE BOTH THEORETICAL AND OPERATIONAL HYPOTHESES.

**Theoretical hypothesis**: The theoretical variables and their presumed relationship are stated in their more abstract form (e.g the theoretical construct A will be related to the theoretical construct B).

**Operational hypothesis:** Takes the same form as the theoretical hypothesis but includes references to the actual measures used (e.g. the measure A will be related to the measure B). We will discuss how to design these scales for measurement later.



# MINIMIZING ERROR WHEN OPERATIONALIZING VARIABLES

# HERE WE WILL FOCUS ON HOW TO MINIMIZE ERROR THAT RESULTS FROM A DISCREPANCY BETWEEN THE THEORETICAL CONCEPT AND THE DESIGNED MEASURE.

If the measure matches the theoretical concept well, we say that the measure is **valid.**

But how can we establish validity when nothing can be measured directly?

There are several types of validity to consider:

# FACE VALIDITY

When a measure has high face validity, it is considered a relatively straightforward reflection of the theoretical variable. There is little need to justify the use of a measure with high face validity.

Consider a researcher wishing to study the distribution of meat taboos cross-culturally. The researcher may look for explicit ethnographic reports of food taboos in different societies. This measure would have high face validity. A measure demonstrating low face validity might involve more inference; for example, it may use a compilation of ethnographically-reported customary diet items to then infer that underutilized but edible species have a taboo upon their consumption.

# CRITERION VALIDITY

A measure that has criterion validity is highly correlated with an existing and generally accepted measure (the criterion). This type of validity is more difficult to use in cross-cultural research but can be useful when collecting data in the field.

# CONTENT VALIDITY

This type of validity refers to the degree to which “a specified domain of content is sampled” (Nunnally 1978: 91).

A measure that covers several aspects of a concept is better than one that focuses on too few.

This type of validity is useful when measuring more abstract concepts. For example, when measuring cultural complexity, a measure might be designed that looks at different domains of complexity, such as political heirarchy, class stratification, size of communities, and occupational specialization. However, sampling more aspects does not always improve the measure. An aspect that does not reflect the concept of interest or is not associated with the other aspects should be excluded. Statistical tests can be done to ensure that a set of cultural aspects belong together.

# CONVERGENT VALIDITY

When a measure correlates highly with other independent measures of the same concept, it has convergent validity. If several different scales are correlated with one another, you can choose the scale that best fits your theoretical concept or that is easiest to apply.

# TYPES OF MEASUREMENT

# NOMINAL

Nominal variables (also known as categorical variables) are the simplest form of measurement and involve putting variables or cases into discrete sets.

This type of measurement could be used when looking at subsistence type, for example.

Cases could be sorted into hunter-gatherers, pastoralists, intensive agriculturalists, etc.

# ORDINAL

Ordinal measurement adds a qualifier of “more” or “less” to a scale. Ordinal scales can be as simple as “frequent,” “occasional,” and “rare.”

Subsistence type can also be measured on this type of scale, depending on the information of interest.

We could devise a measurement that would categorize societies from less to more dependence on hunting, gathering, and fishing. This would provide us with the relative, but not exact, dependence on the type of subsistence.

# INTERVAL & RATIO

Interval and ratio measurements allow description of the amount of difference between two cases.

Points on these scales (e.g. temperature) are **equidistant**, and therefore the distances between points hold meaning.

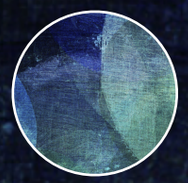
Population density can be measured on a ratio scale. Technically, population density is a ratio scale because it has a “true zero” point (zero people per sq. mile), whereas an interval scale does not. With a zero point you can say that one society having 20 people per square mile has twice the population density of one with 10 people per square mile. However, there is little statistical  difference between interval and ratio scales.

**Ratio scales** are considered the "highest" levels of measurement, followed by interval, then ordinal, and lastly nominal. Statistics are generally more powerful when higher levels of measurement are used.

In some cases, we can transform a measurement to a higher type of scale. For example, if we are interested in the occurrence of polygyny, we can use all three types of scales:

- **Nominal**: Typical form of marriage (e.g. polygyny, monogamy, polyandry).
- **Ordinal**: Frequency of polygynous households (rare, sometimes, frequent, very frequent).
- **Ratio**: Percentage of households that are polygynous.

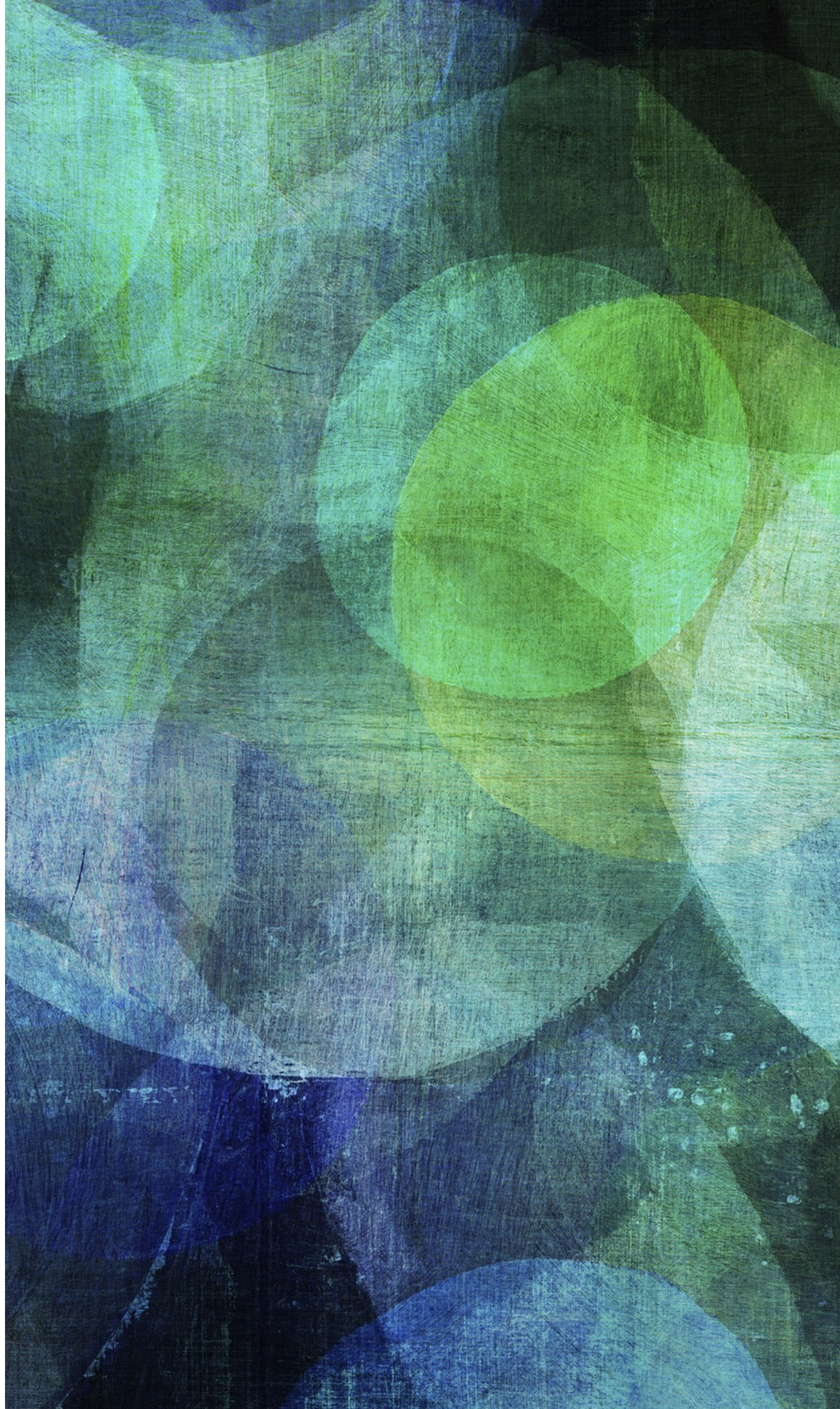
Although it is preferable to use the highest level of measurement, you must be careful not to use a higher level of measurement than warranted by the data. Higher levels of measurement are frequently inapplicable to cross-cultural research because ethnographers often do not provide quantitative information.



# OTHER IMPORTANT NOTES ABOUT CODING AND MEASUREMENT DESIGN

**Once you have derived** your hypothesis or hypotheses, specified the measures and operational procedures for all of the variables you want to test, pretested, and selected a sample of cases to study, you are ready to collect data.

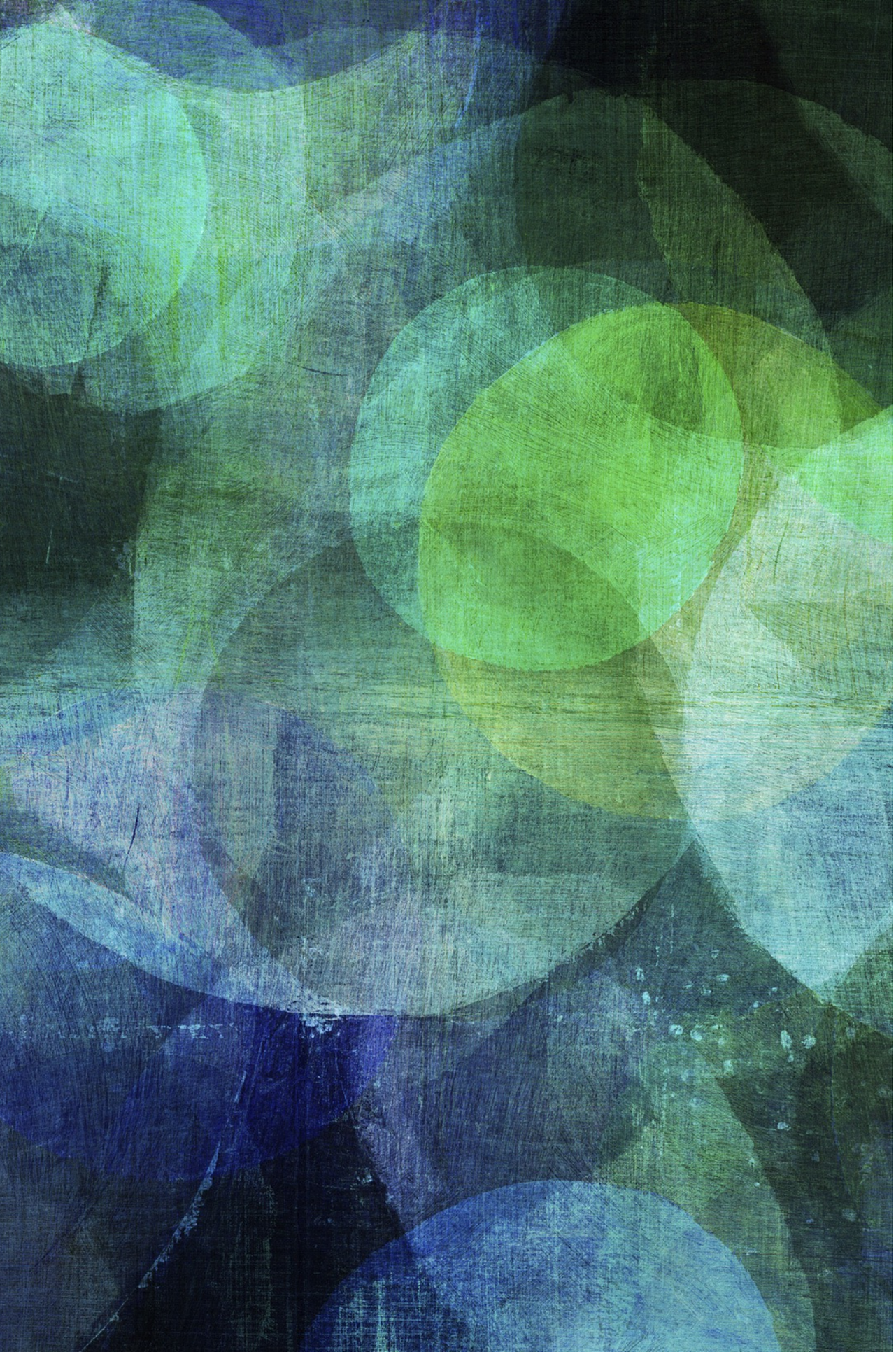
Collecting data for a cross-cultural study using secondary data involves a process called **coding**.



# WHAT IS CODING?

- The term "coding" in cross-cultural research typically refers to a way of **transforming qualitative data** (usually from ethnography) into data that can be **tested quantitatively**.
- This process involves creating measures to represent the variables you want to statistically test (discussed in the "What will you measure?" section), and having coders collect the relevant ethnographic information and rate each case in your sample according to those measures.

# THE CODING PROCESS



## OFTEN, THE RESEARCHER WHO CREATED A MEASUREMENT DOES NOT PERSONALLY DECIDE WHERE EACH CASE FALLS ON THE SCALE.

Instead, that process becomes the responsibility of people who are designated as "**coders**." The coders rate each case in the researcher's sample based on the "**code sheet**" that is provided to them by the researcher. This code sheet includes the scales for each variable to be measured, as well as explicit instructions on how to rate each case on the scale.

# WHO SHOULD CODE THE DATA?

- It's always better to have **more than one individual** code data. This minimizes the possibility of bias affecting the results and allows for reliability checks between coders.
- Two well-trained coders are usually sufficient. Ideally, training should be conducted on cases that will not be included in the final sample.
- It is preferable for at least one coder to be unaware of the hypothesis or hypotheses being tested. These coders are referred to as **"naive" coders**. Having different sets of coders code the independent and dependent variables will help insure that your coders remain unaware of the purpose of the study.
- If only one person is coding and that person knows the hypothesis, it is important to provide explicit references for coding decisions (such as the relevant passage upon which decisions were made).

# KEEPING THIS INFORMATION IN MIND, THERE ARE A FEW OTHER IMPORTANT POINTS TO CONSIDER WHEN CREATING A SCALE:

- **You must provide** your coders with a way to indicate either a lack of information or contradictory information. It is a good idea to use discontinuous numbers (such as 88 or 99) for these options. It will help you remember that these numbers should not be used in your analysis.

- **It is important** that your coders know how to infer whether a trait is rare or absent. For example, if marriage and family life are described in detail but polygyny is not mentioned, the coder might be allowed to infer that polygyny is not present. However, if marriage and family life are not described in any detail and polygyny is not mentioned, the coder should mark "don't know." Again, make these instructions as clear and explicit as possible.

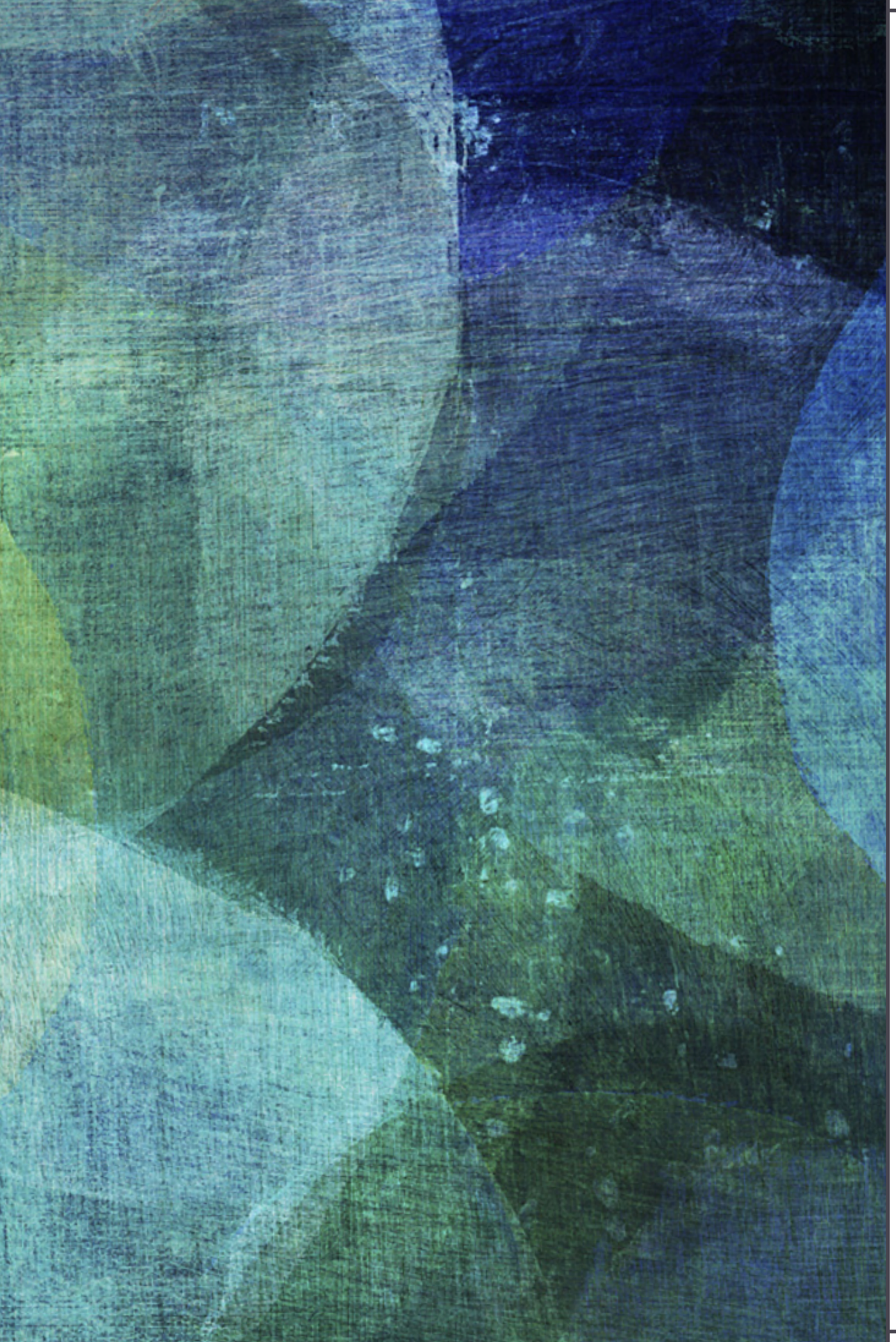
- **You need to specify** where to find the appropriate information for each case. You should inform coders of the appropriate ethnographies to use as well as the appropriate time and place foci that codes should pertain to.

    - If you wish to allow your coder to deviate from the indicated time and place focus, clearly indicate that a deviation is allowed. A separate code can be added to indicate a deviation, allowing you to test whether or not the deviation affected the results.

- **Pretesting your scales is essential.** Creating scales and coding is an iterative process. It is likely that you will find flaws in your measures when you test them on cases. For example, the points on the scale may be too ambiguous or you may need to increase or decrease the number of points on the scale when distinctions are too difficult to code reliably. It is better to detect and resolve these problems before you begin coding in earnest. Pretesting should also be done with coders who were not involved in the creation of the scale. This will ensure that the researcher spells out their assumptions of how the measurement should operate.

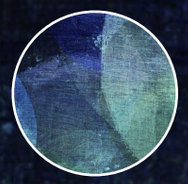
# FORM YOUR CODE SHEET

YOUR CODE SHEET CAN BE PAPER OR ELECTRONIC AND SHOULD ALWAYS ALLOW SUFFICIENT ROOM FOR NOTES AND COMMENTS. WHEN USING EITHER FORMAT, IT IS VERY IMPORTANT TO BACK UP YOUR DATA.



# THERE ARE THREE MAJOR COMPONENTS IN A CODE SHEET:

- Identifying information
- Operational procedures for measuring the variable
- A place for indicating the coding decisions and pertinent information



# USING EXISTING MEASURES

# IN SOME CASES, THE VARIABLE YOU ARE INTERESTED IN WILL HAVE ALREADY BEEN STUDIED AND CODED BY ANOTHER RESEARCHER.

If you are using the same sample of cases as the previous researcher, using existing measures can be convenient. However, there are some important points to consider before you make this decision.

# ADVANTAGES

- Using another researcher's codes for variables can help to minimize the unconscious biasing of the results. The more familiar you or your coders are with the sample of cases or the hypotheses being tested, the more likely you or your coders will be to rate cases in a way that is influenced by knowledge of the project or expectations of its results.
- Using existing codes can greatly reduce the amount of time and money spent on coding.
- It is recommended that you code a portion of another researcher's cases and verify that you are comfortable with their coding decisions.

# DISADVANTAGES

- You will need to make sure that your time and place foci, or allowance for deviation from these foci, match those of the code that you wish to use.
- The existing codes may not actually measure the concept that you want to be measuring. As we discussed above, using a measure that does not match your theoretical concept can be a serious source of error.

**Note:** If you use another researcher's measure, your operational definition will be a description of the original author's code, where to find the coded information, and how you used the existing scale or modeled it. You can refer the reader to the published definitions, if available, to avoid repeating all of the definitions in your study.

# SUMMARY

- To test hypotheses, **all variables need to have clearly specified measures**.
- **All measures are indirect**, but a research needs to strive for measures with **high validity and reliability.**
- The most important types of validity in cross-cultural research using secondary data are: **face validity, content validity, and convergent validity**
- **The four main types of measurement** are nominal, ordinal, interval and ratio scales.
- If possible, a researcher should try to have at least one coder who does not know the hypothesis. Two coders are ideal.
- Whether you use a code sheet or enter decisions directly into a database, it is critical that the coders have identifying information, the sources and procedures for finding data, the operational measures for every variable and scale score, and places to enter decisions and notes.
- There are both advantages and disadvantages to using precoded data; it is **highly recommended that at least a portion of the cases with precoded data be recoded** to better understand the variable.

# REFERENCES

Nunnally, Jum C. 1978. Psychometric Theory. 2nd ed. McGraw-Hill.

# CITATION AND TERMS OF USE

**Using this course:** The material in this course is intended for educational purposes only--either by individuals for personal use or for classroom use by instructors with appropriate attribution. For any commercial use or other uses please contact HRAF. (Email: <hraf@yale.edu>).

**This chapter is from:** Carol R. Ember. 2016. Introducing Cross-Cultural Research. Human Relations Area Files. <http://hraf@yale.edu/ccc/>