

19 Análisis del acuerdo de inter-codificación

En los proyectos de investigación cualitativa, a menudo surgen preguntas sobre la intersubjetividad del análisis. Dado el mismo pasaje de la entrevista, por ejemplo, ¿mi colega investigador ve los mismos temas tratados que yo, y saca las mismas conclusiones? ¿Hasta qué punto estamos de acuerdo en nuestra comprensión de las categorías? Con estas preguntas entramos en el campo de los criterios de calidad, que no deben descuidarse en la investigación cualitativa. En los enfoques basados en categorías, el enfoque se centra en la cuestión de hasta qué punto dos personas identifican los mismos temas, aspectos y fenómenos en los datos y los asignan a las mismas categorías. Es muy posible que dos personas se pongan de acuerdo en cuanto al contenido, pero asignen diferentes categorías a un fenómeno, porque las definiciones de las categorías aún no se han formulado claramente. MAXQDA ofrece numerosas funciones (parcialmente interactivas), que permiten el análisis sistemático, la mejora y la verificación del acuerdo entre codificadores. Se pueden identificar categorías problemáticas, instrucciones engañosas y definiciones de categorías confusas para poder mejorar la calidad de su análisis paso a paso.

En este capítulo:

- ✓ Conocer los objetivos y áreas de aplicación de las pruebas de acuerdo
- ✓ Entender los procedimientos para probar el acuerdo intercodificador en MAXQDA
- ✓ Realización de pruebas a nivel de documento y de acuerdo por segmentos
- ✓ Cálculo del acuerdo de porcentaje y establecimiento de umbrales útiles
- ✓ Tener en cuenta el acuerdo por casualidad

Objetivos y áreas de aplicación para analizar el acuerdo

La medida en que diferentes investigadores producen el mismo resultado es un criterio de calidad importante de los estudios empíricos. Por lo tanto, la investigación cualitativa implica la cuestión del acuerdo entre codificadores: «¿En qué medida están de acuerdo los diferentes codificadores a la hora de asignar categorías a textos, imágenes y vídeos?» El control y la mejora del acuerdo entre codificadores está estrechamente vinculado a la tradición de investigación del análisis cualitativo de contenidos y se considera un factor central de calidad (Kuckartz, 2014b; Mayring, 2014; Schreier, 2012). En los métodos cualitativos que siguen un enfoque más interpretativo, la cuestión del acuerdo entre codificadores rara vez se plantea, o no se plantea en absoluto, ya que las categorías desempeñan un papel subordinado o incluso ninguno como instrumentos de análisis. Aunque los proyectos de investigación que

siguen el enfoque de la teoría fundamentada implican un trabajo intensivo con códigos y categorías, no es común realizar pruebas de acuerdos entre codificadores en este contexto. El enfoque de la teoría fundamentada se centra explícitamente en el desarrollo continuo de conceptos y categorías y no en su aplicación al material de acuerdo con reglas definidas con precisión.

Entonces, ¿cuál es exactamente el beneficio de un análisis de acuerdo y para qué se utiliza? Bueno, hay muchas respuestas a esto: así como los nuevos empleados en los departamentos de control de Facebook tienen que aprender las reglas para evaluar las contribuciones para eliminar las inapropiadas, también es tarea de los nuevos codificadores en un proyecto de investigación entender las pautas para aplicar categorías. Una prueba de acuerdo con un ejemplo modelo de asignaciones de código revela hasta qué punto el entrenamiento de un codificador ha sido exitoso. El análisis de intercodificadores puede utilizarse no solo para determinar los efectos de la formación de los codificadores, sino también para afinar el sistema de categorías y las instrucciones de codificación. Con la ayuda del análisis de intercodificadores, se pueden identificar y reducir los problemas con las categorías individuales y sus definiciones, se pueden rastrear los problemas relacionados con la delimitación de categorías y, además, se pueden identificar codificadores cuyo trabajo difiere sistemáticamente de los de los demás. Por lo tanto, el análisis de intercodificadores puede ser utilizado igualmente como una herramienta para verificar, mejorar y asegurar la calidad de los procesos de codificación.

Si varias personas deben codificar los mismos datos para el análisis de un acuerdo, surgen inevitablemente preguntas sobre la mejor manera de organizar el flujo de trabajo. En principio, se puede distinguir entre tres variantes separadas:

Codificación en dos momentos diferentes: Aquí los datos son primero codificados por una persona y luego sus asignaciones de código son verificadas por una o más personas posteriormente. Por ejemplo, como primer paso, un estudiante asistente codifica los textos, el director de proyecto examina sus asignaciones de código, hace correcciones y discute los casos dudosos con el asistente o en el equipo de investigación. Este procedimiento solo es adecuado en el caso de que ya existan definiciones de categorías bien desarrolladas. Cuando se trabaja solo, a veces también puede ser útil repetir el proceso de codificación en una etapa posterior; de dos a cuatro semanas después, la misma persona vuelve a comprobar su propio trabajo de codificación o la misma persona vuelve a codificar los datos sin tener en cuenta el trabajo de codificación realizado anteriormente.

Codificación colaborativa simultánea: En el caso de codificación concurrente, los datos pueden ser discutidos a través de todo el equipo de investigación o en pares de dos y las categorías adecuadas pueden ser asignadas colectivamente a los segmentos de datos. Este enfoque es particularmente adecuado para el examen inicial de los datos y el desarrollo o la prueba inicial de los marcos de codificación. Sin embargo, está expuesto al riesgo de la influencia de las jerarquías de los equipos de investigación o del dominio de personalidades particularmente extrovertidas.

Codificación independiente simultánea: El método utilizado con más frecuencia consiste en que los investigadores codifican los datos independientemente unos de otros y luego comparan sus resultados. La codificación independiente es generalmente obligatoria para

el cálculo del porcentaje de acuerdo y para los coeficientes de acuerdo con corrección del azar, como Kappa de Cohen o Alfa de Krippendorff (véase más adelante). Sin embargo, en nuestra opinión, es importante y deseable para la investigación cualitativa que se calculen y publiquen algo más que coeficientes de acuerdo. Más bien, deberían, junto con los lugares en los que se han producido incoherencias, constituir la base de un debate sistemático sobre las incoherencias y sus consecuencias para el sistema de categorías y las instrucciones de codificación. Basándonos en esta aspiración a los análisis cualitativos, en este capítulo preferimos el término «acuerdo de intercodificación» al de «fiabilidad de intercodificación». La fiabilidad es uno de los tres criterios clásicos de calidad de la investigación cuantitativa, representa el derecho a la precisión y la replicabilidad de las mediciones y se sitúa principalmente en el contexto del análisis cuantitativo del contenido. Además, la transferibilidad de los criterios de calidad clásicos a la investigación cualitativa debe ser cuestionada críticamente (Kuckartz, 2014b, pp. 151-155).⁴

Es sorprendente que la codificación independiente simultánea sea particularmente adecuada para llevar a cabo análisis de acuerdos *sistemáticos*. Las dos primeras variaciones de análisis anteriores representan otras formas de aumentar la calidad de la codificación y pueden combinarse de forma provechosa con la tercera.

MAXQDA ofrece funciones especialmente desarrolladas que apoyan la determinación del acuerdo entre codificadores, el control de los desacuerdos y la mejora del acuerdo. Describiremos estas funciones con respecto a diferentes codificadores y por lo tanto usaremos el término «intercodificador» a lo largo de este capítulo. Sin embargo, estos métodos también se pueden utilizar en el caso de codificación repetida por una misma persona, lo que siempre es útil si desea analizar la estabilidad de su propio trabajo de codificación, el llamado «acuerdo intracodificador».

Antes de que pueda empezar a analizar el acuerdo intercodificador en MAXQDA, primero debe clarificar para qué documentos (y en qué orden) se debe llevar a cabo el análisis. Si la cantidad de datos a analizar es pequeña, todos los documentos pueden ser codificados por una segunda persona. Este es el caso, por ejemplo, si ha realizado diez entrevistas de media hora que han sido codificadas temáticamente usando un sistema de categorías simple. La mayoría de las veces, sin embargo, los datos serán más extensos que eso, en cuyo caso necesitará una muestra. A veces se sugiere en la literatura sobre el tema que un cierto porcentaje de los datos debe ser codificado por una segunda persona. Este porcentaje del 10% puede proporcionar una indicación inicial, pero debido a las cantidades de datos potencialmente muy diferentes y a la diversidad de condiciones entre los proyectos, deberían incluirse definitivamente otros criterios en la decisión:

- ❖ El número esperado de segmentos codificados — por ejemplo, no tiene sentido limitarse a muy pocos documentos si solo espera que se utilicen algunas de las categorías disponibles en los documentos seleccionados.

4 Krippendorff distingue sistemáticamente entre acuerdo y fiabilidad en el contexto del análisis de contenido (en su caso, de orientación más bien clásica): «Para ser claros, el acuerdo es lo que medimos; la fiabilidad es lo que queremos inferir de él. En el análisis de contenido, la reproducibilidad es posiblemente la interpretación más importante de la fiabilidad» (2004, p. 414).

- ❖ La diversidad de casos — la muestra debe incluir un amplio espectro de datos disponibles. Se recomienda una selección bien pensada de documentos según el principio del máximo contraste (por ejemplo, textos cortos frente a textos largos o entrevistas con personas entrevistadas que cuentan historias frente a personas que contestan brevemente) o una selección aleatoria.
- ❖ La etapa de desarrollo del sistema de categorías — especialmente cuando se aplica por primera vez un marco de codificación de nuevo desarrollo, debe iniciarse el análisis del acuerdo intercodificador relativamente pronto para poder detectar deficiencias en el marco de codificación.
- ❖ Los recursos disponibles — los análisis intercodificadores llevan tiempo. Los proyectos a menudo se llevan a cabo bajo presión de tiempo y no siempre hay personas disponibles que estén dispuestas y sean capaces de realizar una segunda ronda de codificación. Sin embargo, cuando se realiza un análisis cualitativo de contenido, nunca se debe prescindir por completo de una comprobación de la intercodificación. A veces se sobreestima la cantidad de esfuerzo que implica, pero el lema «un poco es mejor que nada» se aplica claramente aquí.

En general, es mejor verificar la consistencia de la codificación con un conjunto de datos manejables en una etapa temprana para evitar descubrir, al final del proyecto, que las instrucciones de codificación estaban incompletas o eran engañosas. Para poder informar sobre mejoras en el proceso de codificación, es necesario realizar verificaciones múltiples, es decir, empezar con dos documentos muy diferentes, discutir cualquier inconsistencia que se haya producido y, a continuación, continuar con otros dos documentos.

Antes de iniciar el análisis en MAXQDA, tiene sentido no solo pensar en la selección de los documentos sino también en los códigos que desea incluir. Cuando se analiza el acuerdo intercodificador, normalmente no tiene sentido comprobar todos los códigos al mismo tiempo. En primer lugar, códigos como «Pasaje de texto interesante», «Cita adecuada» y potencialmente también el código «Otro» se excluyen a menudo del análisis. Luego, el analista que ha llevado a cabo la codificación en dos pasos —primero aplicando temas de gran alcance antes de diferenciarlos posteriormente— debe proceder con el análisis del acuerdo en dos etapas equivalentes: primero para los segmentos de gran alcance, seguido de un análisis más detallado de las asignaciones de subcódigos. Además, deben tenerse en cuenta los tipos de códigos asignados. Los códigos fácticos simples (por ejemplo, si una persona afirma ser partidario de una parte o no) solo deben mezclarse con códigos sofisticados en el contexto de un análisis argumentativo complejo si no está interesado en el cálculo de un valor total de acuerdo.

El procedimiento para analizar el acuerdo intercodificador en MAXQDA

MAXQDA le permite determinar el acuerdo entre dos codificadores para los documentos seleccionados. Para realizar el análisis en MAXQDA, los documentos deben existir dos veces en el proyecto —una vez codificados por la persona 1 en un grupo o conjunto de documentos, y otra vez codificados por la persona 2 en otro grupo o conjunto de documentos. Ade-

más, se debe haber definido previamente un marco de codificación y las instrucciones correspondientes para los codificadores. ¿Cuál es la mejor manera de organizar el proceso de análisis de acuerdo? Los siguientes pasos ilustran un procedimiento apropiado:

- ❖ Paso 1 — Crear un proyecto con todos los documentos relevantes y el sistema de códigos completo. Especifique las instrucciones de codificación para categorías individuales en las notas de códigos. Si las unidades que ambas personas deberían codificar han sido definidas de antemano, pueden ser etiquetadas con un código llamado «Segmentos a codificar». Si el material de datos tiene una estructura uniforme, puede ser suficiente instruir a ambos codificadores para que siempre codifiquen el párrafo completo o la respuesta completa a una pregunta.
- ❖ Paso 2 — Activar protecciones de escritura para el sistema de código y textos de documento en el sistema de Administración de Usuario de MAXQDA en el caso de cada codificador para proteger el proyecto de cambios no deseados (ver Capítulo 18). La selección de documentos que deben ser codificados por una segunda persona puede hacerse visible fácilmente, por ejemplo, asignando un determinado color al documento en la ventana «Sistema de documentos».
- ❖ Paso 3 — Proporcione una copia del archivo del proyecto maestro a ambos codificadores.
- ❖ Paso 4 — Ambas personas codifican los documentos seleccionados y añaden sus abreviaturas o nombres al nombre de un grupo de documentos (o conjunto) para poder identificar sus asignaciones de código más tarde.
- ❖ Paso 5 — Fusionar los dos archivos de proyecto en uno, que contendrá dos copias de todos los documentos a comparar, utilizando la función **Inicio > Fusionar proyectos** descrita en el Capítulo 18. Seleccione aquí la opción **No importar documentos existentes**, de modo que solo se añadan los documentos codificados por la segunda persona. Una vez completada la importación, los nombres de los documentos indicarán quién ha codificado qué documento.

Una vez que un proyecto contiene los documentos seleccionados que han sido codificados por ambas personas, el análisis del acuerdo intercodificador se puede hacer *para dos documentos o conjuntos*. Para iniciar el procedimiento, seleccione primero los códigos que desea incluir en el análisis activándolos y, a continuación, abra la función **Análisis > Acuerdo de inter-codificación**. El cuadro de diálogo que aparece (Fig. 19.1) permite especificar los dos documentos que se desean analizar y distinguir entre tres clases de acuerdos.

MAXQDA puede verificar la codificación consistente de los documentos seleccionados en tres niveles diferentes, donde las dos primeras clases de consistencia se refieren al «nivel de documento» y la tercera al «nivel de segmento»:

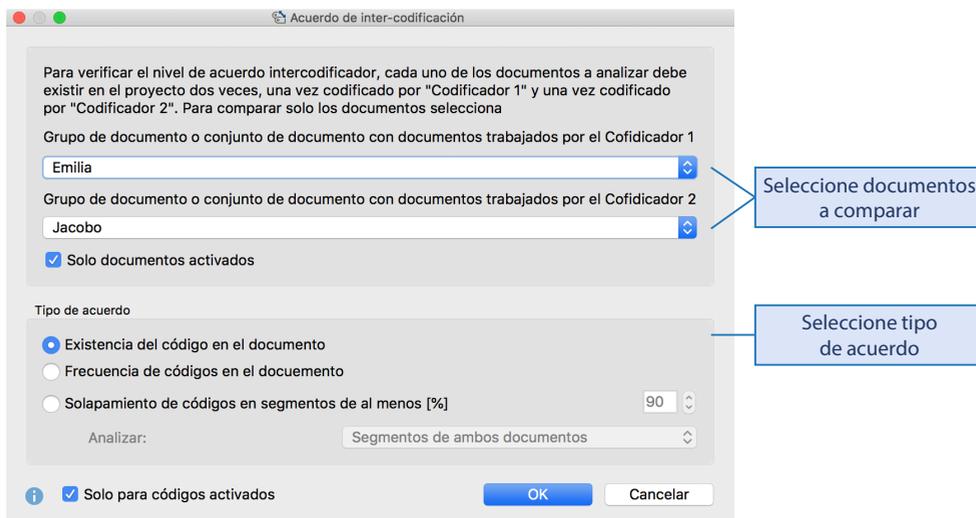


Fig. 19.1: Cuadro de diálogo de opciones para el análisis de acuerdo intercodificador

- ❖ **Existencia del código en el documento** — Se cuenta una coincidencia si ambos codificadores han asignado el mismo código al documento. No importa en este caso si una persona asignó el código tres veces y la otra solo una vez. La ubicación del código en el documento también es irrelevante, siempre y cuando el código exista en algún lugar del documento. Por lo tanto, el desacuerdo con respecto a un código determinado solo se produce si una persona ha asignado el código *una o varias veces* y la otra persona *no lo ha asignado en absoluto* en el documento. Este nivel de verificación de acuerdo es interesante, por ejemplo, para categorías que se refieren a todo el documento. Si en un estudio con pacientes de rehabilitación se va a asignar un código «Experiencia previa en rehabilitación», puede que no importe dónde y con qué frecuencia se asignó el código en el documento — el punto principal es que ambos codificadores lo *han* asignado.
- ❖ **Frecuencia de códigos en el documento** — Se cuenta una coincidencia si ambos codificadores han asignado el mismo código en el documento el mismo número de veces. Si una persona ha asignado el código «autoconfianza» tres veces en el documento y la otra solo dos veces, no hay acuerdo para este código. De nuevo, las ubicaciones de los segmentos codificados en el documento no juegan ningún papel aquí.
- ❖ **Solapamiento de códigos en segmentos de al menos [%]** — Se cuenta una coincidencia si ambos codificadores han asignado el mismo código a un segmento de datos dado. Los segmentos no tienen que ser cien por cien idénticos en su posición; se puede fijar un intervalo de tolerancia.

Para analizar correctamente el acuerdo intercodificador en MAXQDA, los documentos comparados deben ser idénticos. Si este no es el caso, MAXQDA muestra un mensaje de advertencia que incluye una referencia a la primera ubicación donde los documentos difieren. En

este caso, recomendamos abandonar el procedimiento para examinar primero las diferencias entre cada documento para evitar resultados incorrectos.

El análisis del acuerdo se puede llevar a cabo para todos los tipos de documentos, es decir, para textos, archivos PDF, tablas e imágenes, así como para archivos de audio y vídeo.

Acuerdo intercodificador a nivel de documento

En la Fig. 19.2 se puede ver un ejemplo del resultado de un análisis intercodificador. Camilla (codificador 1) e Isabel (codificador 2) codificaron independientemente la entrevista 1 y luego se realizó una comprobación de intercodificación utilizando la segunda opción anterior, es decir, cuando se cuenta una coincidencia cuando el código ha sido asignado el mismo número de veces. Para comparar sus asignaciones de código, MAXQDA utiliza el Visualizador de la matriz de códigos, en el cual los dos documentos se muestran uno al lado del otro en dos columnas. En la Fig. 19.2, se activa la visualización de los números para proporcionar una comparación precisa: por ejemplo, tanto Isabel como Camilla han asignado el código «Intereses» dos veces, lo que significa que existe un acuerdo de intercodificación para este código. En la barra de título, MAXQDA muestra el acuerdo porcentual entre los dos codificadores en relación a todos los códigos mostrados en el Visualizador de la matriz de códigos. El valor del porcentaje se calcula por el número de códigos para los cuales existe un acuerdo entre codificadores dividido por el número de todos los códigos analizados. Esto es lo mismo que el número de líneas con concordancias dividido por el número total de líneas cuando se desglosan todos los códigos. El acuerdo del 60% que se muestra en el título del ejemplo se calcula de la siguiente manera: para tres códigos («Intereses», «Dinero y Asuntos Financieros», «Religión y Espiritualidad») los dos codificadores están de acuerdo, para los otros dos códigos no lo están. El número de códigos coincidentes dividido por el número total de códigos es $3 / (3+2) = 60\%$. Este valor indica que los dos codificadores están de acuerdo en el 60% de los códigos y que difieren en el 40% de los códigos. Por cierto, la variación de análisis «existencia de código en el documento» resultaría en un porcentaje de acuerdo del 100% entre estos dos codificadores, ya que ambos están completamente de acuerdo en que los mismos cuatro códigos se aplican a este documento y uno no.

Por favor, tenga en cuenta: Por defecto, solo los códigos activados se muestran en una lista lineal sin sus códigos principales. Al seleccionar la opción *Mostrar códigos con jerarquía* en la barra de herramientas, también se pueden listar los códigos principales no activados que son necesarios para la correcta visualización del árbol de códigos. El cálculo del acuerdo porcentual ignora estos códigos.

Normalmente, los aspectos interesantes de un análisis intercodificador son los desacuerdos, que pueden revelar problemas relacionados con los códigos individuales, las instrucciones de codificación o los enfoques de los codificadores individuales. MAXQDA soporta el análisis de los desacuerdos a través de la visualización interactiva de los resultados: haciendo doble clic en una celda se listan todos los segmentos codificados correspondientes

en la ventana «Segmentos recuperados». Así, en el ejemplo, se puede ver qué segmento ha codificado Camilla con «Emociones», pero no Isabel.

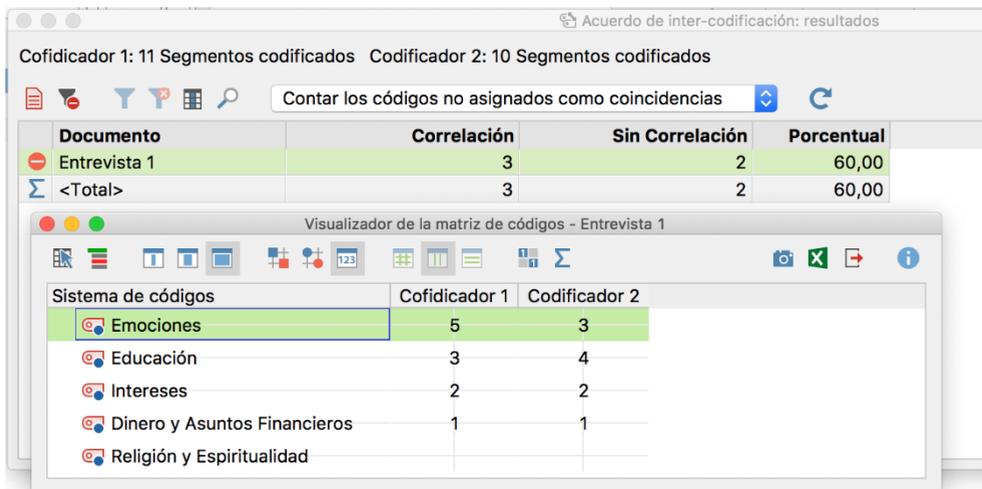


Fig. 19.2: Tablas de resultados de un análisis de acuerdo intercodificador a nivel de documento

Acuerdo entre codificadores a nivel de segmento

Aunque un análisis a nivel de documento proporciona indicaciones iniciales de diferencias sistemáticas en el uso de las categorías, en la mayoría de los casos será necesaria una comprobación del acuerdo por segmentos específicos. Puede hacerlo seleccionando el tercer tipo de acuerdo en el cuadro de diálogo de opciones (*Solapamiento de códigos en segmentos de al menos [%]*, Fig. 19.1). En la práctica, los codificadores a menudo no solo asignan categorías a segmentos predefinidos, sino que de hecho definen los segmentos que necesitan ser codificados a medida que avanzan. En otras palabras, si se adopta este último enfoque, a menudo puede ocurrir que una persona codifique un carácter o una palabra extra en comparación con la otra, y es igual de probable que dos codificadores difieran en uno o dos segundos al codificar la misma escena en un vídeo. Para ignorar estas diferencias menores e intrascendentes en los límites de los segmentos durante el análisis intercodificador, puede establecer un solapamiento mínimo de dos segmentos codificados que se compararán en el cuadro de diálogo de opciones de un valor porcentual. Una superposición de código de 100% significa que los límites de los segmentos deben coincidir con precisión para ser contados como segmentos idénticos en MAXQDA.

Tenga en cuenta: El porcentaje establecido en el cuadro de diálogo de opciones no debe confundirse con el acuerdo porcentual que se muestra como resultado del análisis. El cuadro de diálogo de opciones solo permite establecer el solapamiento mínimo para el que dos segmentos codificados se consideran idénticos.

Debe especificar un solapamiento mínimo cada vez que ejecute este análisis. Se debe establecer un valor de 100% si los codificadores trabajan en segmentos predefinidos. Este sería el caso si, por ejemplo, las instrucciones de codificación fueran a codificar cada párrafo de un texto, o si todos los segmentos a codificar fueran previamente etiquetados con un código específico y luego asignados a códigos temáticos. En la mayoría de los casos, sin embargo, se le aconsejaría comenzar con una superposición mínima de aproximadamente el 95% como prueba y reducir este valor paso a paso si resulta en un número excesivo de desacuerdos *insignificantes*.

Después de iniciar un análisis de acuerdo intercodificador a nivel de segmento, MAXQDA procesa primero las asignaciones de código del primer documento y luego las del segundo documento. Cada segmento codificado se comprueba para ver si la otra persona ha asignado el mismo código a este segmento. Suponiendo que una persona haya codificado 10 segmentos en el documento y la otra persona haya codificado 12 segmentos, se verifica el acuerdo entre $10 + 12 = 22$ segmentos. La Fig. 19.3 muestra parte de la tabla de resultados que MAXQDA produce para este análisis. Cada fila contiene el resultado de la verificación de un único segmento, donde las marcas de verificación verdes representan acuerdos y las señales de *stop* representan desacuerdos. En la primera fila se proporciona la siguiente información: en el párrafo 5, codificador 2 asignó el código «Emociones», pero codificador 1 no lo hizo. Por consiguiente, falta la marca de verificación en la columna «Acuerdo» y en la primera columna aparece una señal de *stop*.

Consejo: Haciendo clic en el encabezamiento de la columna «Acuerdo», se puede ordenar la tabla para esa columna, de modo que todos los desacuerdos se muestren en la parte superior de la tabla.

Documento	Código	Codificador 1	Codificador 2	Acuerdo	Principio	Final
Entrevista 3 [Codificador 2]	Temas\Emociones	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	5	5
Entrevista 3 [Codificador 1]	Temas\Intereses	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	8	8
Entrevista 3 [Codificador 2]	Temas\Intereses	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	8	8
Entrevista 3 [Codificador 2]	Temas\Emociones	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	16	16
Entrevista 3 [Codificador 1]	Temas\Educación	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	16	16
Entrevista 3 [Codificador 2]	Temas\Educación	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	16	16
Entrevista 3 [Codificador 1]	Temas\Educación	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	22	22
Entrevista 3 [Codificador 2]	Temas\Educación	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	24	24
Entrevista 3 [Codificador 1]	Temas\Educación	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	26	27
Entrevista 3 [Codificador 2]	Temas\Educación	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	26	27
Entrevista 3 [Codificador 1]	Temas\Educación	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	28	29

Fig. 19.3:Tabla de resultados para un análisis de acuerdos de intercoder a nivel de segmento (tabla de segmentos)

En la investigación cualitativa en general, especialmente al principio del proceso de análisis, se investigan todos los desacuerdos y se identifican sus causas. ¿Se le pasó algo a la segunda persona? ¿Deberían separarse o incluso fusionarse más claramente dos códigos? ¿Asignaron

las dos personas el mismo código, pero divergieron en dónde colocar los límites del segmento? La interactividad de la tabla de resultados le ayuda a identificar estos y otros problemas similares. Haciendo doble clic en las columnas «Cofidicador 1» o «Cofidicador 2» se abren los documentos respectivos en la ubicación correspondiente en el «Visualizador de documento» y permite examinar las diferencias entre los codificadores.

La cuestión de cómo tratar rápidamente los desacuerdos se plantea a continuación. Como hemos enfatizado anteriormente, los proyectos de investigación cualitativamente orientados buscan utilizar estas diferencias como una ocasión para discutir sobre las asignaciones de códigos, el marco de codificación y los segmentos del documento. Para llegar a un acuerdo sobre los segmentos codificados problemáticos, a veces es útil incluir el caso completo como información contextual o investigar otros segmentos codificados bajo la misma categoría. Por regla general, vale la pena registrar este proceso de resolución de problemas y consolidación, así como los argumentos y puntos de vista presentados, ya que estas discusiones a menudo pueden dar lugar a información valiosa. No es raro que se puedan hacer descubrimientos relevantes muy interesantes para su proyecto de investigación y el proceso de análisis en su conjunto mientras resuelve el problema de definir y demarcar categorías —incluso si estas son solo hipótesis que necesitan ser probadas con datos adicionales.

Para optimizar los segmentos codificados en el archivo de proyecto MAXQDA, uno de los dos documentos debe ser definido como el «documento maestro». En este documento, las asignaciones de código pueden mejorarse cuando sea necesario, para que contenga la versión optimizada. Una vez finalizado el análisis intercodificador, puede eliminar la segunda versión de cada documento para seguir trabajando en el proyecto optimizado y fusionado.

Tabla de resultados específicos por código

MAXQDA no solo muestra la tabla de resultados de segmentos, sino también la llamada tabla de resultados específicos de código. Esta última muestra no solo el porcentaje de acuerdo entre todos los códigos, sino también para cada código individualmente (Fig. 19.4). El porcentaje de acuerdo resulta de la proporción de acuerdos en el código evaluado. En la Fig. 19.4, en la fila «<Total>», con 18 acuerdos y un total de 22 segmentos codificados, el porcentaje total es del 81,82%.

El acuerdo de porcentaje proporciona información valiosa para la identificación de códigos problemáticos, pero debe evaluar su tamaño con un buen grado de precaución. Debe prestarse especial atención al número absoluto de segmentos evaluados por código en la columna «Total» y al número de desacuerdos. Por ejemplo, si cada codificador creara cinco segmentos, un solo desacuerdo ya reduciría el porcentaje de acuerdo en 20 puntos porcentuales. Y con dos segmentos (como con el código «Dinero y Asuntos Financieros» en la Fig. 19.4) la tabla solo incluiría los valores 0% y 100%. Además, adicional al número absoluto, también debe considerarse el número total de segmentos por codificador. Si es muy desigual, usualmente habrá una diferencia sistemática entre los codificadores, por ejemplo porque uno de ellos codificó los datos en partes más pequeñas y desglosadas que el otro, lo que llevará a un pequeño porcentaje de acuerdo.

Código	Correlación	Sin Correlación	Total	Porcentual
Temas\Emociones	6	2	8	75,00
Temas\Educación	6	2	8	75,00
Temas\Intereses	4	0	4	100,00
Temas\Dinero y Asuntos Financieros	2	0	2	100,00
<Total>	18	4	22	81,82

Fig. 19.4: Tabla de resultados de un análisis de acuerdo intercodificador a nivel de segmento (tabla de códigos)

Si se incluye el acuerdo porcentual para un estudio cualitativo en un informe o publicación, siempre se debe indicar también qué segmentos causaron las diferencias restantes y qué relevancia tienen para el estudio. No es particularmente útil presentar solo el «porcentaje global» para todos sus códigos. Por el contrario, podría ser mejor proporcionar al menos los porcentajes mínimo y máximo con respecto a todos los códigos probados o, mejor aún, incluir toda la información en la tabla de resultados por códigos específicos de la publicación. Esto es aconsejable porque, por las razones mencionadas anteriormente, también debe indicarse el número absoluto de códigos y desacuerdos evaluados. Por lo tanto, durante el análisis es necesario documentar los códigos problemáticos y las discrepancias para poder incluirlos posteriormente en los informes y publicaciones.

La cuestión sigue siendo qué debe considerarse como un porcentaje bajo y qué como un porcentaje alto de acuerdo. Desafortunadamente, esta pregunta no puede ser contestada con los umbrales establecidos, porque el acuerdo porcentual entre codificadores no solo depende del número de segmentos codificados absolutos como se ha descrito, sino también de otros factores. Estos incluyen, en particular, el número y la varianza de las diferentes (sub)categorías, así como el grado de dificultad del proceso de codificación en sí (es más difícil codificar el argumento de un entrevistado, por ejemplo, que asignar códigos fácticos o temáticos). Una simple inversión de la conclusión normalmente le ayudará a evaluar si el porcentaje es demasiado bajo: si un código tiene un acuerdo del 80%, por ejemplo, esto significa que el 20% de las asignaciones de código difieren. Por lo general, este nivel de desacuerdo no se consideraría ideal, y sería necesario establecer umbrales más estrictos. Sin embargo, el modo de determinar los valores adecuados debe estar siempre relacionado con el contenido de los datos.

Unidades de codificación vs. segmentos codificados

Cuando se realiza un análisis intercodificador usando MAXQDA, es importante distinguir entre unidades de codificación y segmentos codificados. En muchos enfoques para el análisis de los acuerdos entre codificadores, se supone que se asigna un código único a cada unidad codificadora. Este es el caso, por ejemplo, si los médicos califican una imagen de rayos X por la presencia de una enfermedad o si artículos de periódico enteros se califican

de bajos, medianos o altos en una escala de «racismo latente». Cada codificador habrá asignado exactamente un código a cada unidad codificadora.

Este no es el caso en muchos análisis realizados con MAXQDA. Es posible que los dos codificadores hayan asignado diferentes cantidades de códigos a los mismos pasajes del documento. Además, es común que las unidades de codificación no hayan sido predefinidas, y los codificadores a menudo deciden sobre qué pasajes codificar por sí mismos, lo que puede resultar en segmentos codificados superpuestos o segmentos codificados contenidos dentro de otros segmentos. Para realizar un análisis intercodificador en estos casos, MAXQDA sigue el mismo proceso descrito anteriormente: primero se comprueban los códigos del codificador 1, luego los del codificador 2, para verificar si están de acuerdo con la otra persona, respectivamente. La Fig. 19.5 muestra algunos ejemplos que ilustran cuántas coincidencias resultan de este procedimiento en diferentes casos. La primera fila, en la que ambos codificadores asignaron solo el «Código A» al mismo segmento, indica que se han contado 2 acuerdos. El caso donde los codificadores asignaron un código al segmento cada uno, pero estos códigos no son los mismos, resulta en 2 desacuerdos, como se muestra en la segunda fila.

	Segmentos codificados	Acuerdos	Desacuerdos	Porcentaje
A Segmento A	2	2	0	100%
A Segmento B	2	0	2	0%
B, A Segmento A, B	4	4	0	100%
B, A Segmento A	3	2	1	67%
B, A Segmento C	3	0	3	0%
B, A Segmento C, D	4	0	4	0%

Fig. 19.5: Número de acuerdos, en diferentes constelaciones, entre dos codificadores

Cuando usted abre la función de **Acuerdo de inter-codificación**, un menú desplegable llamado **Analizar**: es visible en la caja de diálogo de opciones que se abre (Fig. 19-1), donde usted puede instruir a MAXQDA para que analice solamente los segmentos codificados en un solo documento, es decir, el documento codificado por uno solo de los codificadores, para evaluar su nivel de acuerdo con el otro codificador. Este ajuste puede ser usado donde ambos codificadores han codificado segmentos predefinidos con un solo código. En este caso, el número de segmentos codificados evaluados corresponde al número de unidades de codificación de la tabla de resultados.

Cálculo de coeficientes de acuerdo con corrección del azar como Kappa

«Mi supervisor me dijo que calculara Kappa. ¿Cómo hago esto en MAXQDA? «Necesito Kappa para mi análisis, ¿no?» Estas preguntas no solo están en la mente de los muchos doc-

torandos a los que asesoramos en los talleres, sino que también se formulan regularmente en el foro de discusión en el sitio web de MAXQDA. Esto se basa a menudo en el deseo de legitimar el propio enfoque cualitativo de una comunidad científica que se adhiere principalmente a la tradición de la investigación cuantitativa. Aquí kappa es un coeficiente conocido, por lo que los supervisores lo exigen a menudo. Un coeficiente como kappa cuantifica la calidad del análisis cualitativo como una figura y traduce el trabajo en una forma comprensible (y familiar). El cálculo de tales medidas ayuda a mantenerle conectado con algunas comunidades científicas y aumenta la posibilidad de publicación en ciertas revistas.

Incluso si se pusiera en duda un énfasis excesivo en un coeficiente con corrección del azar a expensas de otros criterios de calidad importantes para la investigación cualitativa, el cálculo de los coeficientes con corrección del azar sí tiene sentido. Al determinar el acuerdo porcentual, se plantea la cuestión de la probabilidad de que ese acuerdo pudiera haber surgido por casualidad. Para responder a esta pregunta, se desarrollaron coeficientes con corrección del azar que restan posibles coincidencias aleatorias del acuerdo en bruto. Su idea central consiste en determinar hasta qué punto los codificadores humanos pueden codificar un texto o vídeo, con un conjunto existente de categorías, mejor que una máquina que funcione al azar.

La siguiente fórmula básica se utiliza a menudo para calcular los coeficientes corregidos por azar: $(Po - Pc) / (1 - Pc)$, donde Po es el porcentaje observado de acuerdo y Pc es el acuerdo esperado por casualidad. Siguiendo este método de cálculo, el valor resultante indica hasta qué punto el acuerdo de los codificadores excede el acuerdo por casualidad. Coeficientes como Kappa (Cohen, 1960), Pi (Scott, 1955) y Alpha (Krippendorff, 1970) difieren principalmente en cómo se calcula el acuerdo con corrección del azar. Normalmente una matriz «categorías por categorías», como se muestra esquemáticamente en la Tabla 19.1, se utiliza como base para el cálculo de estos coeficientes. Las celdas indican con qué frecuencia los dos codificadores han asignado las categorías respectivas a una unidad codificadora. La diagonal primaria contiene las celdas con los acuerdos entre los dos codificadores (las celdas a, e, i). El acuerdo por azar se calcula utilizando las sumas marginales. Para las siguientes explicaciones es importante enfatizar que generar una tabla de este tipo usualmente requiere que los codificadores hayan asignado solo una categoría a cada segmento.

Tab. 19.1: Tabla «categorías por categorías»; los acuerdos se encuentran en la línea diagonal (a, e, i)

		Codificador 2			
Codificador 1	Cat. 1	Cat. 2	Cat. 3	SUMA	
Cat. 1	a	b	c	a + b + c	
Cat. 2	d	e	f	d + e + f	
Cat. 3	g	h	i	g + h + i	
SUMA	a + d + g	b + e + h	c + f + i	N	

Aquí nos limitaremos al cálculo de coeficientes con corrección del azar para acuerdos a nivel de segmento, ya que rara vez se determinan a nivel de documento y no se pueden realizar automáticamente en MAXQDA. Para solicitar el cálculo de un coeficiente corregido por azar a nivel de segmento en MAXQDA, haga clic en el símbolo Kappa en la tabla de resultados (Fig. 19.3) después de realizar el análisis intercodificador. MAXQDA entonces genera un conjunto de resultados como se muestra en la Fig. 19.6.

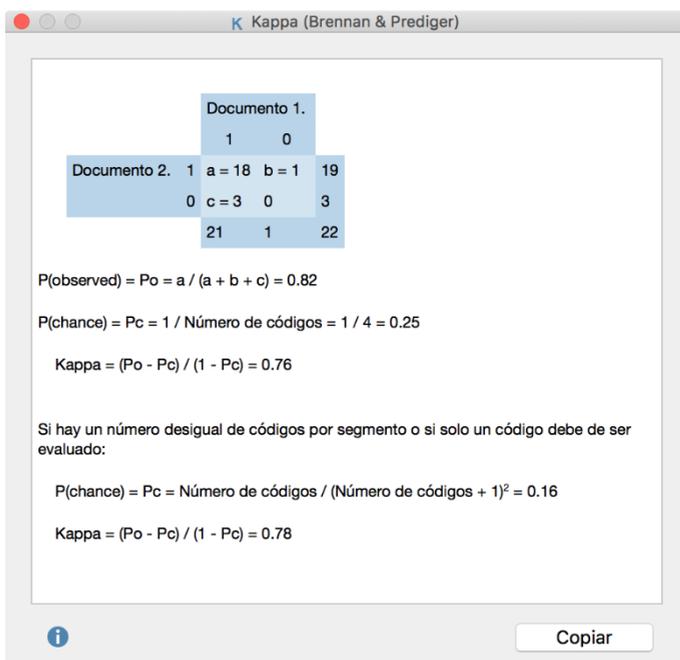


Fig. 19.6: Ventana de resultados para el cálculo de Kappa según Brennan y Prediger (1981)

¿Qué información se incluye en los resultados que produce MAXQDA y cómo se puede interpretar? En la práctica, en los proyectos de investigación cualitativa, raramente se podrá crear una matriz como se muestra en la Tabla 19.1, ya que los codificadores a menudo asignan más de un código a un segmento y, además, también sucede que una persona asigna uno o más códigos a un segmento donde la otra persona no asigna ninguna categoría en absoluto. Esta es la razón por la cual MAXQDA pasa por los segmentos codificados de un codificador, y luego los segmentos del otro, como se describe arriba, y los cuenta como coincidentes si la otra persona ha asignado el mismo código al mismo segmento. Para determinar un coeficiente con corrección del azar para este procedimiento, MAXQDA genera una tabla 2x2 como se muestra en la Fig. 19.6. La celda superior izquierda (a) indica con qué frecuencia los dos codificadores han asignado el mismo código a un segmento. La celda superior derecha (b) y la inferior izquierda (c) indican para cuántos segmentos los dos codifi-

codadores difieren en su asignación. La cuarta celda (d) es siempre cero, porque debido al método utilizado no hay segmentos predefinidos que no hayan sido codificados por ninguno de los dos codificadores.

Debajo de esta tabla, se calculan tanto el acuerdo en bruto P_o como el acuerdo esperado por casualidad P_c . P_o corresponde a la salida de valores de la fila «<Total>» de la tabla de códigos específicos. Puesto que las distribuciones marginales de la tabla 2x2 están siempre desigualmente distribuidas debido a la celda $d = 0$, el acuerdo por casualidad no puede calcularse como para el Kappa de Cohen, el Pi de Scott o el Alfa de Krippendorff. Las distribuciones marginales desiguales pueden llevar a valores abstrusos y paradójicos en la Kappa de Cohen, que es un punto de crítica frecuentemente articulado (p. ej. Feinstein & Cicchetti, 1990; o Gwet, 2008). Al calcular P_c , el cálculo en MAXQDA sigue por lo tanto un concepto propuesto por Brennan y Prediger (1981)⁵. En lugar de determinar el acuerdo esperado por casualidad utilizando la distribución marginal, se utiliza aquí el número de categorías. Se puede ver fácilmente que la probabilidad de acuerdo disminuye a medida que aumenta el número de categorías. Computacionalmente, P_c es $1/n$, donde n corresponde al número de categorías utilizadas. Esto se ilustra gráficamente en la tabla de la izquierda en la Fig. 19.7: el número de celdas grises con coincidencias en la diagonal principal corresponde al número de categoría n , el área total de la tabla corresponde a las celdas n^2 , resultando en una coincidencia aleatoria de $n/n^2 = 1/n$.

	Cat. 1	Cat. 2	Cat. 3
Cat. 1			
Cat. 2			
Cat. 3			

	Cat. 1	Cat. 2	Cat. 3	X
Cat. 1				
Cat. 2				
Cat. 3				
X				0

	Cat. 1	X
Cat. 1		
X		0

Fig. 19.7: Determinación del acuerdo por casualidad (gris); «X» significa «no codificado»

En el caso frecuente de que los codificadores difieran en el número de categorías asignadas por segmento, el acuerdo aleatorio esperado puede ser ligeramente corregido hacia abajo. Como muestra la tabla central de la Fig. 19.7, en este caso se añade otra categoría «X», que representa «no codificado». Dado que el número de celdas con coincidencias (por casualidad) sigue correspondiendo al número de categoría n , P_c se calcula ahora con $n/(n+1)^2$. MAXQDA también muestra el valor calculado de esta manera. En el ejemplo, es 0,78, ligeramente superior al valor Kappa «normal» de 0,76. Como ilustra la tabla del extremo derecho de la Fig. 19.7, esta forma de cálculo también debe utilizarse si solo se evalúa un único código para el análisis de acuerdo.

5 Krippendorff (2004, p. 417) señala que esta propuesta ya fue formulada por Bennett et al. en 1954 y que luego fue «reinventada» con ligeras variaciones por varios autores, entre ellos Brennan y Prediger (1981).

Nota: MAXQDA no proporciona Kappa de Cohen, sino Kappa según Brennan y Prediger (1981), quienes nombraron su coeficiente con una kappa griega con el subíndice n: K_n . Si utiliza los resultados de un cálculo de MAXQDA en una publicación, debe hacer referencia a Brennan y Prediger para evitar confusiones.

Claramente, a medida que aumenta el número de categorías, el acuerdo por casualidad calculado de esta manera disminuirá. Supongamos que hay un 90% de coincidencia. Entonces la corrección del azar para dos categorías lleva a una Kappa de Brennan y Prediger de 0,80 y para diez categorías Kappa sería tan alta como 0,89.

Pero, ¿cómo se debe evaluar el nivel de Kappa? La Kappa de Brennan y Prediger puede tomar valores entre -1.00 y +1.00; un valor de 0 corresponde a una paridad con el azar, un valor de +1.00 corresponde a un acuerdo perfecto de los codificadores —esto es lo más lejos que se puede llegar de un acuerdo por azar. 1.00 se alcanza si el porcentaje de acuerdo entre los dos codificadores es del 100%. La interpretación del valor puede basarse en las notas de referencia establecidas para la Kappa de Cohen: de acuerdo con Landis y Koch (1977) uno puede etiquetar un resultado como bueno («sustancial») desde 0,61 y como muy bueno desde 0,81 («casi perfecto»). Sin embargo, cualquier umbral de este tipo podría ser engañoso. En primer lugar, en muchos casos la Kappa de Cohen nunca puede alcanzar el valor de 1,00 debido a su método de cálculo, razón por la cual los valores umbral para la Kappa de Brennan y Prediger podrían aumentar, pero nunca deberían reducirse. En segundo lugar, como se ha explicado anteriormente para el acuerdo porcentual, la definición de este umbral y la interpretación de su valor también deberían justificarse en relación con el contenido, por ejemplo, explicando qué incoherencias restantes fueron aceptadas por los investigadores.

Queremos concluir esta sección con dos notas importantes: *en primer lugar*, pensamos que el cálculo y la publicación de un coeficiente de corrección del azar no debería distraer a los investigadores cualitativos del proceso de mejora de su sistema de categorías. *En segundo lugar*, es necesario para el cálculo de los coeficientes de corrección del azar que los segmentos que deben codificarse, es decir, las unidades de codificación, se definan *a priori*. Si los codificadores son libres de establecer los límites del segmento, no tiene sentido calcular coeficientes de acuerdo con corrección del azar. La razón de esto es obvia: incluso para un texto de una página de 2.000 caracteres, la probabilidad de que dos codificadores seleccionen al azar exactamente los mismos caracteres y les asignen el mismo código tiende a cero. Por lo tanto, no es necesaria una corrección del azar.

Bibliografía

- Brennan, R. L. & Prediger, D. J. (1981). Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 41(3), 687-699. doi: 10.1177/001316448104100307
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46. doi: 10.1177/001316446002000104

- Feinstein, A. R. & Cicchetti, D. V. (1990). High agreement but low Kappa: I. the problems of two paradoxes. *Journal of Clinical Epidemiology*, 43(6), 543-549. doi: 10.1016/0895-4356(90)90158-L
- Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *The British Journal of Mathematical and Statistical Psychology*, 61(Pt 1), 29-48. doi: 10.1348/000711006X126600
- Krippendorff, K. (1970). Bivariate agreement coefficients for reliability of data. *Sociological Methodology*, 2, 139-150. doi: 10.2307/270787
- Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, 30(3), 411-433. doi: 10.1111/J.1468-2958.2004.TB00738.X
- Kuckartz, U. (2014). *Qualitative text analysis: A guide to methods, practice & using software*. Thousand Oaks, CA: SAGE. doi: 10.4135/9781446288719
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. doi: 10.2307/2529310
- Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*. Recuperado a partir de <https://nbn-resolving.de/urn:nbn:de:0168-ssoar-395173> [07.12.2019]
- Schreier, M. (2012). *Qualitative content analysis in practice*. Thousand Oaks, CA: SAGE.
- Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly*, 19(3), 321-325. doi: 10.1086/266577