The Routledge Reviewer's Guide to Mixed Methods Analysis

Edited by
Anthony J. Onwuegbuzie and R. Burke Johnson

First published 2021

ISBN: 978-1-138-30526-7 (hbk)

ISBN: 978-1-138-30527-4 (pbk)

ISBN: 978-0-203-72943-4 (ebk)

Chapter 24

The Case Comparison Table: A Joint Display for Constructing and Sorting Simple Tables as Mixed Analysis

Judith Schoonenboom and R. Burke Johnson

(CC BY-NC-ND 4.0)

DOI: 10.4324/9780203729434-24

The Case Comparison Table

A Joint Display for Constructing and Sorting Simple Tables as Mixed Analysis

Judith Schoonenboom and R. Burke Johnson

1 Definition of a Case Comparison Table

This chapter introduces a specific type of joint display table called the case comparison table. Construction of a case comparison table is a form of data analysis in mixed methods research (henceforth: *mixed analysis*). We define a case comparison table as a simple table, in which each row contains information on one case of the research study (e.g., one student, one hospital, or one legal case), with part of the information coming from qualitative data sources, such as in-depth interview data, and another part coming from quantitative data sources, such as structured questionnaire data (see Table 24.1 and Table 24.2 in this chapter). In the first section of our chapter, we describe the case comparison table, and we clarify the five components of this chapter's title: *simple tables*, *analysis*, *mixed analysis*, *constructing*, and *sorting*.

Joint Displays as the Parent Category of Case Comparison Tables

Our case comparison table forms a subcategory of visualizations that are used in mixed methods research, commonly called *joint displays* (Guetterman, Creswell, & Kuckartz, 2015; Guetterman, Fetter, & Creswell, 2015). Joint display is an umbrella term for visualizations in which quantitative and qualitative data and/or results are presented together for comparative and integration purposes. Following Fetter, Curry, and Creswell (2013, p. 2143), we define a joint display as “a visual means to draw out new insights beyond the information gained from the separate quantitative and qualitative results.” Joint displays need not be tables: Examples of joint displays other than tables are the mixed flowchart and the mixed concept map.

In mixed methods research, broadly defined, joint displays have various, partly overlapping, purposes. These purposes include, but are not limited to:

- Display meta-inferences from qualitative and quantitative findings
- Juxtapose quantitative results and qualitative findings (Plano Clark & Sanders, 2015, p. 183)
- Summarize links among theory, quantitative and qualitative data, and analysis (Bazeley, 2018)

- Show quantitative and qualitative findings (columns) for multiple research questions (rows) (Schoonenboom & Johnson, 2017)
- Show quantitative and qualitative findings (columns) for multiple independent/causal/predictor variables (rows) (Schoonenboom & Johnson, 2017)
- Show quantitative statistical information per qualitative theme (Guetterman, Fetter, et al., 2015)
- Interrelate specific quantitative results and qualitative findings (Guetterman, Fetter, et al., 2015)

2 When Use of Case Comparison Tables is Appropriate in Mixed Methods Research: The Use of Case Comparison Tables in the Stage of Analysis

Tables, of which the case comparison table is a subcategory, can be used in various stages of mono-method and mixed methods research (Bazeley, 2018; Miles & Huberman, 1994; Miles, Huberman, & Saldaña, 2020; Plano Clark & Sanders, 2015). Tables are most frequently used for presenting research results, for example, the results of statistical analysis. Tables also can be used for the research design to show the flow of the research process (Bazeley, 2018). They can be used in research planning, and to show when the researcher intends to do what. And they can be used in the development of the study's conceptual framework (Maxwell, 2013).

Tables that combine or integrate information from qualitative and quantitative sources are called *mixed tables*. Mixed tables can be used for many purposes, including:

- Show information from various qualitative and quantitative sources
- Note patterns, themes; make contrasts, comparisons; cluster; and count (early in the process). (Miles et al., 2020, p. 117)
- Follow up on surprises, triangulate, make if-then tests, and check out rival explanations (later in the analysis process; Miles et al., 2020, p. 117)

TABLE 24.1

Reconstructed Simple Case Comparison Table based on Lee and Greene (2007)

Descriptive data				Collected data			Analytic data			
ID	Discipline	S	R	C	GPA	Quotation	LP	PA	CS	EG
0607	Humanities	Q	S	4	4,00	I do not have any language problems.	0	0	0	1
0609	technology	Q	S	3	4,00	The professor is old so that his pronunciation is not clear. He usually handed out important contents and I could understand it by reading the textbook.	1	0	1	2
1315	business	I	S	3	4,00	My biggest problem is related with speaking in English. I will get a good grade because my mathematical background is strong.	1	0	1	2
2020	Technology	I	S	3	4,00	I understand 80% of the lectures. Careful reading complements 20% of lack of understanding.	1	0	1	2
0620	Science	Q	S	2	4,00	I understand only 60–70% of the lectures. It has made my scores less than my expectation.	1	1	0	5
1310	business	Q	S	3	3,89	I want to participate and argue the subject. But I cannot find myself enough. This does not affect my grade.	1	0	0	3
0624	Technology	I	S	2	3,80	It is easy to understand the lectures and participate in class discussions. The instructor speaks slowly.	0	0	0	1
2036	Humanities	I	S	3	3,57	The major problem is speaking. I spoke once or twice during the whole semester.	1	0	0	4
0610	technology	I	S	3	3,53	Listening is a problem. Lack of cultural knowledge interferes with understanding the concept.	1	1	0	5
0605	humanities	I	S	2	3,50	Because of my poor listening, I am struggling with catching up with my content courses.	1	0	0	5
0603	Technology	Q	S	3	3,39	I still have some problem in speaking. This difficulty doesn't affect my ability to do well in all the courses I take.	1	0	0	3
2037	business	I	S	2	3,22	Lack of knowledge about idiomatic expressions prevents me from understanding questions on the homework assignment.	1	1	0	5
2025	technology	I	S	3	3,18	I was not able to finish all the reading assignments every week.	1	0	0	4
2031	humanities	Q	S	3	3,11	I have problems with my hearing and how to speak correctly.	1	0	0	4
0608	technology	I	S	4	2,89	I understand almost 100% of the lectures. I understand the professor completely. [...] Grades are not important. That is the least of my worries. After we graduate with a Ph.D., people are not going to ask "Did you get an A or a B in that class?" I might get a B in this course. As long as I am learning, it is OK with me.	0	0	0	1

Table notes:

Descriptive data

ID = Student ID; Discipline = Academic/course discipline; S = Quotation source (Q = questionnaire, I = interview); R = Quotation source role (S = student, F = faculty); note that we simplified Table 1 by removing quotations by faculty (there is no quotation with quotation source value 'F').

Collected data

C = CEEPT score ranging from 2 (worst) to 4 (best): (2=placement in the ESL One class, 3=placement in the ESL Two class, and 4=exempt from any ESL classes (Lee and Greene 2007, p. 371).

GPA = Student first-semester GPA (ranging from 0 (worst) to 4 (best)).

Quotation = Utterance about one student.

Analytic data created during analysis

LP = Student's Language Problems: student indicates having (1) or not having (0) language problems.

PA = Problems affect Achievement: student mentions (1) or does not mention (0) that language problems affect his or her achievement.

CS = Student's Compensation Strategies: student mentions (1) or does not mention (0) compensation strategies;

EG = Effect on student's Grade: how language problems affect a student's achievement: (1) student does not have language problems; (2) student mentions some ability or behavior that compensates for their language problems (3) language problems do not affect their achievement; (4) student indicates language problem, but not their effect on their achievement (5) language problems have a negative effect on achievement.

Sorting order Table sorted first descending by GPA (G), and then descending by CEEPT Score (C).

Highlighting Highlighted are students with CEEPT Score 3 and GPA 4 (see explanation in text).

TABLE 24.2

Reconstructed Simple Case Comparison Table from Shaw et al. (2013)

Descr. data	Collected data		Calculated data		Analytic data					
Practice	CRC-B (%)	CRC-12 (%)	Diff. (%)	High perf.	Team Structure	Leadership	Engagement	Psychological Safety	Intra- communication	Inter- communication
P22	47	71	24	Y	3. Strong	1. Weak	2. Moderate	2. Moderate	2. Moderate	1. Weak
P7	53	73	20	N	3. Strong	1. Weak	2. Moderate	1. Weak	2. Moderate	1. Weak
P21	38	56	18	N	—	—	—	—	—	—
P15	50	67	17	N	2. Moderate	1. Weak	2. Moderate	1. Weak	2. Moderate	1. Weak
P2	14	30	16	Y	3. Strong	2. Moderate	3. Strong	3. Strong	3. Strong	2. Moderate
P8	37	52	15	Y	3. Strong	2. Moderate	3. Strong	2. Moderate	2. Moderate	1. Weak
P11	54	66	12	N	1. Weak	1. Weak	2. Moderate	1. Weak	2. Moderate	NA
P16	43	48	5	Y	3. Strong	3. Strong	3. Strong	3. Strong	3. Strong	1. Weak
P23	93	86	−7	Y	3. Strong	2. Moderate	3. Strong	3. Strong	3. Strong	1. Weak
P19	52	44	−8	Y	3. Strong	3. Strong	3. Strong	3. Strong	3. Strong	NA
P17	41	10	−31	N	—	—	—	—	—	—
P10	71	33	−38	Y	3. Strong	2. Moderate	2. Moderate	2. Moderate	3. Strong	3. Strong

Table notes:

Descriptive data. Practice = ID of medical practice.

Collected data: CRC-B = Screening Rates Baseline (%); CRC-12 = Screening Rates 12-Month Follow-up (%)

Analytic data (summary of qualitative ordinal data) Team; Structure; Leadership; Engagement; Psychological Safety; Intracommunication; Intercommunication

Calculated data (analytic data calculated from other data within this table): High perf. = high performing practices (=overall judgment based on the analytic data); Diff. = difference between follow-up and baseline ratings

Grey cells: values crucial in our analysis.

Cases sorted on Diff.: difference between follow-up and baseline ratings.

- Compare qualitative and quantitative results to see whether they are congruent or not
- Present qualitative and quantitative information side by side, using different theoretical lenses (Guetterman, Fetter, et al., 2015)
- Embed qualitative information into quantitative findings (Guetterman, Fetter, et al., 2015)
- Display integration of qualitative and quantitative results (Guetterman, Creswell, et al., 2015; Johnson & Christensen, 2020)
- Indicate difference/similarity of quantitative and qualitative results (Guetterman, Creswell, et al., 2015; Johnson & Christensen, 2020)

In this chapter, we focus on the use of case comparison tables during analysis. This means that we exclude two other uses of mixed tables: (a) mixed tables that are not case comparison tables, in which the rows are not cases, but are, for example, the research questions or the themes; and (b) the use of mixed tables for purposes other than analysis—for example, for presenting final results. We use a broad definition of analysis, including not only analysis of the data, but also integration of findings (Guetterman, Fetter, et al., 2015) and drawing conclusions.

In addition, our focus is on modern, digital, sortable tables as we know them from word processors and spreadsheets. This focus on digital, sortable tables is not new. The use of such tables in qualitative data analysis has been discussed by Janesick (2016), La Pelle (2004), Meyer and Avery (2009),

Miles and Huberman (1994), Miles et al. (2020), Ruona (2005), Ryan (2004), and in mixed analysis by Combs and Onwuegbuzie (2010). Our chapter builds on and extends these previous publications, while putting a larger emphasis on the process of sorting and resorting.

3 Technical Outline of Case Comparison Tables for Mixed Methods Research

Case Comparison Tables as Simple Tables

Our case comparison table belongs to the class of simple tables. We define a simple table as a two-dimensional grid of single rows and single columns, in which each cell contains information based on the combination or intersection of the rows and columns. This piece of information can be as diverse as one patient's weight, one student's score on a test, one quotation by an interviewed teacher, or a summary of one legal case. Note that *simple* does not necessarily mean *easy*. As we will see, a case comparison table can be very large, in which case, we need specific techniques to manage and understand our data.

Conversely, a table can be complex in three distinct ways: (a) it contains more than one piece of information per cell, such as a combination of height and weight; (b) it contains rows or columns that split (we could, for example, turn Table 24.2 into a complex table by creating one column "Screening rates," which splits into "Baseline" and "After 12 months"); and/or (c) it includes three or more dimensions, in which case it is called

a multidimensional matrix. In this chapter, we will argue that, in the stage of analysis, a simple two-dimensional table is usually preferable to a table or matrix with a complex layout (La Pelle, 2004).

Constructing and Sorting Case Comparison Tables

Case comparison tables will change during analysis, not unlike other joint displays (Miles et al., 2020, p. 113): They are constructed and reconstructed. This analytical and iterative process is often best supported by using simple, sortable tables. Conversely, mixed tables used for presenting results in publications might need to be more complex. This applies to Miles et al.'s (2020, p. 111) meta-matrices and to almost all mixed tables in the literature on joint displays (Bazeley, 2018; Guetterman, Creswell, et al., 2015; Guetterman, Fetters, et al., 2015; Miles & Huberman, 1994; Miles et al., 2020; Plano Clark & Sanders, 2015). Different from Miles et al. (2020, p. 118), our claim is that the, often complex, mixed table that one needs for presenting results is typically not the same as the simple case comparison table that one needs for data analysis.

Throughout this chapter, *sorting* means that the table as a whole is sorted on the basis of the values of one or more sorting columns.

4 Empirical Demonstration of Case Comparison Tables for Mixed Methods Research: Two Real-Life Examples

Before we describe how to construct case comparison tables in detail, we present two case comparison tables that we constructed from two published mixed methods research studies: Lee and Greene (2007) and Shaw et al. (2013). We constructed our simple case comparison tables from complex tables and other information that was presented in the original publications.

Example 1: Lee and Greene (2007)

Lee and Greene's (2007) research purpose was to understand the predictive value of scores on an English as a Second Language (ESL) test—the CEEPT—of international graduate students at a large public university in the United States on their academic performance (measured as grade point average [GPA]) and their language difficulties in courses during their first semester. In addition, 100 students undertook a self-assessment, and 55 faculty evaluated their students. Interviews were conducted with 20 students and 10 faculty members.

We started our analysis from Lee and Greene's (2007, p. 378) observation that students' utterances contained the following two dimensions: the extent to which they found English demands for a particular course difficult and their perceived effect of English demands on their course performance. Similar to Lee and Greene (2007), our research question was whether we could establish a link between students' measured CEEPT and GPA and what students reported about the demands and the effect on their performance.

Our case comparison table included quotations and utterances by students and faculty members that were provided by Lee and Greene (2007); we used a subset of the quotations on which these authors based their analysis. Most of our information came from Table 5 in Lee and Greene (2007), an information-rich and highly complex table. We created a simple case comparison table, in which students are the cases (Table 24.1). We added two additional columns with analytic data, which we explain later.

We created the case comparison table in Table 24.1 by using one row for each person's quotation(s), and by providing the quantitative and any other qualitative information that we had about the case in which the quotation was embedded (i.e., the student to which the quotation referred). Table 24.1 is a case comparison table because each row in Table 24.1 contains information about one case only. Furthermore, it is a simple table because each cell contains only one piece of information and because it does not contain columns that split. For the purpose of demonstration, we removed utterances by faculty members from our case comparison table. As a result, Table 24.1 contains only quotations by students about themselves.

A case comparison table contains at least three types of data. *Descriptive data* provide descriptive characteristics of the case or the quotation of that row, in this case, student ID and discipline of the student (the case) and information on the source of the quotation, namely, its form (questionnaire or interview) and the type of person who uttered the quotation (student or faculty). *Collected data* provide data on something that the case or someone else did, which was relevant to the research question, in this case the student's CEEPT score, his or her GPA, and the quotation that was uttered. *Analytic data* provide values of analytic categories of the researcher that are used to classify and group the quotations and/or students. Analytic categories are developed during analysis.

Our analysis is an alternation between sorting and resorting the case comparison table using various columns as sorting columns, reflecting on the results, adding new analytic categories, and thus columns, which capture the differences resulting from the reflection, sorting, and resorting using various columns as sorting columns, and so forth (Figure 24.1). Next, in Figure 24.1 and the following text, we describe our analysis of Example 1 in more detail. Our analytical process is now described.

Phase 1 Analysis

- *We started by reflecting on a finding in Lee and Greene (2007):* Lee and Greene (2007, p. 378) observed that one half of the students with CEEPT 2, the worst score, found the English demands difficult, nine of 11 with CEEPT 3 found the demands okay, and all of the students with CEEPT 4 found the English demands easy. Then, we performed the following analytical actions:
- *We added an analytic category that captured the reflection result:* We added a new analytic category: LP: mentioning language problems (1 = mentioned by the student, 0 = not mentioned);

FIGURE 24.1 Heuristic for the analysis of case comparison tables.

- *We sorted and resorted:* We sorted our case comparison table various times, using various columns, including the new LP column, to determine whether new patterns would emerge.
- *We reflected on the results:* When we reflected on the quotations that were available to us (which are a subset of what Lee and Greene [2007] used), we noticed that, except for one outlier, students with a CEEPT 2 or 3 reported language problems, whereas none of the students with a CEEPT 4 reported language problems, a finding that differs slightly from Lee and Greene (2007). Thus, we found a discrepancy, which might later turn out to be relevant or might eventually be solved through further analysis.

Phase 2 Analysis

- *In a second round, we started by reflecting on another finding reported by Lee and Greene (2007):* According to Lee and Greene (2007, p. 378), one half of the students with CEEPT 2 revealed that language problems did not affect their grade, whereas the other half reported a negative effect. None of the students with CEEPT 4 reported a negative effect, whereas 2/3 of the students with CEEPT 3 reported a negative effect.
- *We performed a second round of adding analytic categories that captured the reflection result:* We added an analytic category: PA: Language problems affect the student's grades: Student mentions (1) or does not mention (0) that language problems affect achievement.
- *We performed a second round of sorting and resorting:* We sorted our case comparison table several times, using different columns, including the new PA column, to see whether new patterns would emerge.
- *We reflected on our new results:* We discovered, that of those other students who had mentioned language problems, some mentioned that it affected their

grade, whereas others did not, which is in line with Lee and Greene's findings.

Phase 3 Analysis

- *We performed a third round by sorting and resorting:* To search for links among measured CEEPT, GPA, and students' utterances, we first sorted our comparison table on GPA, CEEPT, and combinations of the two.
- *We reflected on the new results:* When we sorted the table first by GPA and then by CEEPT, which is the sorting order displayed in Table 24.1, we discovered an interesting pattern: All of the students with a moderately high CEEPT (3) and a high GPA (4) indicated some sort of compensation strategy for their language problems. These students are highlighted in Table 24.1. Student 0609 compensated his or her misunderstanding of the professor by reading the textbook, Student 1315 compensated problems with speaking English by a strong mathematical background, and Student 2020 compensated by careful reading.
- *We added one new analytic category that captured the reflection result:* We used the analytic category CS: mentioning compensation strategies: Student mentions (1) or does not mention (0) compensation strategies.

Phase 4 Analysis

- *We performed a fourth round of sorting and resorting:* We sorted our case comparison table various times, using various columns, including the new CS column, to determine whether new patterns would emerge.
- *We reflected on the new results:* From sorting and resorting, it became clear, that the three students with a CEEPT of 3 and a GPA of 4 were the only students mentioning compensation strategies: they are the only students with value 1 in the CS (compensation strategies) column. These finding might explain the lack of a unique effect of CEEPT on GPA: successful students have strategies to compensate for their language problems. This relevant observation is not discussed by Lee and Greene (2007).

Phase 5 Analysis

- *We performed another round of reflection on the results:* Using the columns LP, PA, and CS, we were able to divide the students into five non-overlapping categories: (a) students do not have language problems; (b) students mention some ability or behavior that compensates for their language problems (c) language problems do not affect their achievement;

examples was *hypothesis and theory development*. Through a combination of adding analytic category columns and sorting and resorting the table, possible explanations for the main finding in the studies became visible: students' compensation strategies in Example 1 and a possible resistance to change of strong teams in Example 2. Note that these inductively developed hypotheses are not explanations yet, but tentative findings, and starting points for further inquiry.

Category development involves the use of one analytic category column (often a quantitative column) to categorize the contents of another column (often qualitative data: e.g., quotations). Continuously sorting columns by such an emergent analytic category makes visible the merits and flaws of the category-under-construction and leads to continuous updates of the category and its values (e.g., the development of the category Effect on student's Grade (EG) in Example 1 earlier).

Profile Development

Through continuously sorting and resorting of various columns, the researcher might discover that a specific value in one column often co-occurs with a specific value in one or more other columns. Cases with such a specific value combination are said to have a specific profile.

Causal analysis proceeds by putting columns in a chronological order. The researcher may then discover that a value in one column that describes something that happens to the case at an earlier stage ("left column") leads to, or *causes*, a value in another column that describes something that happens to the case at a later stage ("right column").

Case comparison tables are not meant as a device for identifying text fragments/*quotations* that are relevant for qualitative analysis. We assume that a researcher already has created quotations by assigning first, preliminary categories to the textual materials, which might eventually be revisited or revised later in the analysis process. Useful information on the process of coding texts, converting transcripts to tables, and importing files in spreadsheet software can be found in Janesick (2016), La Pelle (2004), Meyer and Avery (2009), and Ruona (2005). A second assumption is that for each case, the researcher has at least one column of quantitative data values (e.g., GPA for each student), and that part of the analysis consists of relating the quantitative data to either qualitative analytic categories and/or quotations.

How to Construct and Sort Case Comparison Tables

Preparing your Case Comparison Table

In this section, we discuss the two iterative key stages of constructing and using case comparison tables: preparing your table and performing the analysis. Depending on the size of your dataset and your needs, you can set up your case comparison table either in a sophisticated software program for qualitative or mixed analysis, or you can use a simple word processor or a spreadsheet. Whatever software you use initially, you can always easily export your case comparison table to one of the other programs. There are three key questions

when preparing a case comparison table for analysis: What are your rows? What are your columns? What will go in the cells?

What are your rows, columns and cells? In case comparison tables, rows are used for the cases that you are studying. Cases can exist at various levels, for example, individuals, groups, locations, setting, events, and discussion utterances (Miles et al., 2020, p. 29). Each row contains information about one case only, such as one student (Example 1) or one practice (Example 2). Columns are typically used to describe characteristics of your cases: Each column describes one type of characteristic of your cases. As a result, each cell describes the quantitative or qualitative value of one characteristic of one case. Although swapping rows and columns is possible, our general advice is to use entities as rows and characteristics as columns throughout the data analysis phase, as you would in a statistical program (this advice differs from Miles et al., 2020, p. 113). This advice reflects the sorting unit in spreadsheets and word processors, which is the column, not the row.

In mixed methods research, cases occupy a special role. They are typically not examined in the quantitative strand of the research, where each respondent commonly contributes to establishing an effect at the group level, and are not examined as individual cases (Yin, 2018). Thus, in Example 1, CEEPT and GPA were used by Lee and Greene (2007) to calculate a mean effect. In a case comparison table, we take the quantitative information of each respondent, and put it side-by-side with qualitative information about this respondent, thereby turning the respondent from a point that contributes to establishing a mean effect, into a case that is investigated for its own sake, a process called "casing" by Ragin (1992, p. 217).

In Tables 24.1 and 24.2, each row provides information about one case only. There are occasions when cases require more than one row. First, your case may be at a higher level than your *unit of observation*. For example, a case comparison table with the interaction in dyads as its case would preferably have one row for each member of the dyad and identify the dyad to which individuals belong in a separate column. It is likely, that characteristics of the dyad interaction are influenced by characteristics of the individuals that make up the dyad, and sorting tables using characteristics of the individual members can bring this to the fore. This can only be undertaken, though, when each individual member has its own row.

Second, you might have similar information about each case from different sources, for example, judgments by students and by faculty members in Example 1. In this case, each judgment should have its own row, in order to be able to investigate differences between, for example, faculty and student judgments. Third, more than one quotation in a text may refer to the same aspect of the same case. In this case, one row should be used for each quotation, which makes it possible to investigate differences among quotations. In Example 1, it is thinkable that two or more quotations refer to the same individual student. As a result, each individual student could have one or more rows in the table, depending on the number of quotations for that student.

Often, you will have information about your cases at higher and lower levels. Pupils, for example, belong to a particular school class, school classes belong to a particular school, schools are located in a particular municipality, and so forth.

In case comparison table analysis, the higher level information is provided in separate columns (e.g., one column for class, one column for school, and one column for municipality). For most studies, this is more convenient than using relational databases to store the characteristics of the class in a separate table, making it impossible to do one sorting analysis on all characteristics of the cases at once (see also La Pelle, 2004, p. 87).

Our advice is this: Use only one table for all information, as long as you are focused on the same unit of analysis. Thus, as long as the student is our unit of analysis (Example 1), we can provide higher level information such as discipline, or information about the students coming from faculty member in the same table. Only when, in more complex case studies, the study contains multiple units of analysis, for example, cases at multiple levels, is it advised to create tables at the level of each case separately.

The information in the cells of a case comparison table will usually be a mix of what can be considered raw data, and data that are the result of qualitative or quantitative analysis. The following data types may be part of your case comparison table:

Static data types

Descriptive data

- Name or ID
- Descriptive data of the unit of analysis or of higher level units, such as age and gender of a child, the teacher of its class, and the size of its school

Collected data

- Scores obtained by direct measurement or self-reporting
- Quotations
- Qualitative observations

Analytic data

- Scores calculated from other scores (e.g., length x width; summative score)
- Scores derived from other scores other than by calculation (e.g., classification as high or low; overall score on the basis of qualitative observations)

Some data types are considered *static*, which means that they will not change after they have been collected, except for correcting errors. They are divided into *descriptive data*, characteristics of a case, and *collected data*, the result of something the case did during the study. *Analytic* data are obtained by some form of data analysis, some form of further processing the static data.

As our analysis of the cases proceeds, increasingly more data tend to be added to each case. In the case comparison table, these show up as additional columns. As a result, the tables tend to become larger during analysis, as an increasing number of columns are added (See Example 1). In our

analysis, we often want to focus on a restricted number of characteristics of each case, and/or on a restricted number of cases, for example if the total number of cases is high. We recommend keeping the analysis manageable by focusing on a specific part of the table (see more details on this below). This advice differs from the advice given by La Pelle (2004, p. 87) and Miles et al. (2020, p. 113), who recommend the construction of several smaller tables. The disadvantage of splitting larger tables into smaller tables, however, is that this cannot easily be undone. In our solution, re-shuffling columns or re-sorting rows to change the focus of our analysis can easily be reverted.

Performing the Analysis.

What can we see? A case comparison table can be used for many purposes. Below are a few examples of the many kinds of connections that we can see and/or explore by adding analytic categories and sorting data columns:

- *Explanation.* The quantitative values can be explained by processes mentioned in the qualitative quotations (Example 1).
- *Multi-dimensionality.* A quotation that had been classified using one analytic category (e.g., high achievement) is shown to consist of two dimensions (high language achievement vs. high academic achievement, represented as two different columns in Example 1).
- *Profiles.* The values of several analytic categories are shown to hang together (demonstrated in Example 2).
- *Relationships among continuous, ordinal, or nominal variables.* A high value in one column tends to go together with either a high or a low value in the other column.
- *Interaction effects.* Values of interaction effects, in which there are different effects for different groups can be difficult to interpret. However, they are relatively easy to identify in a case comparison table if we order the results according to group.
- *Thresholds*—an effect of one variable on another occurs only if the independent variable has reached a specific level (e.g., only for more advanced pupils).
- *Ceiling/floor effects*—one variable does not affect the other, because the value of the other variable is already high/low for all or most participants.
- *Profiles*—values on various variables hang together in specific ways, thereby creating groups.
- *Exceptions (outliers)*—one case or individual has a value in one column that deviates from the remainder of the groups or one individual does not match the profiles that had been discovered.
- *Causation and prediction*—one value occurring earlier in a process, seems to be followed by a specific value of a column that describes an event or characteristic later in the process.

How Do We Perform the Analysis?

Our procedure of continuously sorting and resorting columns can be performed via standard word processing or spreadsheet software. In Microsoft Word, for example, it is possible to use up to three sorting columns simultaneously, and sort the data first on the basis of Column 1, sort within the entries that have the same value in Column 1 on the basis of Column 2, then on the basis of Column 3. If more than three sorting columns are needed, sorting can be undertaken in two or more steps. This advice differs from (La Pelle, 2004), who recommends combining cells, which would create a complex table. Although this might be sound advice in a situation in which much of analysis has been undertaken, our advice for case comparison tables is always to use only one type of information per cell, to keep the freedom of sorting and resorting data in different columns independently.

For qualitative values that have a natural order, it makes sense to add digits to enable meaningful sorting, for example 1. Weak, 2. Moderate, 3. Strong (see Table 24.2). This advice differs from the numerical codes of La Pelle (2004) and Ruona (2005), which are used in conjunction with a codebook, in which the numerical codes are described. Like Meyer and Avery (2009), we prefer to use meaningful qualitative labels in combination with digits for a meaningful sorting order.

Sorting and resorting is the foundation of a case comparison approach to mixed data analysis (Figure 24.1). By adding and using analytic, often qualitative, categories and descriptive, often quantitative, data to sort and compare quotations (as we did in Example 1), we can discover patterns. For example, we might discover links among the descriptive characteristics of participants, the qualitative quotations, and the analytic categories of the researcher (Combs & Onwuegbuzie, 2010; La Pelle, 2004). More importantly, we view sorting and comparing as a way to develop further the qualitative analytic categories themselves. This “thinking with your data” (Ruona, 2005, p. 259) includes questions such as: Do all quotations listed under a specific value of one analytic category really belong there? Does the analytic category and its assigned values provide an adequate description of what happens in the quotations? Do other quotations categorized under a different value of the analytic category really differ? These are the well-known questions for developing grounded theory (Glaser & Strauss, 1967). Using columns with analytic categories to sort cases and to determine whether there is some relationship between analytic categories and quotations is an excellent method to accomplish this because it brings the quotations related to one value of one analytic category next to each other. As we described in Example 2, we prefer meaningful codes (such as “low” or “high”) to numerical codes because it better enables us to see whether the value assigned to the analytic category fits a specific quotation, or whether the quotation should be moved, or the analytic category adapted. In addition, different from Ruona (2005, p. 259), we consider sorting tables on the basis of various columns an excellent method to discover connections among analytic categories. If, for example, the analytic category “language problems (Table 24.1)” is displayed in one column, the analytic category “effect on grade” in another column, and if quantitative measures indicate the

degree of presence or absence of what is described by the analytic category, we can discover various groups of students with different profiles.

Different from grounded theory, however, we emphasize the ideas and preconceptions that researchers bring to their studies (Figure 24.1). Theories do not just emerge from the data, but they arise through an interplay between the researcher’s ideas/theories and the empirical world (Emmel, 2013).

In sorting, we use the values in one or more columns to sort or re-sort the cases—that is, to sort and re-sort the rows. In addition to re-sorting cases, re-shuffling columns, that is changing the order of columns, is a powerful tool for various reasons:

- Reshuffling the columns enables a researcher to focus on those columns that are relevant for that step of the analysis only, and leave other columns out of sight (but do not delete them!).
- Putting columns in the order of their appearance in a process can be used to reflect on process and causation.

As analytic categories change and new analytic categories are added during analysis, two issues are likely to emerge. The first is that one quotation relates to two (or more) values within one analytic category. There are three solutions to this problem: (a) split the quotation into two quotations and assign a different value to each split quotation; (b) duplicate the quotation and assign a different code to each of the two identical quotations; or (c) create a new analytic column and divide the two values over the two columns. Solution (a) and (b), described by Combs and Onwuegbuzie (2010), La Pelle (2004), and Ruona (2005), will lead to additional rows, whereas solution (c), described by Meyer and Avery (2009), will lead to additional columns. Our advice is to use additional columns instead of additional rows whenever possible. The reason for this is that every column that we add during analysis, enhances the possibilities for sorting the cases.

The second issue concerns the descriptions that we use to describe the values in our cells. Often, these descriptions change during analysis. In that case, we recommend changing old descriptions to new descriptions by means of sorting. Sorting can be used to list the cases with the same cell value underneath each other, so that a copy-paste, including all the cells that contain the code, can be undertaken easily, while it invites one to have one more look at the quotations that this code describes. Our advice differs from that of La Pelle (2004) and Ruona (2005), who recommend the use of Find and Replace. A well-known risk of Find and Replace, though, is that of unwanted changes being made in the text that one does not want to change.

What Do We Do with the Outcomes of these Sorting Exercises?

By performing these sorting exercises, we begin to see patterns that might otherwise remain hidden (see Examples 1 and 2 below). However, we know that humans are good at detecting

patterns, including patterns that are actually random fluctuations. Therefore, it is important to put every pattern that we discover to the strongest possible test. Such tests include:

- Identify negative cases and attempt to explain these.
- Test whether and how an explanation of a negative case might apply to positive cases as well.
- Reflect on each of the new categories that have arisen as a result of sorting and ask whether these categories are the best possible ones.
- Use statistical tests.
- Attempt to formulate mechanisms that would explain the patterns found.
- Attempt to find alternative explanations.
- Determine how widely/narrowly the pattern generalizes to the different participants in the study.
- Determine whether the pattern depends on or varies according to some other characteristic or variable.

By repeatedly applying these sorting exercises and tests, the researcher will arrive at a conclusion that is both well-fitting and nuanced.

6 Strengths and Limitations of Case Comparison Tables in Mixed Methods Research

We have shown that creating simple case comparison tables and sorting and resorting them enables researchers to identify patterns that might otherwise remain hidden. Through sorting and resorting, we discovered that students with the highest GPA, but a low CEEPT score, mentioned the use of compensation strategies (Example 1), and that the intervention to raise colorectal cancer screening rates appeared to be the least successful in practices with a strong team culture (Example 2). In addition to the increased knowledge based on additional analysis, both cases provide clear directions for future research to explain further these patterns.

Case comparison tables seem less suited for narrative analysis. By making sorting quotations a key feature of our analysis (Example 2), the connections among quotations, the story, can get lost. That makes sortable tables a perfect tool for finding themes across cases or individuals, but not for analyzing what happens within a case.

7 Resources for Learning More About Tables and Joint Displays in Mixed Analysis

- Bustamante, C. (2019). TPACK and teachers of Spanish: Development of a theory-based joint display in a mixed methods research case study. *Journal of Mixed Methods Research, 13*, 163–178. doi:10.1177/1558689817712119
- Guetterman, T. C. (2019, September 4). *Joint displays to facilitate integration of qualitative and quantitative research* [Video file]. *Mixed methods webinar series*. International Institute for Qualitative Methodology. Retrieved from www.youtube.com/watch?v=U6KvCN-7ZKM

- Johnson, R. E., Grove, A. L., & Clarke, A. (2019). Pillar integration process: A joint display technique to integrate data in mixed methods research. *Journal of Mixed Methods Research, 13*, 301–320. doi:10.1177/1558689817743108
- McCrudden, M. T., Schraw, G., & Buckendahl, C. W. (2015). *Use of visual displays in research and testing: Coding, interpreting, and reporting data*. Charlotte, NC: Information Age Publishing.
- Younas, A., Pedersen, M., & Durante, A. (2020). Characteristics of joint displays illustrating data integration in mixed-methods nursing studies. *Journal of Advanced Nursing, 76*, 676–686. doi:10.1111/jan.14264

REFERENCES

- Bazeley, P. (2018). *Integrating analyses in mixed methods research*. Los Angeles, CA: Sage.
- Combs, J. P., & Onwuegbuzie, A. J. (2010). Describing and illustrating data analysis in mixed research. *International Journal of Education, 2*(2), E13. doi:10.5296/ije.v2i2.526
- Emmel, N. (2013). *Sampling and choosing cases in qualitative research: A realist approach*. London, England: Sage.
- Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health Services Research, 48*, 2134–2156. doi:10.1111/1475–6773.12117
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York, NY: Aldine de Gruyter.
- Guetterman, T., Creswell, J. W., & Kuckartz, U. (2015). Using joint displays and MAXQDA software to represent the results of mixed methods research. In M. T. McCrudden, G. Schraw, & C. W. Buckendahl (Eds.), *Use of visual displays in research and testing* (pp. 145–175). Charlotte, NC: Information Age Publishing.
- Guetterman, T. C., Fetters, M. D., & Creswell, J. W. (2015). Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Annals of Family Medicine, 13*, 554–561. doi:10.1370/afm.1865
- Janesick, V. J. (2016). *"Stretching" exercises for qualitative researchers*. Los Angeles, CA: Sage.
- Johnson, R. B., & Christensen, L. B. (2020). *Educational research: Quantitative, qualitative, and mixed approaches* (7th ed.). Los Angeles, CA: Sage.
- Kegan, R., & Lahey, L. L. (2009). *Immunity to change: How to overcome it and unlock the potential in yourself and your organization*. Boston, MA: Harvard Business Review Press.
- La Pelle, N. (2004). Simplifying qualitative data analysis using general purpose software tools. *Field Methods, 16*, 85–108. doi:10.1177/1525822x03259227
- Lee, Y.-J., & Greene, J. (2007). The predictive validity of an ESL placement test: A mixed methods approach. *Journal of Mixed Methods Research, 1*, 366–389. doi:10.1177/1558689807306148
- Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Los Angeles, CA: Sage.
- Meyer, D. Z., & Avery, L. M. (2009). Excel as a qualitative data analysis tool. *Field Methods, 21*, 91–112. doi:10.1177/1525822x08323985

- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). Los Angeles, CA: Sage.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London, England: Sage.
- Plano Clark, V. L., & Sanders, K. (2015). The use of visual displays in mixed methods: Research strategies for effectively integrating the quantitative and qualitative components of a study. In M. T. McCrudden, G. Schraw, & C. W. Buckendahl (Eds.), *Use of visual displays in research and testing* (pp. 177–206). Charlotte, NC: Information Age Publishing.
- Ragin, C. C. (1992). "Casing" and the process of social inquiry. In H. S. Becker & C. C. Ragin (Eds.), *What is a case? Exploring the foundations of social inquiry* (pp. 217–226). Cambridge, England: Cambridge University Press.
- Ruona, W. E. A. (2005). Analyzing qualitative data. In R. A. Swanson & E. F. Holton, III (Eds.), *Research in organizations: Foundations and methods of inquiry* (pp. 233–263). San Francisco, CA: Berrett-Koehler Publishers.
- Ryan, G. W. (2004). Using a word processor to tag and retrieve blocks of text. *Field Methods*, 16, 109–130. doi:10.1177/1525822x03261269
- Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69(2 Supplement), 107–131. doi:10.1007/s11577-017-0454-1
- Shaw, E. K., Ohman-Strickland, P. A., Piasecki, A., Hudson, S. V., Ferrante, J. M., McDaniel, R. R., Jr., . . . Crabtree, B. F. (2013). Effects of facilitated team meetings and learning collaboratives on colorectal cancer screening rates in primary care practices: A cluster randomized trial. *Annals of Family Medicine*, 11, 220–228. doi:10.1370/afm.1505
- Yin, R. K. (2018). *Case study research and applications* (6th ed.). Los Angeles, CA: Sage.