MARSHALL

COLLEGE OF EDUCATION AND PROFESSIONAL DEVELOPMENT

COEPD OFFICE OF ASSESSMENT

WWW.MARSHALL.EDU/COEPD-ASSESSMENT

Content Validity & Reliability Procedure Guide

DR. TIMOTHY MELVIN

DIRECTOR OF ASSESSMENT

Introduction

The College of Education and Professional Development (COEPD) at Marshall University has established a content validity procedure for all Education Preparation Provider (EPP) created assessments and surveys, including key assessments, performance tasks, clinical evaluations, and national board-certified exams. The EPP adopted the procedure to evaluate its assessments in Spring 2022 and beyond. The content validity and reliability procedures are used by both initial- and advanced-level programs. Procedures follow the guidelines outlined in the CAEP Evaluation Framework document for EPP-Created Assessments to design, pilot, and judge the adequacy of the assessments created by the EPP.

The purpose of the content validity procedure is to provide guidance for collecting evidence and to document the adequate technical quality of assessment instruments and rubrics used to evaluate candidates in the COEPD.

CAEP Defined Assessments

CAEP uses the term “assessments” to cover content tests, observations, projects or assignments, and surveys – all of which are used with candidates. Surveys are often used to gather evidence on candidate preparation and candidate perceptions about their readiness to teach. Surveys are also helpful to measure the satisfaction of graduates or employers with preparation and the perceptions of clinical faculty about the preparedness of EPP completers.

Assessments and rubrics are used by faculty to evaluate candidates and provide them with feedback on their performance. Assessments and rubrics should address relevant and meaningful candidate knowledge, performance, and dispositions, aligned with CAEP standards. An EPP will use assessments that comprise evidence offered in accreditation self-study reports to examine candidates at various points from admission through completion consistently. These are assessments that all candidates are expected to complete as they pass from one stage of preparation to the next or that are used to monitor the progress of candidates’ developing proficiencies during one or more stages of preparation.

Validity Vs. Reliability

Validity

- ✓Process by which a test developer collects evidence to support the types of inferences that are to be drawn from test scores.
- ✓Purpose to assess whether the items adequately represent a performance domain or construct of specific interest.

Reliability

- ✓Confidence that a test score earned is a good representation of actual knowledge of the content.

COEPD 10-Step Procedure

Validity

- ✓Form Working Group and Identify Performance Domains & Rubric Items
- ✓Q-Sort Methodology
- ✓Lawshe's Content Validity Ratio (CVR)
- ✓Create or Modify Rubric
- ✓Expert Panel Distribution
- ✓Content Validity Index (CVI)

Reliability

- ✓Cronbach's Alpha
- ✓Pilot Rubric
- ✓Calibrate Rubric
- ✓Cohen's Kappa

VALIDITY

Step 1: Working Group, Performance Domains, & Rubric Items

✓**Form a Small Working Group**

- The working group should consist of faculty, students, program completers, and external representatives.
- Generally, 5-7 individuals are sufficient.
- Please be mindful that the external representative must be an expert in the field.
- Be sure to document the name, employment title, years of service in the field, and employer of all completers and external representatives

Step 1: Working Group, Performance Domains, & Rubric Items

✓Identify Performance Domains

- Identify Performance Domains with your Working Group.
- Your performance domain may be taken directly from and aligned to your specialized accreditation standards
- The performance domain should contain an operational definition.

•Example

- **COMMITMENT TO STUDENTS** (CAEP RA1.1): The creation of a learning environment and community to promote successful teaching and learning. Advanced candidates of the COEPD shall:

Step 1: Working Group, Performance Domains, & Rubric Items

✓Compile Rubric Items

- Items in your rubric should align to and measure the Performance Domain.
- Think carefully about the realm of possibilities when creating your items. You cannot have too many items

✓Example

COMMITMENT TO STUDENTS
All Possible Items (as defined by Working Group):

- ✓ Respects the rights of all stakeholders
- ✓ Promotes collaboration and teamwork to improve learning
- ✓ Selects, uses, adapts, and promotes evidence-based practices to meet the needs of learners
- ✓ Selects and uses valid assessment instruments to inform professional decisions
- ✓ Promotes policies and practices which facilitate a positive learning environment
- ✓ Demonstrates flexibility and adaptability to novel or unexpected situations
- ✓ Engages in tasks in a manner showing preparation and organization
- ✓ Promotes system level change to better meet the needs of student's and their families
- ✓ Advocates for student's needs

Step 2: Q-Sort Methodology

✓**Identify Panel of Experts**

- The working group should now identify a Panel of Experts consisting of as many external representatives (content experts) as possible.
- The Panel of Experts should consist of at least 15 members. Although faculty, students, and program completers will be a part of the Expert Panel, strive to have as many external content experts as possible.
- Be sure to document the name, employment title, and employer of all completers and external representatives.

Step 2: Q-Sort Methodology

✓Distribute Virtual Q-Sort

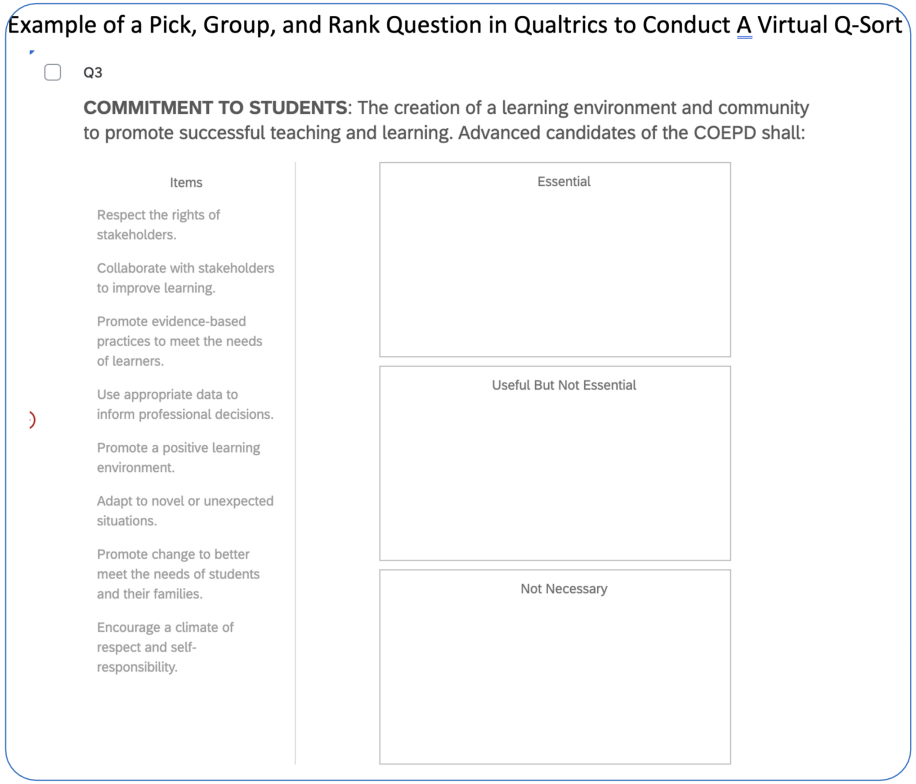
- Distribute a Virtual Q-Sort to the Panel of Experts. A Q-Sort is a sorting technique designed to study subjectivity (views, opinions, beliefs, values, etc.). For this process, you will conduct a Virtual Q-Sort using Qualtrics.
- Using Qualtrics, you will use the **Pick, Group, and Rank** question type in your survey. If you do not have a Qualtrics account, please request a faculty account at www.marshall.edu/qualtrics.
- Essentially, the working group takes a hard look at each possible item aligned to a domain, then ranks whether it is an **Essential** Item, an Item that is **Useful but Not Essential**, or an item that is **Not Necessary**.

Step 2: Q-Sort Methodology

✓Distribute Virtual Q-Sort

- Steps to a Virtual Q-Sort Survey Design:
- In a new survey, write clear instructions to the Panel of Experts with expectations.
  - Select a Pick, Group, and Rank Question Type
    - Include the Domain, Operational Definition, and ALL items identified by the Working Group in Step 1. The Panel of Experts will be able to drag and drop each item into the three scales: Essential, Useful but Not Essential, or Not Necessary.
  - Distribute Virtual Q-Sort to Panel of Experts

Example of a Pick, Group, and Rank Question in Qualtrics to Conduct A Virtual Q-Sort



Step 3: Lawshe's Content Validity Ratio

✓Q-Sort Results

- Once you receive your Q-Sort Results in Qualtrics, you can obtain Lawshe's Content Validity Ratio (CVR) for each item. You first look for the total number of individuals who ranked the item as Essential (***Example #1***).
- CVR calculates a proportionate level of agreement for each item.

QID1 - Groups

#	Field	Respect the rights of stakeholders.	Collaborate with stakeholders to improve learning.	Promote evidence-based practices to meet the needs of learners.	Use appropriate data to inform professional decisions.	Promote a positive learning environment.	Adapt to novel or unexpected situations.	Promote change to better meet the needs of students and their families.	Encourage a climate of respect and self-responsibility.	Total
1	Essential	11.05% 20	11.60% 21	16.57% 30	16.02% 29	16.02% 29	5.52% 10	9.94% 18	13.26% 24	181
2	Useful But Not Essential	17.19% 11	15.63% 10	3.13% 2	3.13% 2	4.69% 3	28.13% 18	17.19% 11	10.94% 7	64
3	Not Necessary	12.50% 1	25.00% 2	0.00% 0	0.00% 0	0.00% 0	25.00% 2	25.00% 2	12.50% 1	8

Showing rows 1 - 3 of 3

Step 3: Lawshe's Content Validity Ratio

✓**Find CVR**

- Download and use the CVR Calculator to obtain your CVR.
- **CVR = (ne-N/2)/(N/2)**
    1. ne = total number of respondents who rated item as Essential.
    2. N = total number of participants.

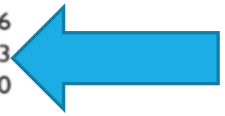
	A	B	C	D
1	Content Validity and Reliability (CVR) Calculator			
2	Item	# of respondents who rated essential (n_e)	Total # of participants (N)	CVR
3	1: Respects the rights of stakeholders	20	39	0.026
4	2: Collaborate with Stakeholders	21	39	0.077
5	3: Promote Evidence Based Practices	30	39	0.538
6	4: Use Appropriate Data	29	39	0.487
7	5: Promote Positive Learning Env.	29	39	0.487
8	6: Adapt to Situations	10	39	-0.487
9	7: Promote Change	18	39	-0.077
10	8: Encourage Respect	24	39	0.231
11	9			

Step 3: Lawshe's Content Validity Ratio

✓Retain Items Meeting CVR

- Use the CVR Chart (on the next slide) to identify the number of participants and the CVR critical value associated with the number of participants. Compare the CVR obtained in Excel with the CVR Chart to determine the minimum CVR value required for an item to be valued based on the number of participants.
- In the example (on the next slide), you will see that since there were 39 respondents, we used **.333** as the CVR critical value. Therefore, we only retain and use items that have a CVR at or above .333.
- **NOTE**: If you used items directly from your specialized standards, and the expert panel did not rank them as essential or were NOT retained with CVR – USE THE ITEM ANYWAY because the item is from your standards.

Panel Size	$N_{critical}$ (Minimum Number of Experts Required to Agree an Item Essential for Inclusion)	Proportion Agreeing Essential	$CVR_{critical}$
5	5	1	1.00
6	6	1	1.00
7	7	1	1.00
8	7	.875	.750
9	8	.889	.778
10	9	.900	.800
11	9	.818	.636
12	10	.833	.667
13	10	.769	.538
14	11	.786	.571
15	12	.800	.600
16	12	.750	.500
17	13	.765	.529
18	13	.722	.444
19	14	.737	.474
20	15	.750	.500
21	15	.714	.429
22	16	.727	.455
23	16	.696	.391
24	17	.708	.417
25	18	.720	.440
26	18	.692	.385
27	19	.704	.407
28	19	.679	.357
29	20	.690	.379
30	20	.667	.333
31	21	.677	.355
32	22	.688	.375
33	22	.667	.333
34	23	.676	.353
35	23	.657	.314
36	24	.667	.333
37	24	.649	.297
38	25	.658	.316
39	26	.667	.333
40	26	.650	.300



Step 4: Create or Modify Rubric

✓**Working Group Creates or Adjust Rubrics**

- The working group has now identified the items to keep in a rubric based on the CVR.
- The working group may need to create a new rubric, adjust an existing rubric, and possibly modify the accompanying assessment so that it focuses on the rubric.

Step 4: Create or Modify Rubric

✓**Rubrics**

- As many standards, if not most, align to candidate performance, most assessment measures will align to a performance-based rubric which will allow for:
  1. Common framework and language for assessment purposes.
  2. Performance or behavior examinations.
  3. Standard and criteria evaluations.
  4. Substantive faculty discussion on improvement.
  5. Collaboration promoting shared expectations and grading practices.

Step 4: Create or Modify Rubric

✓**Rubric Contents**

- Rubric should contain the following:
  1. Domain or Standard.
  2. Operational Definition or Task.
  3. Items for Rating (skills, content knowledge, dispositions, etc.) aligned to a standard and standard component (if applicable).
  4. Levels of Performance or Mastery.
  5. Description of each characteristic at each level of performance/mastery.

Step 4: Create or Modify Rubric

✓**Rubric Development**

- When creating or adjusting a rubric, discuss the following with your working group:
  1. Identify what you want to assess.
  2. Identify the items you obtained from your Q-Sort/CVR.
  3. Identify and Describe each Level of Performance/Mastery.
     1. Describe the best work you could expect using these levels (top/positive level)
     2. Describe an unacceptable product (lowest/negative level).
     3. Describe levels of intermediate or mediocre products (middle levels).

Step 5: Expert Panel Distribution

✓**Distribute Information to Panel of Experts**

- Distribute the Assignment, Assignment Instructions, and the Evaluation Rubric to the same Panel of Experts used with the Q-Sort.
- Looking for Construct Validity, the panel members will rate the representativeness and clarity of each item as it relates to the overall construct using the following scales: Item is Representative/Clear, Item Needs Minor Revisions, Item Needs Major Revisions, Item is Not Representative/Clear.
- Representativeness refers to how well the item measures the domain and Clarity refers to how clearly the item is stated.
  - **NOTE**: The panel of experts will not be rating the performance indicators or descriptors. Only the items.
- Provide each expert with either a paper copy response form or a form created in Qualtrics. Example #1 demonstrates using Qualtrics to collect responses.
  - A template for paper copies can be found the COEPD Resources Microsoft Team under Assessment.

Step 6: Content Validity Index

✓**Content Validity Index (CVI)**

- With data obtained from the Panel of Experts on Item Representativeness and Clarity, complete the CVI for each item:

$$\text{CVI} = \frac{\text{Number of Experts Who Rated the Item a 3 or 4}}{\text{Number of Total Experts}}$$

- Retain only items with CVI ≥ .80.
  - If under, convene the working group to determine how the item fits with the domain or standard or how it is worded.

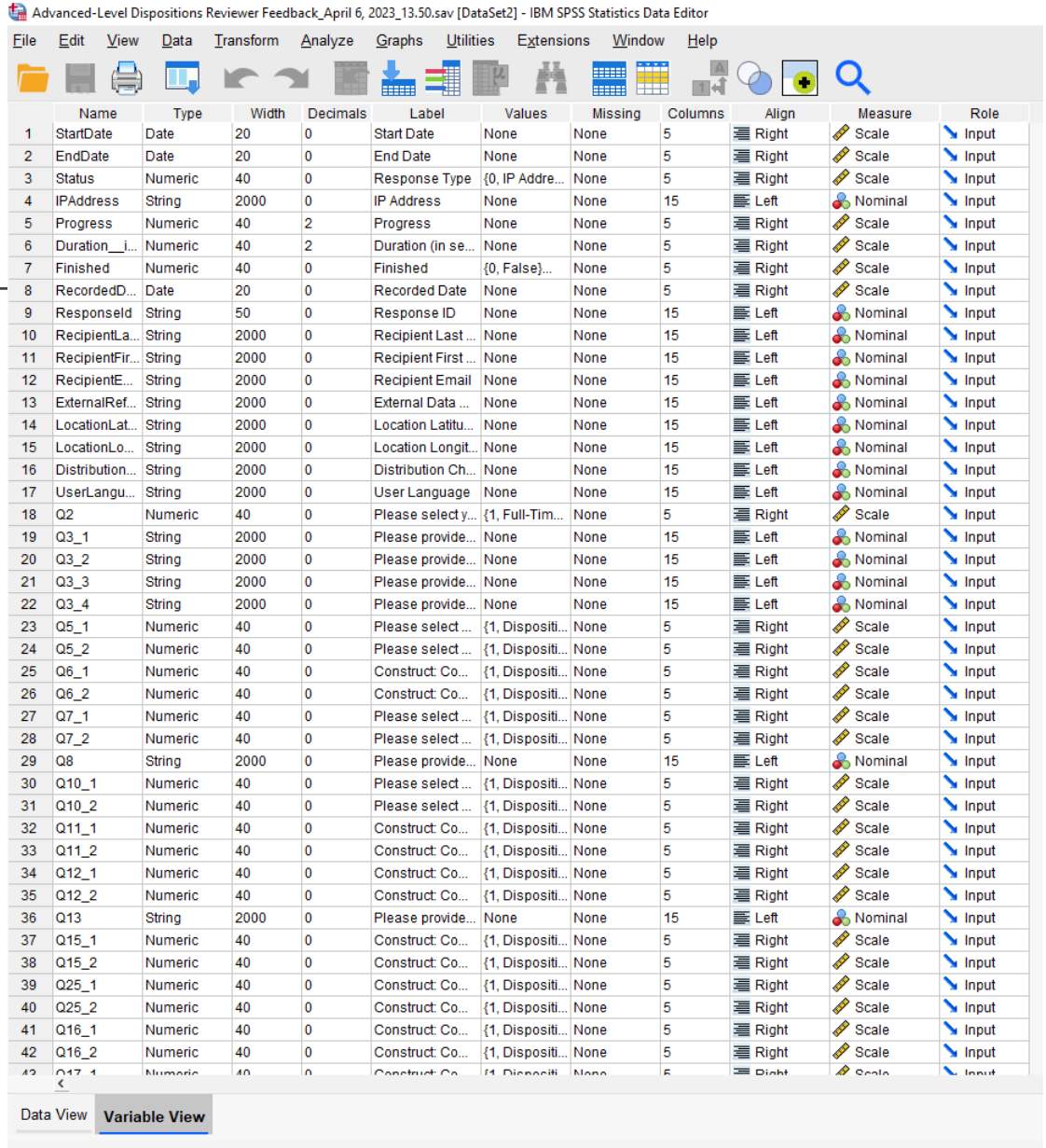
RELIABILITY

Step 7: Chronbach's Alpha

- ✓ **Chronbach's Alpha**
  - ✓ Measure used to assess the reliability, or internal consistency, of a set of scale or test items and indicates whether an item measures the same construct.
  - ✓ Chronbach's Alpha is the most common measure of reliability.
  - ✓ Two alpha's will be obtained:
    - ✓ One: Representativeness
    - ✓ Two: Clarity
- ✓ Chronbach's Alpha closer to 1 is considered highly reliable. The closer to zero is less reliable.
- ✓ General rule that .80 or above is high reliability.

Step 7: Chronbach's Alpha

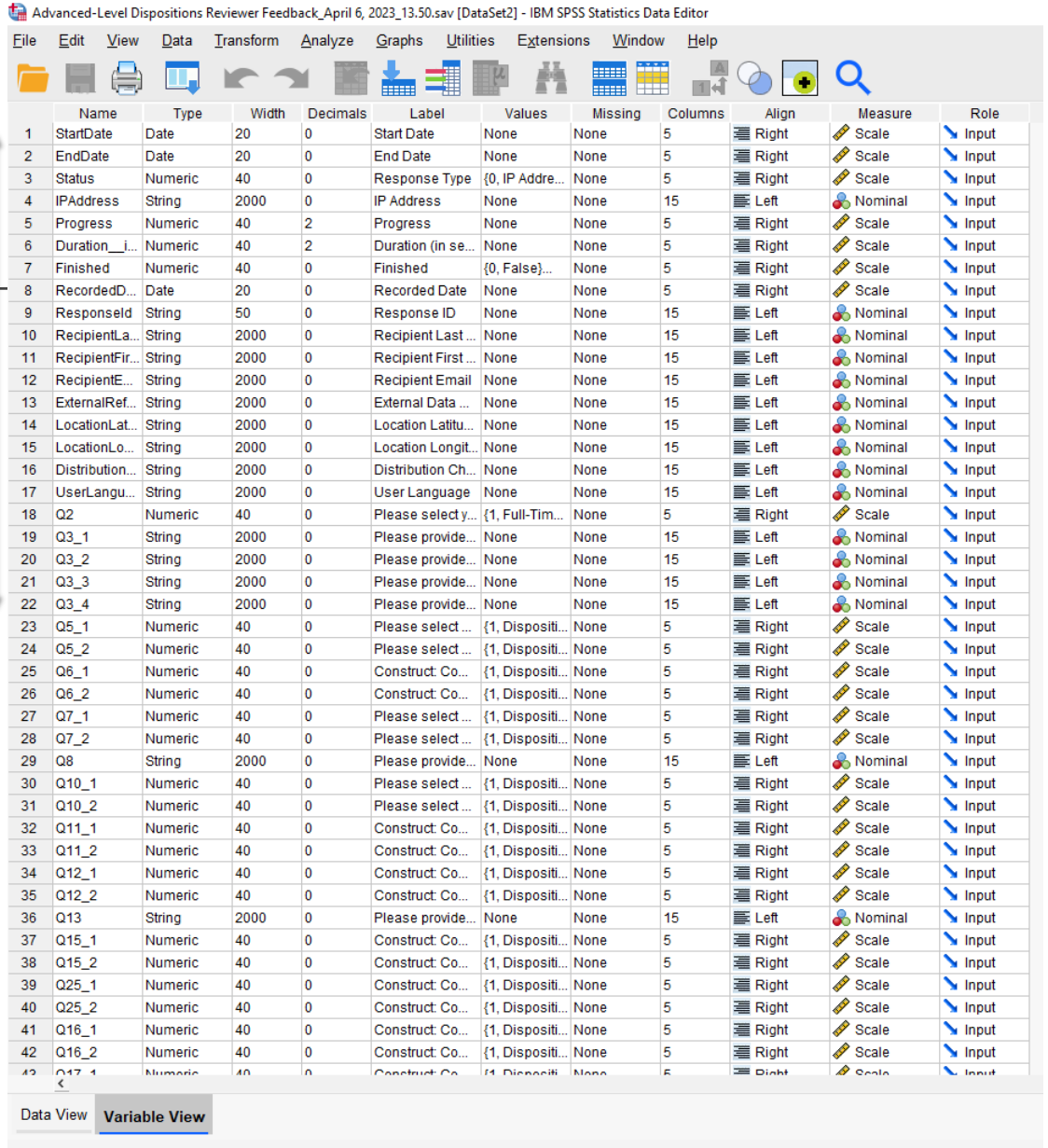
- ✓**Raw Data Set Exported to SPSS from Qualtrics**



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	StartDate	Date	20	0	Start Date	None	None	5	Right	Scale	Input
2	EndDate	Date	20	0	End Date	None	None	5	Right	Scale	Input
3	Status	Numeric	40	0	Response Type	{0, IP Addre...	None	5	Right	Scale	Input
4	IPAddress	String	2000	0	IP Address	None	None	15	Left	Nominal	Input
5	Progress	Numeric	40	2	Progress	None	None	5	Right	Scale	Input
6	Duration__i...	Numeric	40	2	Duration (in se...	None	None	5	Right	Scale	Input
7	Finished	Numeric	40	0	Finished	{0, False}...	None	5	Right	Scale	Input
8	RecordedD...	Date	20	0	Recorded Date	None	None	5	Right	Scale	Input
9	ResponseId	String	50	0	Response ID	None	None	15	Left	Nominal	Input
10	RecipientLa...	String	2000	0	Recipient Last ...	None	None	15	Left	Nominal	Input
11	RecipientFir...	String	2000	0	Recipient First ...	None	None	15	Left	Nominal	Input
12	RecipientE...	String	2000	0	Recipient Email	None	None	15	Left	Nominal	Input
13	ExternalRef...	String	2000	0	External Data ...	None	None	15	Left	Nominal	Input
14	LocationLat...	String	2000	0	Location Latitu...	None	None	15	Left	Nominal	Input
15	LocationLo...	String	2000	0	Location Longit...	None	None	15	Left	Nominal	Input
16	Distribution...	String	2000	0	Distribution Ch...	None	None	15	Left	Nominal	Input
17	UserLangu...	String	2000	0	User Language	None	None	15	Left	Nominal	Input
18	Q2	Numeric	40	0	Please select y...	{1, Full-Tim...	None	5	Right	Scale	Input
19	Q3_1	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
20	Q3_2	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
21	Q3_3	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
22	Q3_4	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
23	Q5_1	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
24	Q5_2	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
25	Q6_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
26	Q6_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
27	Q7_1	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
28	Q7_2	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
29	Q8	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
30	Q10_1	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
31	Q10_2	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
32	Q11_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
33	Q11_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
34	Q12_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
35	Q12_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
36	Q13	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
37	Q15_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
38	Q15_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
39	Q25_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
40	Q25_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
41	Q16_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
42	Q16_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input

Step 7: Chronbach's Alpha

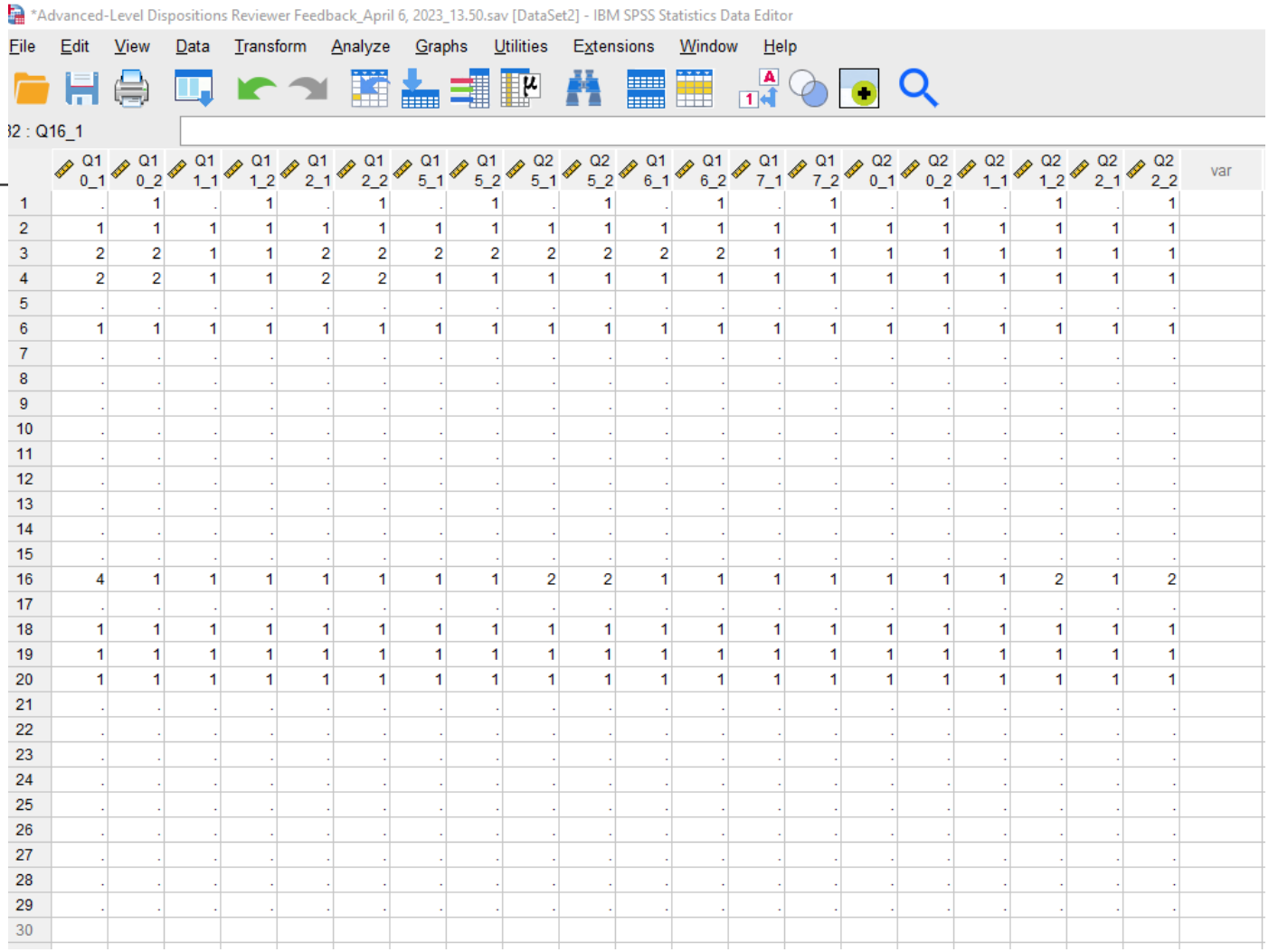
- ✓Delete lines of Data exported from Qualtrics that may be deleted. You only need specific question responses. Notice this is in Variable view.
  - ✓ In this example, lines 1-22 will be deleted.
  - ✓ Lines 1-17 is Qualtrics collected data.
  - ✓ Line 18 was asking participants if they were faculty, staff, etc.
  - ✓ Lines 19-22 were asking for name and employer information.
- ✓To delete, right-click on the line and select "clear"
- ✓Note: Also deleted were lines 7, 14, 23, and 30 because those were question boxes with no responses.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	StartDate	Date	20	0	Start Date	None	None	5	Right	Scale	Input
2	EndDate	Date	20	0	End Date	None	None	5	Right	Scale	Input
3	Status	Numeric	40	0	Response Type	{0, IP Addre...	None	5	Right	Scale	Input
4	IPAddress	String	2000	0	IP Address	None	None	15	Left	Nominal	Input
5	Progress	Numeric	40	2	Progress	None	None	5	Right	Scale	Input
6	Duration__i...	Numeric	40	2	Duration (in se...	None	None	5	Right	Scale	Input
7	Finished	Numeric	40	0	Finished	{0, False}...	None	5	Right	Scale	Input
8	RecordedD...	Date	20	0	Recorded Date	None	None	5	Right	Scale	Input
9	ResponseId	String	50	0	Response ID	None	None	15	Left	Nominal	Input
10	RecipientLa...	String	2000	0	Recipient Last ...	None	None	15	Left	Nominal	Input
11	RecipientFir...	String	2000	0	Recipient First ...	None	None	15	Left	Nominal	Input
12	RecipientE...	String	2000	0	Recipient Email	None	None	15	Left	Nominal	Input
13	ExternalRef...	String	2000	0	External Data ...	None	None	15	Left	Nominal	Input
14	LocationLat...	String	2000	0	Location Latitu...	None	None	15	Left	Nominal	Input
15	LocationLo...	String	2000	0	Location Longit...	None	None	15	Left	Nominal	Input
16	Distribution...	String	2000	0	Distribution Ch...	None	None	15	Left	Nominal	Input
17	UserLangu...	String	2000	0	User Language	None	None	15	Left	Nominal	Input
18	Q2	Numeric	40	0	Please select y...	{1, Full-Tim...	None	5	Right	Scale	Input
19	Q3_1	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
20	Q3_2	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
21	Q3_3	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
22	Q3_4	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
23	Q5_1	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
24	Q5_2	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
25	Q6_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
26	Q6_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
27	Q7_1	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
28	Q7_2	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
29	Q8	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
30	Q10_1	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
31	Q10_2	Numeric	40	0	Please select ...	{1, Dispositi...	None	5	Right	Scale	Input
32	Q11_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
33	Q11_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
34	Q12_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
35	Q12_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
36	Q13	String	2000	0	Please provide...	None	None	15	Left	Nominal	Input
37	Q15_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
38	Q15_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
39	Q25_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
40	Q25_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
41	Q16_1	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input
42	Q16_2	Numeric	40	0	Construct: Co...	{1, Dispositi...	None	5	Right	Scale	Input

Step 7: Chronbach's Alpha

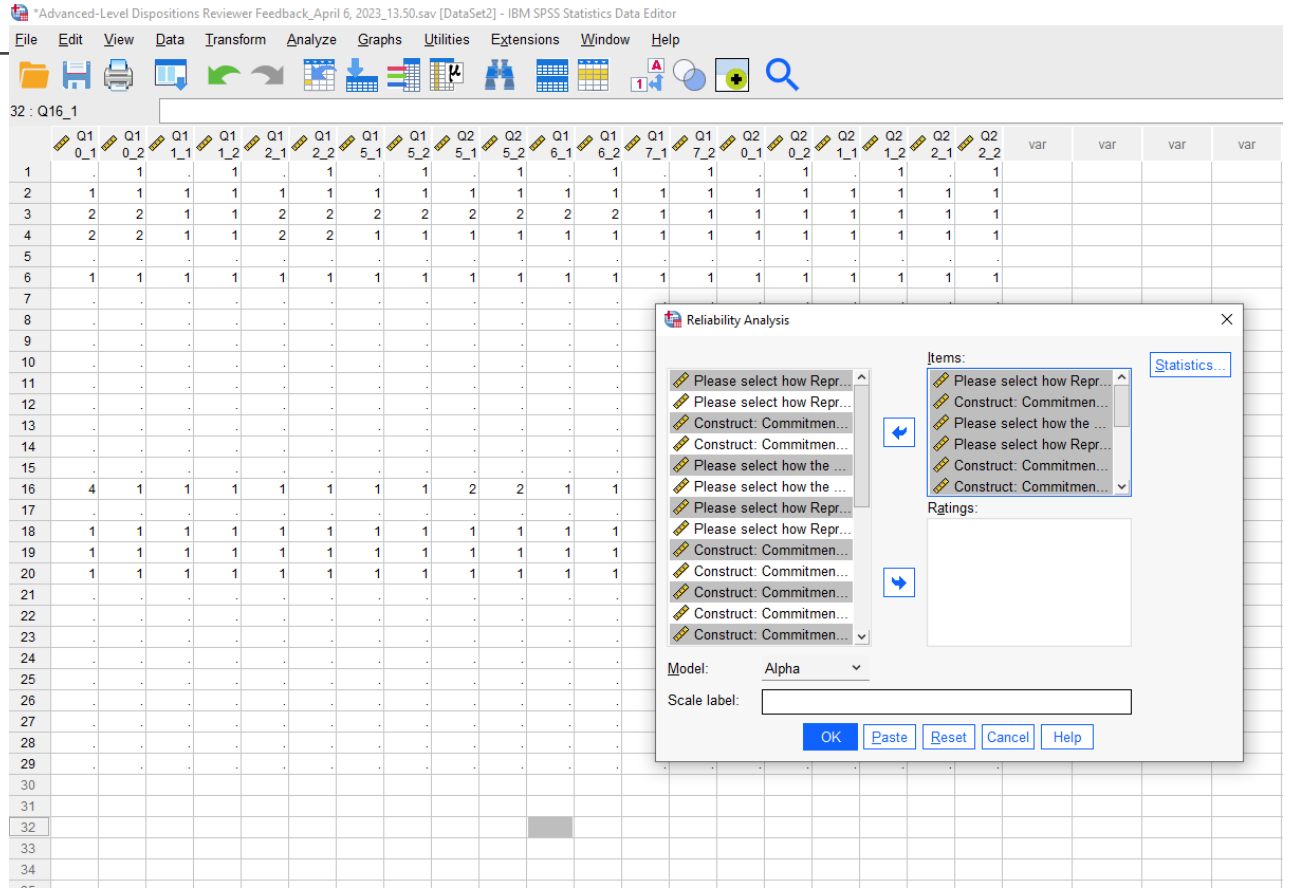
- ✓You'll be left with a clean data sheet with responses to only:
  - ✓Representativeness
  - ✓Clarity
- ✓Things to notice:
  - ✓29 responses, but only 8 lines of data
  - ✓Participants may have clicked on survey but did not participate.



	Q10_1	Q10_2	Q11_1	Q11_2	Q12_1	Q12_2	Q15_1	Q15_2	Q25_1	Q25_2	Q16_1	Q16_2	Q17_1	Q17_2	Q20_1	Q20_2	Q21_1	Q21_2	Q22_1	Q22_2	var
1	.	1	.	1	.	1	.	1	.	1	.	1	.	1	.	1	.	1	.	1	
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
3	2	2	1	1	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	
4	2	2	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
9	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
16	4	1	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	2	1	2	
17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
21	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
26	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
29	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
30																					

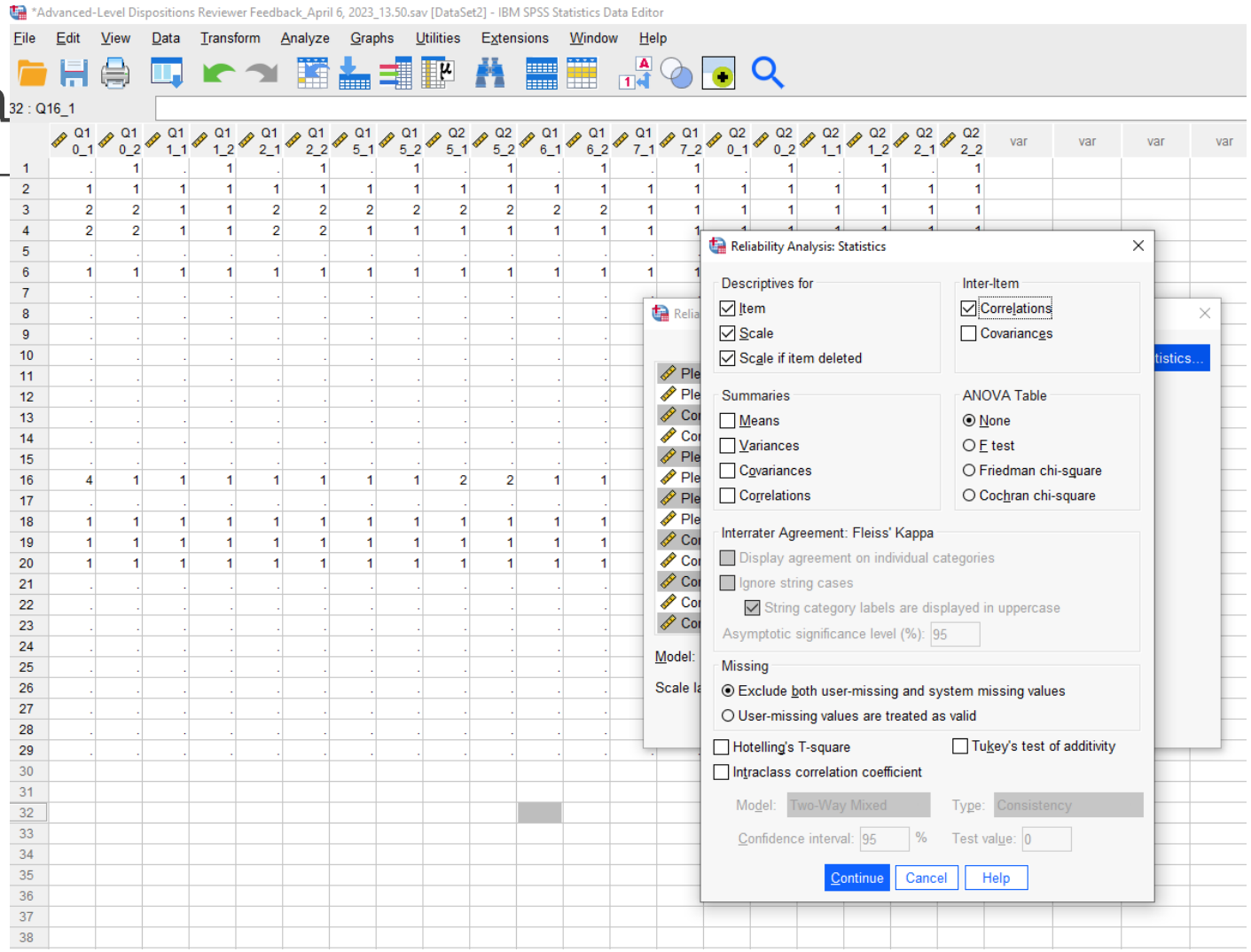
Step 7: Chronbach's Alpha

- ✓Select Analyze > Scale > Reliability Analysis
- ✓Since Chronbach's Alpha is looking for internal consistency, we want to separate our two measures:
  - ✓ (1) Representativeness
  - ✓ (2) Clarity
- ✓I know that the first question item, and every other item, relates to Representativeness
  - ✓ Beginning with the first item, select every other item and move them into the Items box.
  - ✓ Select Statistics



Step 7: Chronbach's Alpha

- ✓Select the following:
  - ✓Item
  - ✓Scale
  - ✓Scale if item deleted
  - ✓Correlations
- ✓Select Continue
- ✓Select Ok



Step 7: Chronbach's Alpha

- ✓Representativeness:
  - ✓Chronbach's Alpha = .845
  - ✓.80 or above is high reliability.

Scale: ALL VARIABLES

Case Processing Summary

		N	%
Cases	Valid	8	27.6
	Excluded[a]	21	72.4
	Total	29	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.845	.929	8

- ✓Repeated Steps for Clarity:
  - ✓Chronbach's Alpha = .903
  - ✓.80 or above is high reliability.

Scale: ALL VARIABLES

Case Processing Summary

		N	%
Cases	Valid	9	31.0
	Excluded[a]	20	69.0
	Total	29	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.903	.905	10

Step 8: Pilot Your Rubric

- ✓Pilot Your Rubric!
  - ✓Make Changes as needed depending on any prior information CVR, CVI, or Chronbach's Alpha
  - ✓If no changes are needed, distribute the assessment and use the rubric to score submissions in the next available section of your course.

Step 9: Calibrate, or "Norm" the Rubric

- ✓Calibrating the Rubric simply means that the rubric items and scales are clear enough to lead to a consensus among faculty who are grading student work.
  - ✓Important when various faculty or adjuncts are evaluating assessments among multiple course sections.
- ✓Calibrating helps to ensure that the rubric use is consistent, and for the most part, free from bias.
- ✓Calibrating helps to ensure that assessment data reflects the expectations of the program, not an individual faculty or instructor.

Step 9: Calibrate, or "Norm" the Rubric

- ✓ Steps to Calibrate your Rubric:
  - ✓ Prepare Scorer Materials
    - ✓ Assessment Instructions
    - ✓ Evaluation Rubric
    - ✓ Student Artifact
  - ✓ Assignment Ratings
    - ✓ After reading assessment instructions, rubric, and artifact, scorers will circle or indicate the candidate performance on each rubric item that they feel defines the quality of work.
  - ✓ Share Scores
    - ✓ One at a time, scorers indicate how they ranked the artifact on each item
    - ✓ DO NOT share why scorers scored the artifact the way they did.
  - ✓ Justify Responses
    - ✓ Scorers justify their evaluation by pointing to specific language in the rubric and evidence in the student artifact. Discuss each piece of student work and resolve issues that may be present because of rubric language, or the evidence provided in the student artifact. Scorer consensus should be reached.

Step 9: Calibrate, or "Norm" the Rubric

✓Scenario:

	Distinguished (4)	Proficient (3)	Basic (2)	Unsatisfactory (1)	
Awareness of the impact of motivation and schema on the learning process	Lesson begins with a clear description of method to activate prior knowledge and build schema toward the learning objective(s); Includes questions that generate intellectual curiosity and engage students with the lesson content through **higher order critical thinking and problem solving**	Lesson begins with a clear description of method to activate prior knowledge and build schema toward the learning objective(s); **Includes questions that generate intellectual curiosity and engage students with the lesson content**	Lesson begins with a **clear description of method to activate prior knowledge and build schema toward the learning objective(s); Includes questions,** but questions lack student engagement and/or stimulation of intellectual curiosity	Lesson begins with **a method to introduce the lesson content**, but introduction does not activate prior knowledge and/or build schema toward the learning objective(s); Questions are not included or do not engage students in intellectual curiosity	

SCENARIO: 5 Faculty members are convened to norm a rubric. A component of the rubric is "Awareness of the Impact of Motivation and Schema on the Learning Process."

The faculty members are given the exact student artifact and asked to score the artifact using this rubric.

The Facilitator has asked that each evaluator provide their score for the artifact and Has recorded the following scores…

Step 9: Calibrate, or "Norm" the Rubric

✓Scenario:

A	B	C	D	E	F
	R1	R2	R3	R4	R5
Awareness of the impact of motivation and schema on the learning process	4	4	4	2	4

How should the facilitator proceed?

Step 9: Calibrate, or "Norm" the Rubric

✓Scenario:

A	B	C	D	E	F
	R1	R2	R3	R4	R5
Awareness of the impact of motivation and schema on the learning process	4	4	4	2	4

How should the facilitator proceed?

The facilitator should begin by asking Rater 4 (R4) to their reasoning for giving the rating of "2." The rating could be for numerous reasons (unclear verbiage, misinterpretation, etc. The facilitator should ask for reasons behind the rating of "4."

From that point, R4 may be willing to increase their score.

Step 10: Cohen's Kappa

- ✓Cohen's Kappa assess the extent to which two raters agree
- ✓Ranges from -1 to +1
- ✓Cohen suggested that the Kappa results be interpreted as:
  - ✓values ≤ 0 as indicating no agreement
  - ✓0.01–0.20 as none to slight
  - ✓0.21–0.40 as fair
  - ✓0.41– 0.60 as moderate
  - ✓0.61–0.80 as substantial
  - ✓and 0.81–1.00 as almost perfect agreement

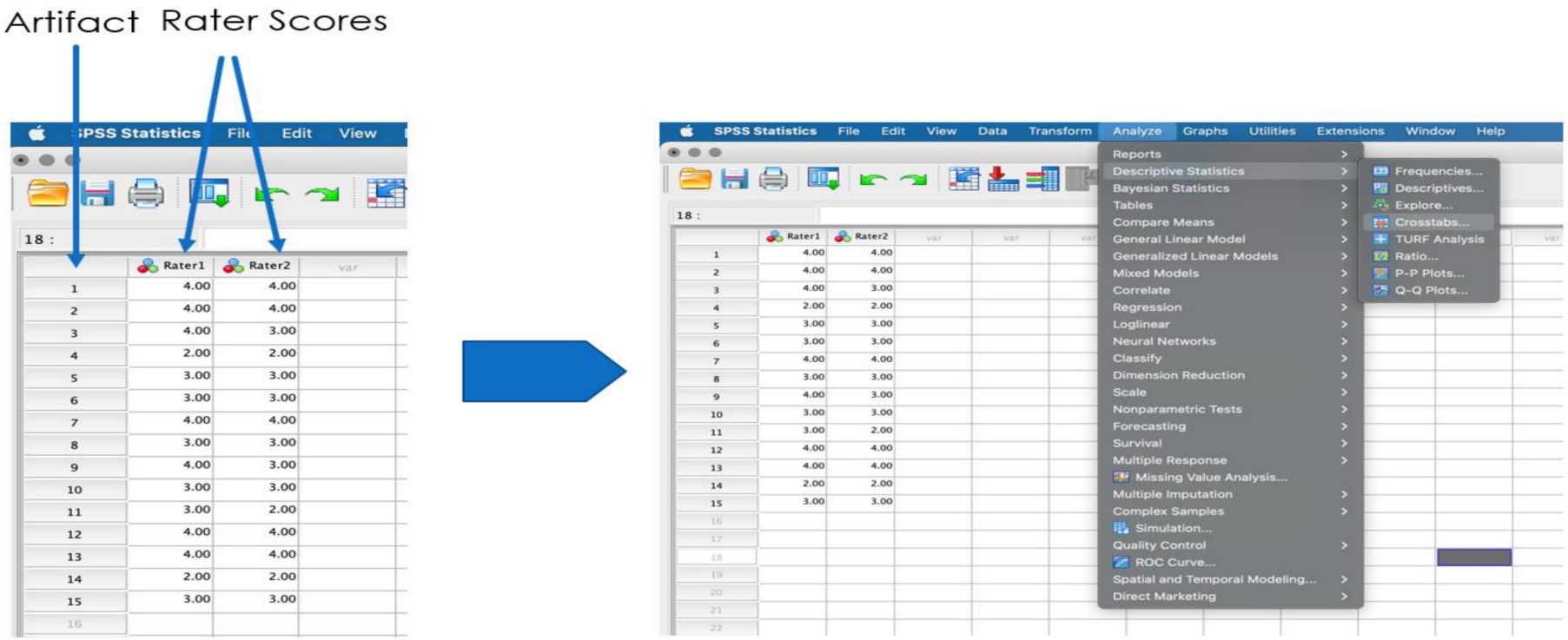
Step 10: Cohen's Kappa

- ✓Artifacts should be randomly assigned to two raters:
  - ✓Artifact One: Kim and Paula
  - ✓Artifact Two: Sandra and Lisa
  - ✓Artifact Three: Tom and Feon
  - ✓Artifact Four: Feon and Sandra
  - ✓Artifact Five: Lisa and Tom
  - ✓Artifact Six: Paula and Lanai
    - ✓ ….and so on.

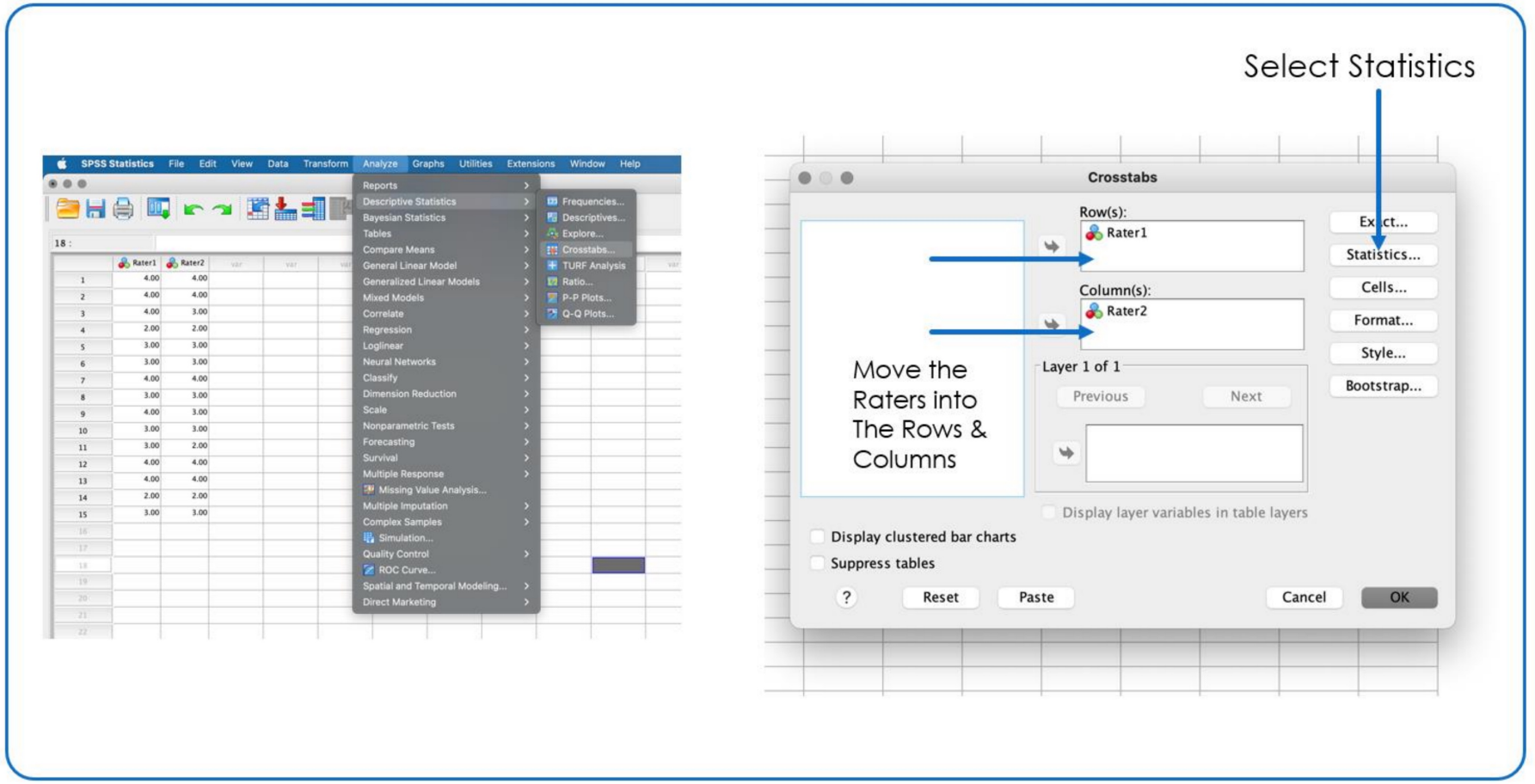
✓**Maintain a Master List** of who the raters are for each artifact but DO NOT share that list with the raters.

Step 10: Cohen's Kappa

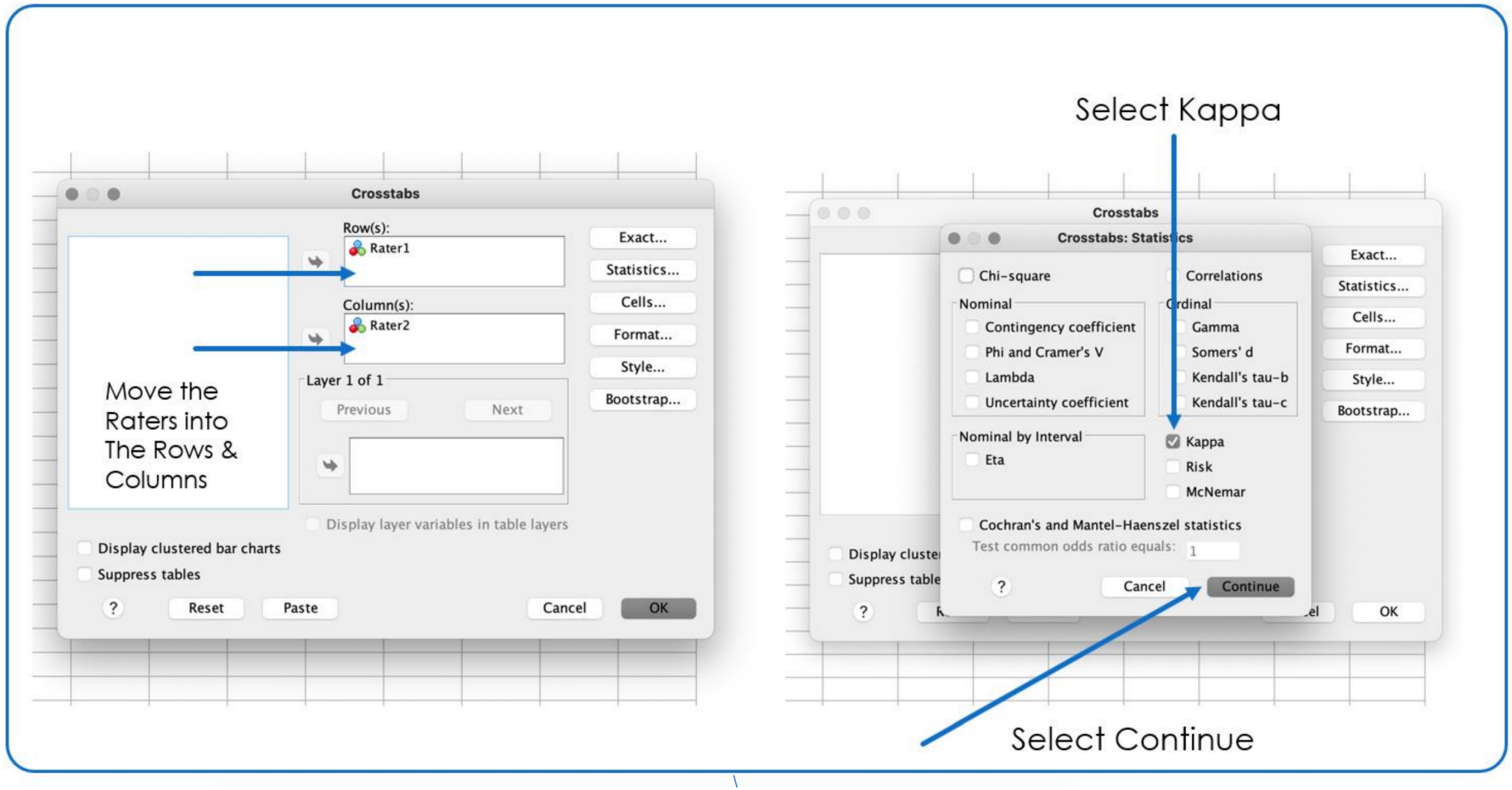
✓Using SPSS to find Cohen's Kappa



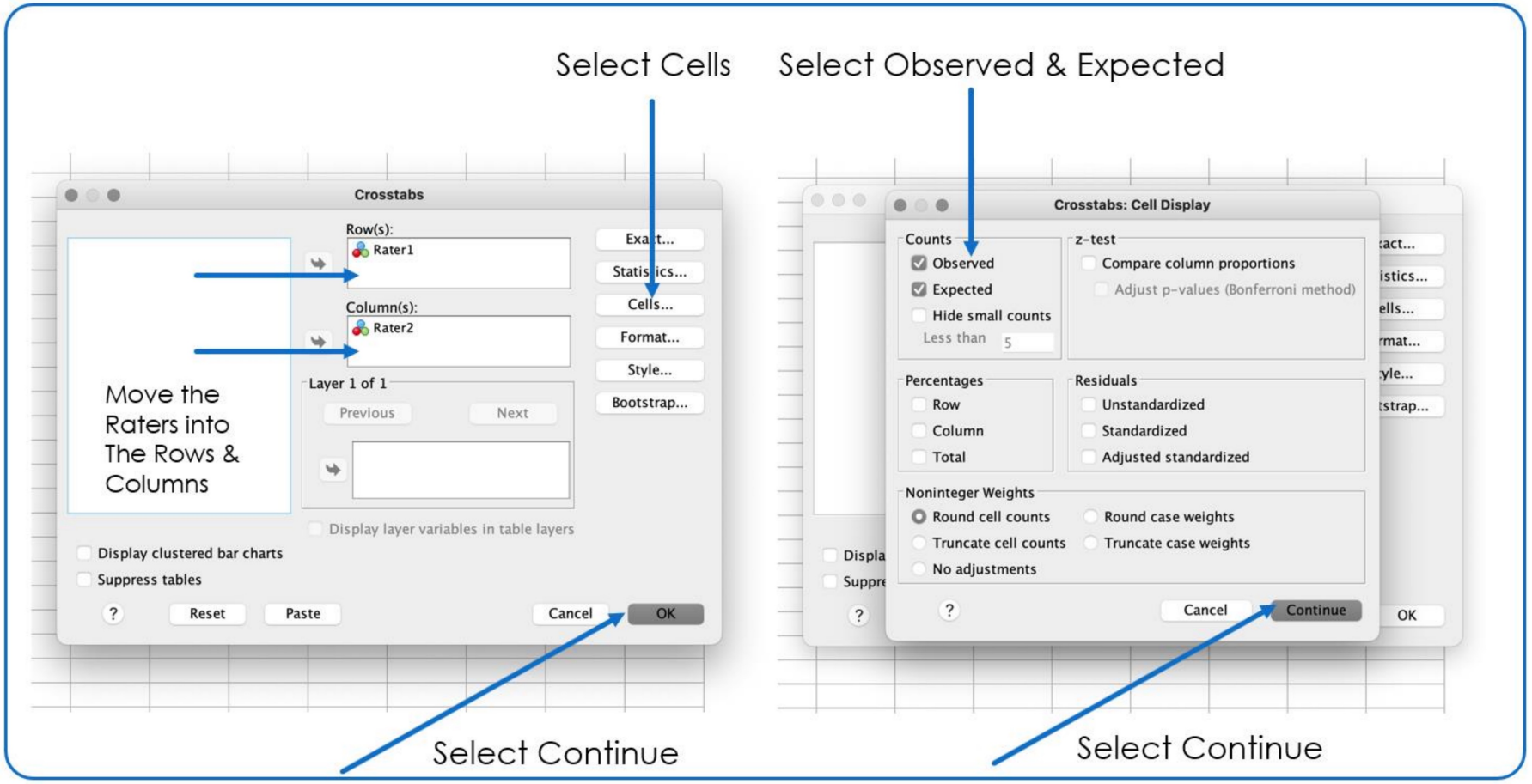
Step 10: Cohen's Kappa



Step 10: Cohen's Kappa



Step 10: Cohen's Kappa



Step 10: Cohen's Kappa

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Rater1 * Rater2	15	100.0%	0	0.0%	15	100.0%

Rater1 * Rater2 Crosstabulation

			Rater2			
			Basic	Proficient	Distinguished	Total
Rater1	Basic	Count	2	0	0	2
		Expected Count	.4	.9	.7	2.0
	Proficient	Count	1	5	0	6
		Expected Count	1.2	2.8	2.0	6.0
	Distinguished	Count	0	2	5	7
		Expected Count	1.4	3.3	2.3	7.0
Total		Count	3	7	5	15
		Expected Count	3.0	7.0	5.0	15.0

Cohen's Kappa = .683
Substantial Agreement

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.683	.162	3.645	.000
N of Valid Cases		15			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Step 10: Cohen's Kappa

- ✓You've now established that your EPP-Created Assessments are Valid and Reliable!
- ✓Make you sure that you document the steps taken for each semester. This documentation will be used as CAEP Evidence.