Coherence relations and DRD identification: theory and analysis

Ted Sanders (Utrecht University)
Wilbert Spooren (Radboud University Nijmegen)

Radboud University

Crash course statistical testing of reliability: Cohen's kappa

Practice in using corpora

- In discourse analytical research we often make use of corpora, to test our hypotheses
- For this we need to analyze corpora
- Frequently a researcher works in isolation
- Often in two or more experts annotate phenomena in a corpus and discuss over unresolved annotation issues

- Question: how solid are these data?

Practice in using corpora

- In empirical research we expect our data (=annotations) to be robust
- One of the criteria for robustness is reproducibility
  - if another researcher replicates the research, (s)he should be able to produce more or less the same data
  - if the same researcher replicates his/her research at a later time, it should produce more or less the same data
- This robustness should be tested

Measures for intercoder agreement

- Statements of the type “the corpus was annotated and coding problems were discussed”
- Percentages observed agreement between two or more coders
  - problem: no indication of chance agreement

Ann1\Ann2	Cat 1	Cat 2
Cat 1	30	5
Cat 2	6	40

Perc. obs. agr. =
(70/81)*100 = 86.4 %

Perc. ch. agr. = ((36/81)* (35/81))
+ (45/81)*(46/81) *100 = 50.7 %

Measures for intercoder agreement

- Chance-corrected measures
  - Measures for two coders (unweighted)
  - Measures for more than two coders (unweighted)
  - Measures for two or more coders (weighted)
- Unweighted: all disagreements weigh equally heavy
  Weighted: some disagreements weigh heavier than others

Two coders, unweighted

- All formulas for coefficient are the same

  (observed agreement – expected agreement) /
  (1 – expected agreement)

- Terminology
  - S
  - (Cohen's) κ
  - (Scott's) π
- they differ in the way expected agreement is calculated
  - (least sophisticated: S; most sophisticated: π)
- Here only κ

Two coders, unweighted

- Example

	Cat 1	Cat 2
Cat 1	30	5
Cat 2	6	40

Ao = (30+40)/81= .86

Ae = ((36/81)* (35/81)) + (45/81)*(46/81) = .51

κ = (Ao-Ae)/(1-Ae) = (.86-.51)/(1-.51) = .71

Interpretation:

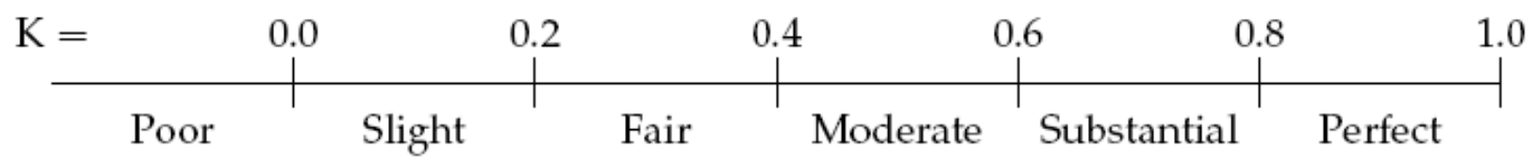
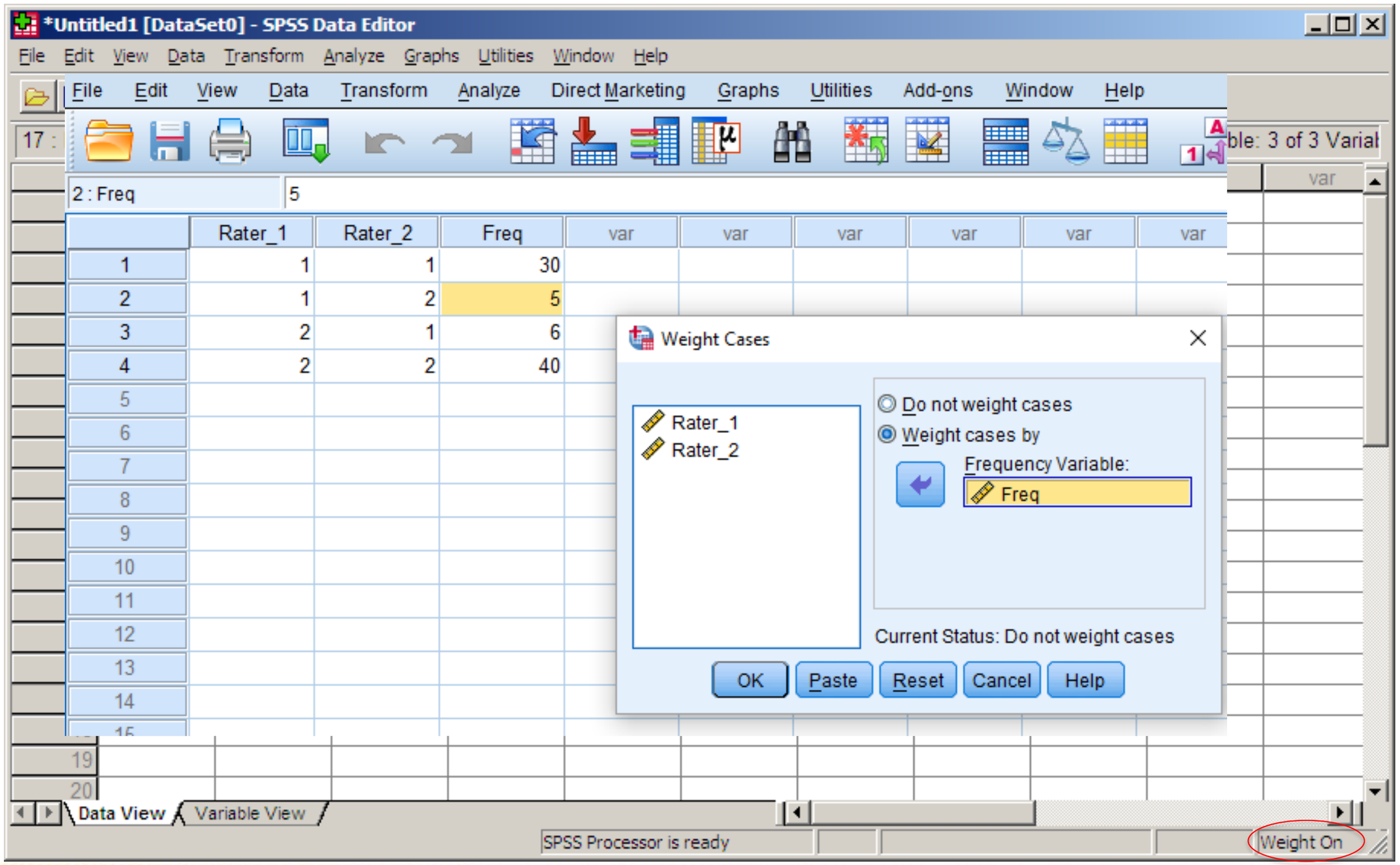


Figure 1
Kappa values and strength of agreement according to Landis and Koch (1977).

Two coders, unweighted, in SPSS



Two coders, unweighted, in SPSS

Rater_1 * Rater_2 Crosstabulation

Count

		Rater_2		
		1	2	Total
Rater_1	1	30	5	35
	2	6	40	46
Total		36	45	81

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	,724	,077	6,520	,000
N of Valid Cases		81			

Continue Cancel Help

NB SPSS calculates κ

Calculating agreement in practice

- ReCal
- Assignment: calculate agreement scores for our data
  - see the instructions page