

## **An Approach for Being Able to Use the Options of Calculating Inter-coder Reliability Manually and Through Software in Qualitative Research of Education and Training in Sports**

**Ali Sevilmiş<sup>i</sup>**

Karamanoglu Mehmetbey University

**Ozer Yildiz<sup>ii</sup>**

Necmettin Erbakan University

### **Abstract**

Reliability that can be proved by numeric indicators in quantitative studies has become a very discussible issue. The reason for this is to be thought that in qualitative researches, reliability is not based on positive perspective and those forming reliability criteria is difficult. However, for testing the reliability of a qualitative study or raising it to the top level, some ways are used. The objective of this study is to introduce an approach to be able to use the options of inter-coder reliability calculating manually and through software in qualitative researches of education and training in sports. This research was conducted by document review method. In this context, how inter-coder reliability in qualitative researches will be formulated, how it will be calculated in software dimension, and the dimensions of inter-coder agreement, code frequency in document, code overlapping rate of segment level, and revealing code frequency were discussed. In the study, document analysis technique among qualitative research designs and in analyses content analysis method were used. The study results revealed with the examples that inter-coder reliability could be calculated in the various ways and that percentage flexibility could differ in every formula. Also, the options of inter-coder agreement were calculated by means of software, and it was identified that each option used had a different flexibility. This study, in terms of its originality, drawing the existing reliability from related to qualitative study from the framework of inter-coder agreement, suggests a synthesis of contemporary viewpoints. It is considered that in inter-coder qualitative studies, especially in sports education and training studies, will guide to the relevant researchers.

**Keywords:** Calculation, Inter-Coder Reliability, Inter-Coder Agreement, Maxqda, Option, Qualitative

**DOI:** 10.29329/ijpe.2021.332.23

---

<sup>i</sup> **Ali Sevilmiş**, Assist. Prof. Dr., Sports Management, Karamanoglu Mehmetbey University, ORCID: 0000-0001-8338-5927

**Correspondence:** alisevilmis.42@hotmail.com

<sup>ii</sup> **Ozer Yildiz**, Prof. Dr., Department of Physical Education and Sports, Necmettin Erbakan University, ORCID: 0000-0002-2470-5457

## INTRODUCTION

Evaluating the reliability of the research results has a great importance for both quantitative and qualitative researches, because reliability gives information about “soundness” of the research related to wholeness of final results. One of the most important criticisms raised to qualitative researches is the opinion that the adopted methods are weakly given justification, that there is no transparency in analytical procedures, that the findings have a total of personal principles only subjecting to the prejudice of researcher, and that they are devoid of scientific accuracy (Sandelowski, 1993). These thoughts are seen among the subjects that can be neglected in qualitative researches (Rädiker & Kuckartz, 2019). However, it is expressed that qualitative research methods are not also based on statistical calculations. But these criticisms are not right for the researches presenting holistic and systematic viewpoints about phenomena, because there are statistical calculations used in evaluating qualitative researches (Rolfe, 2006).

In spite of all these criticisms, qualitative researches also have moral and ethical codes like quantitative researches. Elimination of problems associated with moral and ethical matters is the confirmation of validity of that research (Gomes & Duarte, 2020). The criteria such as triangulation, specialist evaluation, and supervision trace used for providing the validity and reliability of qualitative researches are also the criteria positivism also uses (Hanson et al., 2005). Just in positivism approach, the designs of qualitative researches are also methods, in which statistical calculations are used, and inter-coder reliability can be provided.

Inter-coder agreement among reliability criteria of qualitative researches is to introduce the result of two independent coders according to the scientific essentials (Rädiker & Kuckartz, 2019). Inter-coder agreement is associated with the compliance of the subject studied (Joffe & Yardley, 2004) and whether or not codes can reach the same results as a result of coding for the same documents the coders make or how much they agree with categories.

In qualitative researches, inter-coder reliability is an indispensable criterion in terms of compliance of the codes to be assigned to the research and themes to be formed (Freelon, 2010). Inter-coder reliability approach, the questions of at what extent two experts define the same subjects as similar and whether or not they form similar codes and categories are focused on. It is possible for two experts to agree with the same document and evaluate a phrase or paragraph in the same document as the different categories. For, in qualitative researches, coding and dividing into categories are not clearly formulated (Kuckartz, 2016). In this case, movement point must be that two expert coders firstly deal with the research not only according to its own structure but also based on the result of a finding or conceptual framework. As a result, while inter-coder reliability is provided, scientific criteria must be performed.

For inter-coder reliability, many software programs that can help to evaluate the results (Richards & Richards, 1994). While each program makes certain analyses (coding, reviewing codes), it has its opportunities and weaknesses. However, every program makes contribution to the originality, reliability, and quality of data (Weitzman & Miles, 1995).

Nowadays, inter-coder agreement is realized both manually and through software programs (Weitzman & Miles, 1995). MAXQDA program, one of software program used in analysis programs of qualitative researches, has many characteristics meeting the expectation of qualitative researchers. It has an interface incorporating many features such as searching by qualitative data, fast coding, updating, and recording the data (MAXQDA, 2020). Beside this, in reliability analyses, it is a program that can extract a numeric measure of the agreement between two independent coders by means of the different methods of value (Rädiker & Kuckartz, 2019).

When domain literature is examined, the researches, which calculate inter-coder reliability, show the existing options, and deal with the options of analysis program in detail, are limited. The

objective of this study is to introduce an approach for being able to use the options of calculating inter-coder reliability manually and through software in education and training researches in sports. Therefore, in this study, inter-coder reliability is both dealt with conceptual framework and that all stages of inter-coder reliability are introduced manually and through a software program (MAXQDA, 2020).

## **METHOD**

### **Research Model**

In case that direct observation and interview are not possible, the articles associated with the research problem can be included in the study. Just as document review or analysis can be an analysis method alone, also in case that the other qualitative methods, it can be used as an extra information resource (Kıral, 2020).

In this study, document analysis technique and content analysis method were used. The study data were obtained by means of the articles containing information inter-coder agreement, books, and MAXQDA software program. In content analysis, application of inter-coder agreement reliability was shown through MAXQDA qualitative data analysis software. Among options in MAXQDA software, code occurrence in the document, code overlapping rate of segment, code frequency in document” options were introduced.

### **Collecting data**

The content of domain software was examined by means of information form for how inter-coder agreement is evaluated. Domain literature information about inter-coder agreement is presented in the form of table. Also, inter-coder agreement formulas and options, obtained from domain literature, were given together with their examples. Domain literature content, later collected, was checked by three researchers, and possible errors were prevented.

### **Options of Calculating Inter-coder agreement**

For examination of study case, example project of MAXQDA qualitative data analysis program (ENG/Life Satisfaction.mx20) was utilized. With this example study, inter-coder agreement options in MAXQDA program were examined in terms of code occurrence in document, code overlapping rate of segment, and calculation of code frequency.

## **FINDINGS**

### **Reliability**

Which strategies can qualitative researchers adopt for providing the reliability of research findings? In contrast to quantitative researchers applying statistical methods for the validity and reliability of the research findings, qualitative researchers aims to design and combine strategies for providing reliability of findings.

In qualitative researches, providing the reliability of study is related to consistency. In other word, it is understood whether or not to reproduce findings from this expression (Allan, 2020). Lincoln and Guba (1985) conceptualized reliability as consistency. Just as in each stage of the study, in the stage of coding and theming, introducing the persuasive evidence is a part of reliability (Hennink et al., 2020).

In the stage of theming and coding the data, during comparing the different viewpoints, introducing similarity and differences is to reveal how the data are used and consistency of the research, in which the data are categorized. Namely, for providing the consistency of the research, the decisions in forming the themes and codes belonging to that research are evaluated as a criterion of consistency.

### **Inter-coder agreement reliability**

In a coding framework evaluation of inter-coder reliability is frequently recommended in qualitative research as a good application (O'Connor & Joffe, 2020). Although it is argued that inter-coder agreement is an inappropriate or unnecessary step among the targets of qualitative research, inter-coder agreement helps researchers persuade about the reliability and systematics of coding process (Rädiker & Kuckartz, 2019). For, inter-coder agreement is a stage of associating the codes assigned with dataset. Being able to make an accurate association depends on proper execution of this stage. Inter-coder agreement reliability is a numeric measure of agreement between the different coders regarding how to use the same data (O'Connor & Joffe, 2020) and examination of whether or not which segments and codes match (MAXQDA, 2020). In other words, inter-coder reliability expresses the consistency of that the different people use the same data.

While inter-coder agreement is a standard measure of study quality, a low level inter-coder agreement means that the study is of poor quality. Low level inter-coder agreement is indicator of unclear coding, inadequate agreement of coders on issue, weakness in the theming and coding methods of the research (Lavrakas, 2008). The way of overtaking this problem is possible with keeping coding away from individualism, independent coding of two different experts on the same dataset, and comparing the similarities of two experts to each other in terms of coding and theming (Rädiker & Kuckartz, 2019).

### **Inter-coder agreement in education and training researches in sports**

In education researches, it has begun to be seen that it is not always possible to hypothesize according to a certain theory or model and keep some variables fixed, depending on these hypotheses, hence, to produce generalizable information in educational area, and that the existing physical, social, cognitive, and psychological features in any environment are the most important elements determining the results that emerge (Yıldırım, 1999).

Although qualitative researches introduced several generalizable information, since this information is at very general level, education at applier position remains inadequate in presenting practical information and adequate suggestions to the users. Therefore, educations went towards qualitative researches, considered that they have more explanatory features and are enough in bringing the meaning in the results. Thus, beyond the general findings reached via qualitative research methods, reaching more explanatory and detailed and deep information, and introducing this information for the service of practitioners were aimed.

In the recent years, many methods are used related to education and training researches in sports. Some of these are “theory development”, “phenomenology”, “observation”, “content analysis”, and “discourse analysis”. Consistency used for the data collected in qualitative researches covering a large study process is relatively important (Lester et al., 2020). Inter-coder reliability is a criterion in testing this consistency. Albeit evaluating reliability is an important component for these researches, selecting and using accurate method in inter-coder reliability is an important challenge for the researchers (Olson et al., 2016; Venkatesh et al., 2013).

The place and importance of qualitative researches in educational studies are not discussible but it is necessary for study design, coding, theming, and agreement of codes and themes to each other

to complete and support (Yıldırım, 1999). In the existing analyses, if two independent experts provide so agreement, it prevails the view that the findings of that study will be repeatable, consistent, and reliable (Davies, 2012). However, in providing inter-coder reliability, it is thought that there is a need for more systematic guide that can introduce inter-coder difference.

In qualitative researches, it is seen that there are various studies evaluating inter-coder reliability (ICR) of the studies calculating inter-coder agreement by means of a qualitative research method and deal with (Campbell et al., 2013; Friese, 2019; González-Prieto et al., 2020; MacPhail et al., 2016; O'Connor & Joffe, 2020; Rädiker & Kuckartz, 2019; Wilson-Lopez et al., 2019). But when the domain literature is examined, any resource dealt with evaluating inter-coder reliability (ICR) of education and training in sports was not met.

In education and training researches in sports, if researchers select a suitable coding method or computer software, coding and theming errors will be identified and corrected before introducing the study results (Venkatesh et al., 2013). Using a suitable method in inter-coder agreement will reduce the biasness during coding and theming (Nili et al., 2020).

In this context, the relevant literature regarding the formulas of inter-coder agreement methods was reviewed, and the most suitable approach of calculation method in inter-coder reliability in the education and training researches in sports was introduced. In addition, in education and training researches in sports, an inter-coder agreement program was introduced with the calculation examples of code frequency in document and code overlapping rate of segment.

### **Manually calculating inter-coder agreement**

For reliability of inter-coder agreement, what is important is not only calculating it manually or by means of computer programs but also the data of the same text, coded without giving information to each other is consistent (Silverman, 2005). Beside this, there are very few standards and accessible guide regarding the suitable procedures to be used for evaluating and reporting inter-coder reliability or auxiliary instrument for calculating it (Lombard et al., 2002). "Inter-coder agreement" can be calculated by means of certain methods in the literature (Rädiker & Kuckartz, 2019). But match percentage that is the most used common method of using inter-coder reliability control is the most flexible one (Nili et al., 2017).

### **Inter-coder reliability;**

---

Matches / (Matches /Non-Matches)

---

This agreement conceptualized as inter-coder reliability is calculated by using the formula above. This formula shows similarity with Miles and Huberman (1994) formula.

Document	Agreement	Disagreement	%
Participant -1	9	0	100
Participant -2	9	0	100
Participant -3	8	1	0,88
Total	26	1	0,96

Percentage of agreement is expressed as in the figure above. The agreement and disagreement situations regarding 3 different meetings are shown above. There is no compromise for participant 1 and participant 2. But in participant 3, the disagreement code is 1. Accordingly, the agreement percentage of each participant and the codes of all participants can be calculated with the following formula.

- Matching codes of Participant -3 are calculated in the form of  $8/8+1 = 0.88$  as percentage and

- Matching codes of three participants, in the form of  $26/26+1 = 0.88$  as percentage.

Beside this, the different formulas are used for agreement rates of datasets coded by coders. Therefore, it shows that calculating inter-coder agreement is based on the various assumptions, and that calculations made on the same data i.e. numeric measure of inter-coder agreement is different (Rädiker & Kuckartz, 2019). In the following table, the information belonging to the codes coded by two coders takes place. The details of codes Coder-1 and Coder -2 coded and rejected the same or only one coder accepted are shown.

		Coder 1	
		Codes or variables that are the same	Codes or variables that are not the same
Codes 2	Codes or variables that are the same	A	B
	Codes or variables that are not the same	C	D

A: Codes Coder-1 and Coder-2 accepted as the same-coded.

D: Codes Coder -1 and Coder -2 rejected

B and C: Codes only one coder accepted

As a result, to make numeric the expressions:

		Coder 1	
		Codes or variables that are the same	Codes or variables that are not the same
Coder -2	Codes or variables that are the same	20	5
	Codes or variables that are not the same	10	15

Coding of a study, made by two different coders on the same document regarding educational style of physical education teachers, is shown above. Each coder made coding. For fifty coding made, coding is seen in column for Coder-1 and, in row, for Coder-2. If the numeric values above are placed in each of inter-coder agreement formula, the following results can be obtained.

Sorts of inter-coder agreement	Definition	Formula	Calculation
Simple Agreement	Codes and variables that are the same and codes and variables that are not the same are accepted as agreement.	$(a+b)/(a+b+c+d)$	$(20+5)/(20+5+10+15) = 0.50 \text{ \%} 50$
Jaccard	Codes and variables that are not the same are ignored.	$a / (a+b+c)$	$(20) / (20+10+5) = 0.57 \text{ \%} 57$
Rädiker and Kuckartz	Occurrence is evaluated two times and that is absent is single.	$(2a+d) / (2a+b+c+d)$	$(2.20+5) / (2.20+5+10+15) = 0.64 \text{ \%} 64$
Russel and Rao	Only occurrence is evaluated as agreement. Codes and variables that are not the same reduce similarity.	$a / (a+b+c+d)$	$(20) / (20+5+10+15) = 0.40 \text{ \%} 40$

In the table above, the definitions, formulas, and calculations of inter-coder agreement are shown (Rädiker & Kuckartz, 2019). When all the results are multiplied by 100, from this formula,

inter-coder reliability statistics, known as match percentage. Beside this, inter-coder agreement premise by Rädiker and Kuckartz (2019) is very flexible. It suggests agreement at maximum level. Whichever agreement method is used, it is seen that coding agreement of qualitative researches is suitable for certain features (Nili et al., 2017).

### **Example of inter-coder agreement on program: Maxqda**

While coding data in qualitative studies, presence of certain criteria will increase persuasiveness of the results reached. For example, that coding is not arbitrarily and randomly made, instead of this, that it is made by two experts will enable consistency of these documents coded to be evaluated in compliance with scientific framework (Erlandson et al., 1993).

After the documents having the same content are coded by two independent experts, it will be assumed that this coding will reach a certain level of safety level. MAXQDA software will also enable to compare two people coding the same document thanks to its function of inter-coder agreement (Rädiker & Kuckartz, 2019). Here, the aim of comparing independent coders is to discuss distinctions, understand why the distinctions occur, and improve agreement safety, evaluating all deficiencies. MAXQDA presents a number of code occurrences in the document and code frequency in the document providing to improve and validate systematic analysis of inter-coder agreement.

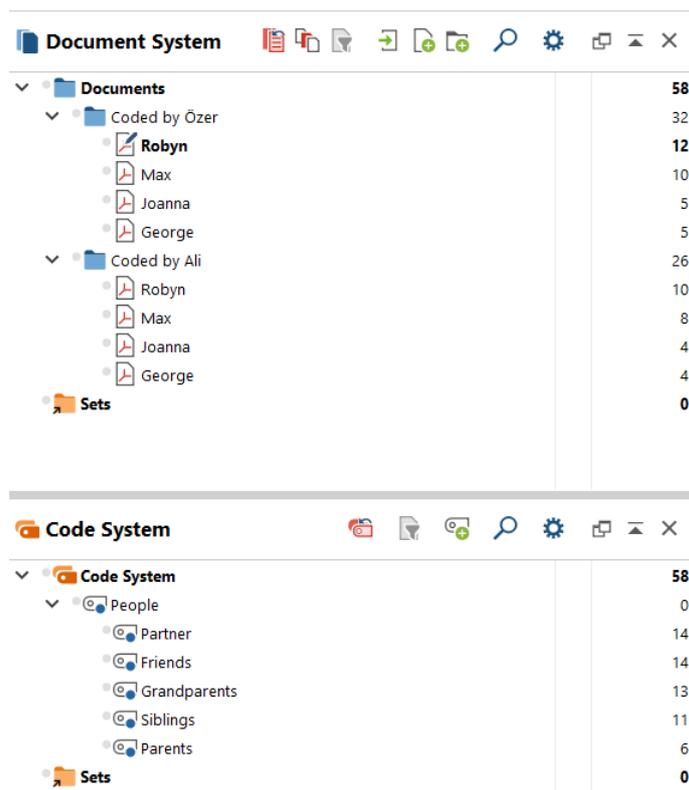
### **The inter- coder agreement procedure**

*Before loading Maxqda software, checking inter-coder agreement includes the following steps:*

“Two coders process the identical document independently and code it according to mutually agreed code definitions. This can be done on the same computer or on separate computers. Of course, it is important that both coders cannot see what the other person has coded.”

“The two identical documents coded by different coders must be contained in the same MAXQDA project. The documents must have the same name, but must be in different document groups or document sets (MAXQDA, 2020).

In MAXQDA, for being able to accurately analyze inter-coder agreement, it is necessary for the names of the documents compared to be the same. If document name is not the same, MAXQDA displays a warning message regarding the place, in which documents are different. In this case, for avoiding to produce erroneous results, first of all, it is necessary to examine the difference in every document. As in the following figure, it is seen that the names of documents coded by two different coders are the same.



**Figure- 1. View of two independent encoders in document and code system**

As seen in the document system above, four interviews with the same content were coded by encoder 1 (Özer) and encoder 2 (Ali). Encoder 1 assigned 32 codes to four interview documents, while encoder 2 assigned 26 codes and two encoders total 58 codes to four interview documents. The coding details of the two independent encoders in Table 1 are given below.

**Table 1. Coding details of two independent encoders**

Coded by Ozer	Partner	Friends	Grandparents	Siblings	Parents	Total
Robyn	3	3	3	2	1	12
Max	2	2	2	2	2	10
Joanna	1	1	1	1	1	5
George	1	1	1	1	1	5
Coded by Ali	Partner	Friends	Grandparents	Siblings	Parents	Total
Robyn	3	3	2	1	1	10
Max	2	2	2	2	0	8
Joanna	1	1	1	1	0	4
George	1	1	1	1	0	4
<b>Total</b>	<b>14</b>	<b>14</b>	<b>13</b>	<b>11</b>	<b>6</b>	<b>58</b>

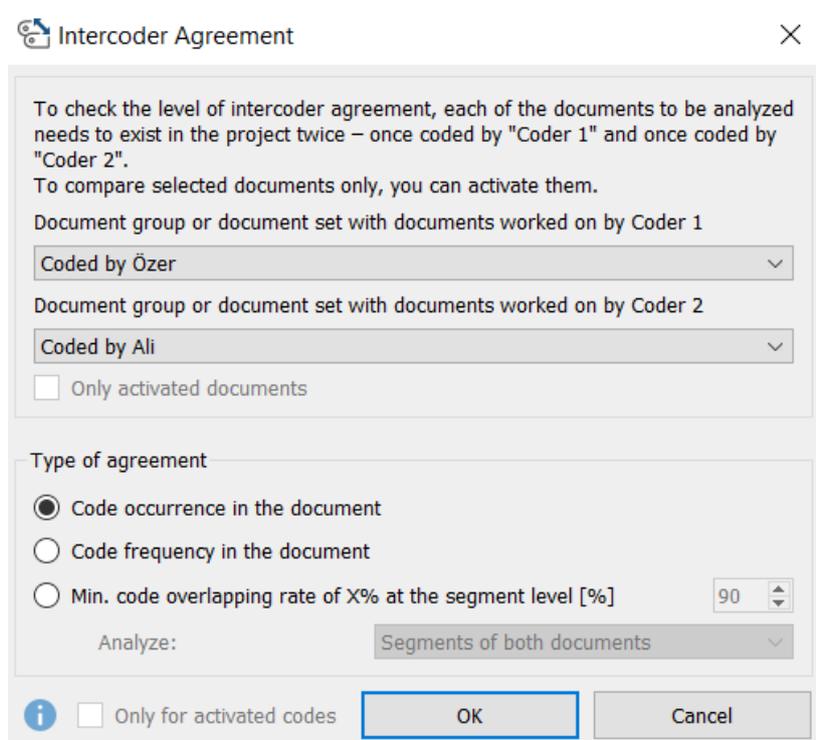
In Figure 1, details of a total of 58 encodings made by encoder 1 (Özer) and encoder 2 (Ali) to documents and sections can be seen. As seen in Table 1, the underlined bold areas show the disagreements. Table 2 shows the summarized situation of the above table.

**Table 2. Coding sum of two independent encoders in code system**

	Partner	Friends	Grandparents	Siblings	Parents	Total
Coded by Ozer	7	7	7	6	5	32
Coded by Ali	7	7	6	5	1	26
Total	14	14	13	11	6	58

The table above shows the sums of encodings assigned to each document made by the two encoders. It is important to understand these tables to have information about how coding content is calculated before seeing the inter-coder negotiation options of the MAXQDA program.

In MAXQDA, while inter-coder agreement is calculated, codes there are the options such as occurrence in document, code frequency in documents, code overlapping rate of segment. Agreement analysis can be made for all sorts of documents i.e. PDFs, tables, images as well as sound and video files.



**Figure 2. Options dialog box for inter-coder agreement analyses**

For inter-coder agreement analysis, when options of dialog box are examined, it is seen that there are three options for agreement. These options are code occurrence in the document, code frequency in the document, and code overlapping rate of the segment level.

## Inter- coder Agreement

### 1. Code occurrence (existence) in the document

Code occurrence (existence) in the document: A match is counted if both coders have assigned the same code to the document. It does not matter in this case whether one person assigned the code three times and the other only once (Rädiker and Kuckartz, 2019). The important thing is that the code is assigned by two coders, even once. For example Ozer has assigned the code “Grandparents” three times in the Robyn document and Ali has assigned the code “Grandparents” twice. There is no

conflict here. There is no question of disagreement here. If Ali had not assigned any code to that code then there would be a conflict. For example Ozer has assigned the code “Parents” in the Max document and Ali has no assigned the code “Parents”. Therefore, there is disagreement here.

The details are shown in Table 3.

**Table 3. Code occurrence (existence) in the document**

Coded by Ozer	Partner	Friends	Grandparents	Siblings	Parents	Total
Robyn	3	3	3	2	1	12
Max	2	2	2	2	2	10
Joanna	1	1	1	1	1	5
George	1	1	1	1	1	5
Coded by Ali	Partner	Friends	Grandparents	Siblings	Parents	Total
Robyn	3	3	2	1	1	10
Max	2	2	2	2	0	8
Joanna	1	1	1	1	0	4
George	1	1	1	1	0	4
Total	14	14	13	11	6	58

Considering the underlined areas as shown in Table 3, Ozer has assigned the code “Parents” once in the document Max, Joanna and George. But Ali has no assigned the code “Parents” in the document Max, Joanna and George. Therefore, the code disagreement exists in the Parents code in Max, Joanna, and George documents. The details are shown below in Table 4.

**Table 4. Code occurrence (existence) in the document summary**

Code	Relevant	Irrelevant	Total	Percentage
Partner	4	0	4	100,00
Friends	4	0	4	100,00
Grandparents	4	0	4	100,00
Siblings	4	0	4	100,00
Parents	1	3	4	25,00
Total	17	3	20	85,00

Table 4 summarizes the Partner, Friends, Grandparents and Siblings themes, while there is a complete consensus in terms of the existence of code in the document, while there are three uncompromised codes in the Parents theme (Table 3 for details). When all these codes are considered together, a total of 17 codes are related and 3 codes are unrelated. According to this, from the formula (Agreement) / (Agreement + Disagreements),  $17/(17+3) = 0.85$ . When this is multiplied by 100, the rate two coders agree with it is calculated as 85%.

## 2. Code frequency in the document

“A match is counted if both coders have assigned the same code in the document the same number of times” (Rädiker and Kuckartz, 2019) If one person has assigned the code "parents" two times in the document and the other one, there is no agreement for this code Again, the locations of the coded segments in the document do not play a role here. For example, in the following figure, while Ozer assigned two codes to Max document in the code “Parenets”, Ali did not assign any code. There is one disagreement for this code of Max document. The details are shown in Table 5.

**Table 5. Showing Code frequency in the document**

Coded by Ozer	Partner	Friends	Grandparents	Siblings	Parents	Total
Robyn	3	3	3	2	1	12
Max	2	2	2	2	2	10
Joanna	1	1	1	1	1	5
George	1	1	1	1	1	5
Coded by Ali	Partner	Friends	Grandparents	Siblings	Parents	Total
Robyn	3	3	2	1	1	10
Max	2	2	2	2	0	8
Joanna	1	1	1	1	0	4
George	1	1	1	1	0	4
Total	14	14	13	11	6	58

Considering the underlined areas as shown in Table 5, while Ozer has assigned the code “Grandparents” three times in the document and Ali has assigned the code “Grandparents” twice. And Ozer has assigned the code “Siblings” twice, Ali has assigned the code “Siblings” once. Ozer has assigned the code “Parents” twice in the Max document. Ali has assigned once. Ozer has assigned the code “Parents” once in the Joanna document. Ozer has assigned the code “Parents” once in the George document. Ali did not assign any code in these two documents.

**Table 6. Code frequency in the document summary**

Code	Relevant	Irrelevant	Total	Percentage
Partner	4	0	4	100,00
Friends	4	0	4	100,00
Grandparents	3	1	4	75,00
Siblings	3	1	4	75,00
Parents	1	3	4	25,00
Total	15	5	20	75,00

A full consensus has been achieved on Partner and Friends themes coded by Ozer and Ali. However, no consensus could be reached on the themes of Grandparents, Siblings and Parents. As a result, a total of five frequencies with no consensus were obtained (Table 5 for details). According to this, from the formula (Agreement) / (Agreement + Disagreements),  $(15)/(15+5) = 0.75$ . When it is multiplied by 100, the rate of view two coders agree with it is calculated as 75%.

### 3. Min. code overlap between segments (%)

“A match is counted if both coders have assigned the code to a given data segment. The segments do not have to be 100% identical in their position; you can set a tolerance range” (Rädiker and Kuckartz, 2019). The difference in the codes assigned here is considered a dispute. For example Ozer has assigned the code “Grandparents” three times in the Robyn document. If Ali has no code assigned. There are three disagreements for this code of Robyn document. In the following table, a total of six irrelevant codes are given in detail in Figure.

**Table 7. Showing Code overlapping rate of X % at the segment level**

Coded by Ozer	Partner	Friends	Grandparents	Siblings	Parents	Total
Robyn	3	3	3	2	1	12
Max	2	2	2	2	2	10
Joanna	1	1	1	1	1	5
George	1	1	1	1	1	5
Coded by Ali	Partner	Friends	Grandparents	Siblings	Parents	Total
Robyn	3	3	2	1	1	10
Max	2	2	2	2	0	8
Joanna	1	1	1	1	0	4
George	1	1	1	1	0	4
Total	14	14	13	11	6	58

In Table 7, in bold marked sections, the relevant codes are given. Ozer has assigned the code “Grandparents” three times in the Robyn document and Ali has assigned the code “Grandparents” twice. Ozer has assigned the code “Siblings” twice in the Robyn document and Ali has assigned the code “Grandparents” once.

Ozer has assigned the code “Parents” twice in the Max document and Ali has assigned the code “Parents” once. Ozer has assigned the code “Parents” once in the Joanne and George document and Ali has not assigned any code to these two documents. In the following table, irrelevant six codes are given.

**Table 8. Code overlapping rate of X% at the segment level summary**

Code	Relevant	Irrelevant	Total	Percentage
Partner	14	0	14	100,00
Friends	14	0	14	100,00
Grandparents	12	1	13	92,31
Siblings	10	1	11	90,91
Parents	2	4	6	33,33
<Total>	52	6	58	89,66

As a result, it is seen that Ali and Özer have assigned 6 codes different from each other. Ozer has assigned the codes “Grandparents” “Siblings” and “Parents” more than one code from Ali in the Robyn, Max, Joanna, George documents. As a result, there was a conflict in six codes. According to this, from the formula (Agreement) / (Agreement + Disagreements), calculating code overlapping rate of segments as  $(52)/(52+6) = 0.89$ , when [this value] is multiplied by 100, it was revealed that two coders agree with it and that agreement rate is 89%.

## CONCLUSION

In this study, inter-coder reliability was attempted to be introduced based on the premises discussed in the relevant literature. Inter-coder agreement keeps an important place in evaluating the reliability of content analysis for qualitative researches. When domain literature is evaluated, the existing views about inter-coder agreement are different (Cho & Trent, 2006; Denzin & Lincoln, 2011; Maxwell, 1992; Rädiker & Kuckartz, 2019). These views were considered and scrutinized but the examples were presented related to formulas introduced by Rädiker and Kuckartz (2019) and it was emphasized that the formula suggested by the researcher himself was more flexible.

Inter-coder reliability has different features from manual calculation. Program (MAXQDA) calculates inter-coder agreement in the context of code frequency in documents, code overlapping rate of segments, and code occurrence in document

When the calculations made through the data are considered, it can be said that code overlapping rate of segments gives a more flexible result of inter-coder reliability. This case can differ

in the different coding in different studies. Therefore, if it is necessary to use a program, while drawing inter-coder agreement reliability, it is suggested to be tested all alternative agreement features of program that are suitable for the study.

It is considered that the study will guide to the relevant researchers in formulating the processes of inter-coder reliability in qualitative studies, especially in education and training researches in sports, in calculations or through qualitative data analysis program.

That qualitative methods are often preferred in educational studies, especially in educational studies in sports, due to the fact that they make scientific contribution to the solution of many problems, bringing a qualitative framework in these studies is possible by making the study acceptable, shortening publication processes in scientific articles, and performing many criteria, especially inter-coder reliability criterion.

In this scope, when domain literature is examined, specific to inter-coder reliability, there are limited numbers of studies. In the tradition of qualitative study, this study introducing a detailed framework regarding the concept of inter-coder reliability is expected to make contribution to the literature about inter-coder reliability. In addition, in the relevant studies, especially in education and training in sports, this study is considered to form a ground for placing in certain criteria inter-coder reliability.

## **DISCUSSION**

In this study, the issue of reliability between coders was discussed based on the propositions that exist in the literature. In this context, qualitative formulas, alternative coding consensus calculation options were examined in qualitative research. It is also given with options for applying consensus between coders on a qualitative program.

As a result, reliability strategies in qualitative research can be defined and applied as in quantitative research methods. It is thought that conducting these processes in a health way will improve the concept of trustworthiness in qualitative research.

Even if a single formula for inter-coder reliability is not used, making acceptable definitions or revealing operational criteria will contribute to the quality of qualitative research. In the light of these discussions, it has been revealed that it is possible to determine criteria for calculating reliability in qualitative research.

Convincing the reader about the legitimacy of qualitative research can be possible with reliability. For this reason, a reliability criterion suitable for the selected research paradigm is important. This situation convinces the researchers of the realism of the research.

## **RECOMMENDATIONS**

As a result of the research, the following suggestions can be made about inter-coder reliability; Broader research on inter-coder reliability can be constructed. The characteristics of qualitative programs can be compared in inter-coder reliability. As in this research, the use of a standard process in inter-coder consensus can contribute to the reliability of research.

### **Limitations**

The limitation of this research can be listed as follows;

The information reached about the inter-coder reliability in the research is limited to the articles or books in the literature. The calculations for inter-coder reliability in the study are limited to

the MAXQDA qualitative data analysis program calculation options. Inter-coder reliability does not include different reliability (such as kappa) calculation methods.

## REFERENCES

- Allan, G. (2020). Qualitative research. In *Handbook for research students in the social sciences* (pp. 177-189). Routledge.
- Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods Research*, 42(3), 294-320. <https://doi.org/https://doi.org/10.1177/0049124113500475>
- Cho, J., & Trent, A. (2006). Validity in qualitative research revisited. *Journal Qualitative Research*, 6(3), 319-340. <https://doi.org/https://doi.org/10.1177/1468794106065006>
- Davies, K. (2012). Content analysis of research articles in information systems (LIS) journals. *Library Information Research*, 36(112), 16-28. <https://doi.org/https://doi.org/10.29173/lirg495>
- Denzin, N. K., & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research*. Sage.
- Erlandson, D. A., Harris, E. L., Skipper, B. L., & Allen, S. D. (1993). *Doing naturalistic inquiry: A guide to methods*. Sage.
- Freelon, D. G. (2010). ReCal: Intercoder reliability calculation as a web service. *International Journal of Internet Science*, 5(1), 20-33.
- Friese, S. (2019). *Qualitative data analysis with ATLAS.ti*. Sage.
- Gomes, S., & Duarte, V. (2020). What about ethics? Developing qualitative research in confinement settings. *European Journal of Criminology*, 17(4), 461-479. <https://doi.org/https://doi.org/10.1177/1477370818801305>
- González-Prieto, Á., Perez, J., Diaz, J., & López-Fernández, D. J. a. p. a. (2020). Inter-coder agreement for improving reliability in software engineering qualitative research. 1-37. <https://arxiv.org/pdf/2008.00977.pdf>.
- Hanson, W. E., Creswell, J. W., Clark, V. L. P., Petska, K. S., & Creswell, J. D. (2005). Mixed methods research designs in counseling psychology. *Journal of Counseling Psychology*, 52(2), 224-234 <https://doi.org/https://doi.org/10.1037/0022-0167.52.2.224>
- Hennink, M., Hutter, I., & Bailey, A. (2020). *Qualitative research methods*. Sage.
- Joffe, H., & Yardley, L. (2004). Content and thematic analysis. In *Research Methods for Clinical Health Psychology*. (pp. 56-68). London, Sage.
- Kıral, B. (2020). Document analysis as a qualitative data analysis method. *Siirt University Journal of Social Science Institute*. 8(15), 170-189.
- Kuckartz, U. (2016). Typenbildung und typenbildende Inhaltsanalyse in der empirischen Sozialforschung. In *Junge Menschen sprechen mit sterbenden Menschen* (pp. 31-53). Springer.
- Lavrakas, P. J. (2008). *Encyclopedia of survey research methods*. Sage.

- Lester, J. N., Cho, Y., & Lochmiller, C. R. (2020). Learning to do qualitative data analysis: A starting point. *Human Resource Development Review*, 19(1), 94-106. <https://doi.org/https://doi.org/10.1177/1534484320903890>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills: Sage.
- Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Journal Human communication Research*, 28(4), 587-604. <https://doi.org/https://doi.org/10.1111/j.1468-2958.2002.tb00826.x>
- MacPhail, C., Khoza, N., Abler, L., & Ranganathan, M. (2016). Process guidelines for establishing intercoder reliability in qualitative studies. *Journal Qualitative Research*, 16(2), 198-212. <https://doi.org/https://doi.org/10.1177/1468794115577012>
- MAXQDA. (2020). *MAXQDA 2020 Manual*. 1-736. <https://www.maxqda.com/download/manuals/MAX2020-Online-Manual-Complete-EN.pdf>
- Maxwell, J. (1992). Understanding and validity in qualitative research. *Harvard Educational Review*, 62(3), 279-301. <https://doi.org/https://doi.org/10.17763/haer.62.3.8323320856251826>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.
- Nili, A., Tate, M., & Barros, A. (2017). A critical analysis of inter-coder reliability methods in information systems research. In Riemer, K, Indulska, M, & Tuunainen, V (Eds.) *Proceedings of the 28th Australasian Conference on Information Systems*. University of Tasmania, Australia, 1-11.
- Nili, A., Tate, M., Barros, A., & Johnstone, D. (2020). An approach for selecting and using a method of inter-coder reliability in information management research. *International Journal of Information Management*, 54, 1-13. <https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2020.102154>
- O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: debates and practical guidelines. *International Journal of Qualitative Methods*, 19, 1-13. <https://doi.org/https://doi.org/10.1177/1609406919899220>
- Olson, J. D., McAllister, C., Grinnell, L. D., Gehrke Walters, K., & Appunn, F. (2016). Applying constant comparative method with multiple investigators and inter-coder reliability. *J Qualitative Report*, 21(1). 26-42.
- Rädiker, S., & Kuckartz, U. (2019). *Analyse qualitativer Daten mit MAXQDA*. Springer.
- Richards, T. J., & Richards, L. (1994). Using computers in qualitative research. *Handbook of qualitative research*, 2, 445-462.
- Rolfe, G. (2006). Validity, trustworthiness and rigour: quality and the idea of qualitative research. *Journal of advanced nursing*, 53(3), 304-310. <https://doi.org/https://doi.org/10.1111/j.1365-2648.2006.03727.x>
- Sandelowski, M. (1993). Rigor or rigor mortis: the problem of rigor in qualitative research. *J Advances in Nursing Science*, 16(2), 1-8. <https://doi.org/10.1097/00012272-199312000-00002>

- Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *Journal MIS Quarterly*, 37 (1), 21-54.
- Weitzman, E., & Miles, M. B. (1995). *Computer programs for qualitative data analysis*. Sage.
- Wilson-Lopez, A., Angela Minichiello, P., & Green, T. (2019). An inquiry into the use of intercoder reliability measures in qualitative research. ASEE Annual Conference proceedings, 1-14. [file:///C:/Users/Windows%2010/Downloads/an-inquiry-into-the-use-of-intercoder-reliability-measures-in-qualitative-research%20\(1\).pdf](file:///C:/Users/Windows%2010/Downloads/an-inquiry-into-the-use-of-intercoder-reliability-measures-in-qualitative-research%20(1).pdf)
- Yıldırım, A. (1999). Nitel araştırma yöntemlerinin temel özellikleri ve eğitim araştırmalarındaki yeri ve önemi [Basic features of qualitative research methods and their place and importance in educational research]. *Education and Science*, 23(112), 7-17.