

Gower program help file v1.0a



**Paul Barrett
Advanced Projects R&D Ltd.
Auckland
New Zealand**

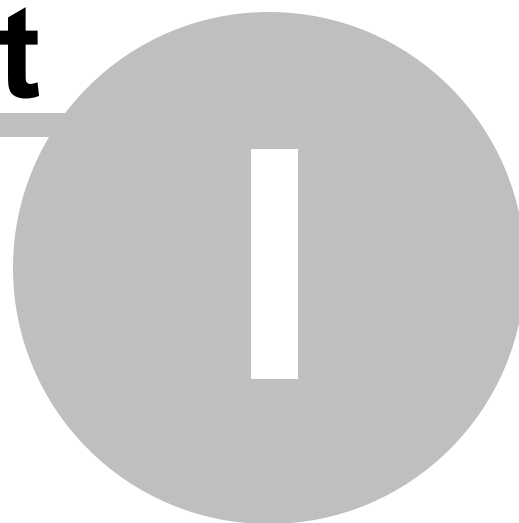
email: paul@pbarrett.net
Web: www.pbarrett.net

Table of Contents

Part I Introduction	3
Part II Datafile format	5
Part III Running the program	7
Part IV The Results Grid	17
Part V The data plots	23
Part VI The Coefficients	26
1 Pearson monotonicity.....	29
2 ICC-1 (intraclass).....	30
3 ICC-2 (intraclass).....	31
4 ICC-3 (intraclass).....	32
5 DSE-s (double-scaled euclidean).....	32
6 KSD-s (kernel-smooth distance).....	34
7 Gower agreement.....	36
Part VII Agreement and Monotonicity	39
Part VIII Common-Metric Rescaling	41
Part IX Relative v Simple Discrepancy	50

Advanced Projects R&D Ltd.

Part



1 Introduction

This program originally started life as a test routine which I was using to explore the [Gower](#) index alongside some other coefficients. But, in order to make it "usable" by others who might wish to look at their own data using the [Gower](#) index, I added a more user-friendly and informative interface - and some useful options.

The program only works with two columns of numbers - providing 7 coefficients computed from such 'bivariate' data. If you need a matrix of such values, for each of these coefficients, you would need to use the IRR-1 program (P.B. 22nd August, 2012: still on my 'to do' list I'm afraid).

* Version 1.0a takes data from an Excel .xls or .xlsx file as input, displaying the data and calculation results in two Excel-style grids, which can be printed or copied through the clipboard into Excel.

It also displays a scatterplot of the data as well as a [case-order plot](#). The latter plot is for the kinds of data which are ordered sequentially; for example where two variables or attributes are sampled over sequential time-points.

If you want a significance test for the Gower, use the Bootstrap program which is available for download from my [website](#). This computes the empirical sampling distribution for a [Gower](#) coefficient, taking into account the maximum possible discrepancy range, the kind of data (integer or reals) and the number of observations over which a [Gower](#) has been computed. It also allows you to test the significance of a difference between two Gowers, by constructing the empirical sampling distribution of the difference between two Gower coefficients, given their respective data characteristics (sample size, number type, and maximum possible discrepancy of observations in each sample).

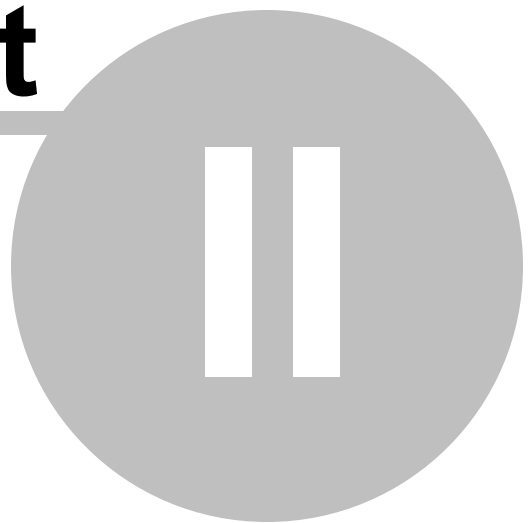
It's an interesting coefficient - rather useful when combined with a measure of monotonicity, the [Pearson](#) Correlation.

The best way to familiarize yourself with the program is to look at accompanying video help available at the [webpage](#) from which you downloaded this software.

* This was released on 14th August, 2012 .. incorporating a new Excel component for reading Excel .xls and .xlsx files.

Paul Barrett, (paul@pbarrett.net)

Part



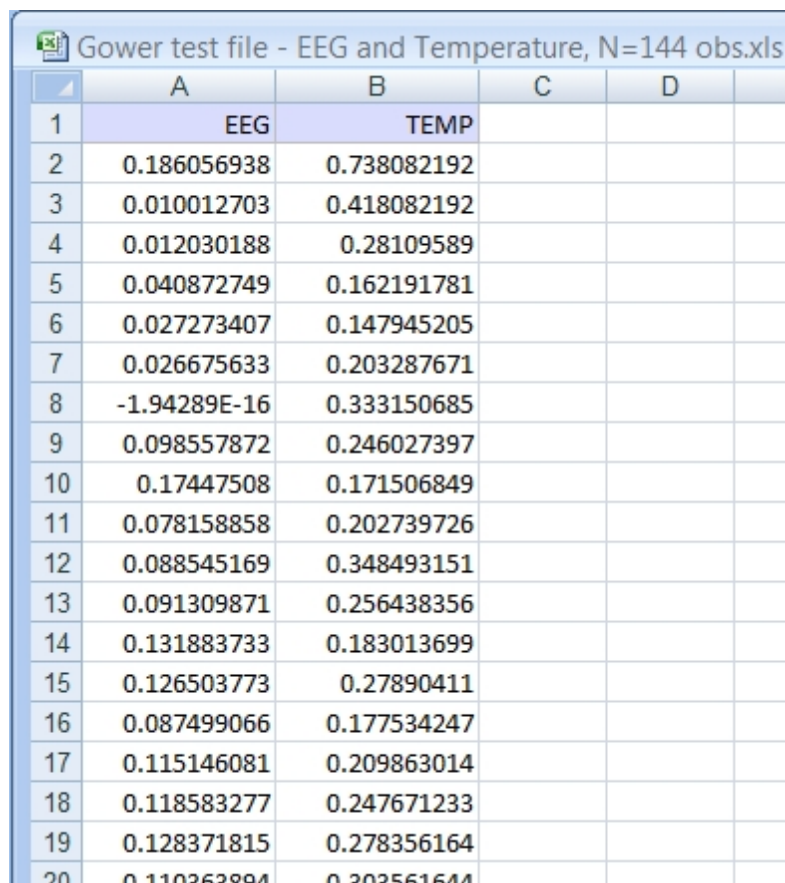
2 Datafile format

Data is input into the program via Excel files only; just two columns of data in a file, with the first line containing the variable/vector names. The program accepts Excel .xls or .xlsx files.

Missing data are expected to be blank cells (either with nothing at all in a cell or just one space character). If you use -9999 or something like that, the values will be read as observations. The missing-data count is provided in the results output, as a double-check for you.

The example file which comes with the program (*Gower test file - EEG and Temperature, N=144 obs.xls*) is a fragment of [common-metric](#) EEG and Temperature data, acquired sequentially over 144 time-points from an animal within a physiology experiment. The EEG data were actually measured in microvolts, with temperature in degree Celsius, but for the sake of graphical display, and for the correct calculation of relative discrepancy, the data were expressed within a common 0-1 metric.

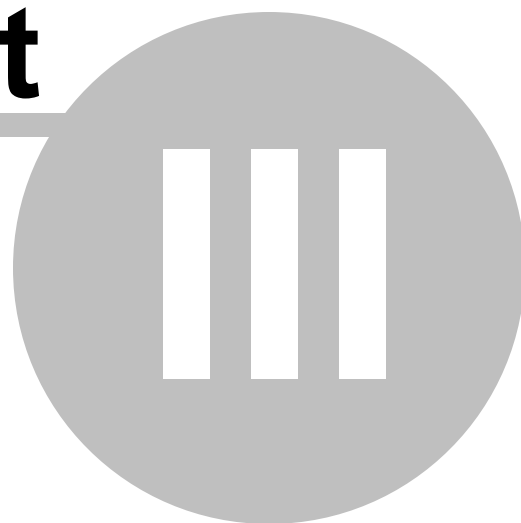
The file looks like:



	A	B	C	D
1	EEG	TEMP		
2	0.186056938	0.738082192		
3	0.010012703	0.418082192		
4	0.012030188	0.28109589		
5	0.040872749	0.162191781		
6	0.027273407	0.147945205		
7	0.026675633	0.203287671		
8	-1.94289E-16	0.333150685		
9	0.098557872	0.246027397		
10	0.17447508	0.171506849		
11	0.078158858	0.202739726		
12	0.088545169	0.348493151		
13	0.091309871	0.256438356		
14	0.131883733	0.183013699		
15	0.126503773	0.27890411		
16	0.087499066	0.177534247		
17	0.115146081	0.209863014		
18	0.118583277	0.247671233		
19	0.128371815	0.278356164		
20	0.110263894	0.302561644		

two columns of data, 1st row contains the vector/variable/attribute names.

Part



3 Running the program

The best way to familiarize yourself with the program is to look at accompanying video help available at the [webpage](#) from which you downloaded this software.

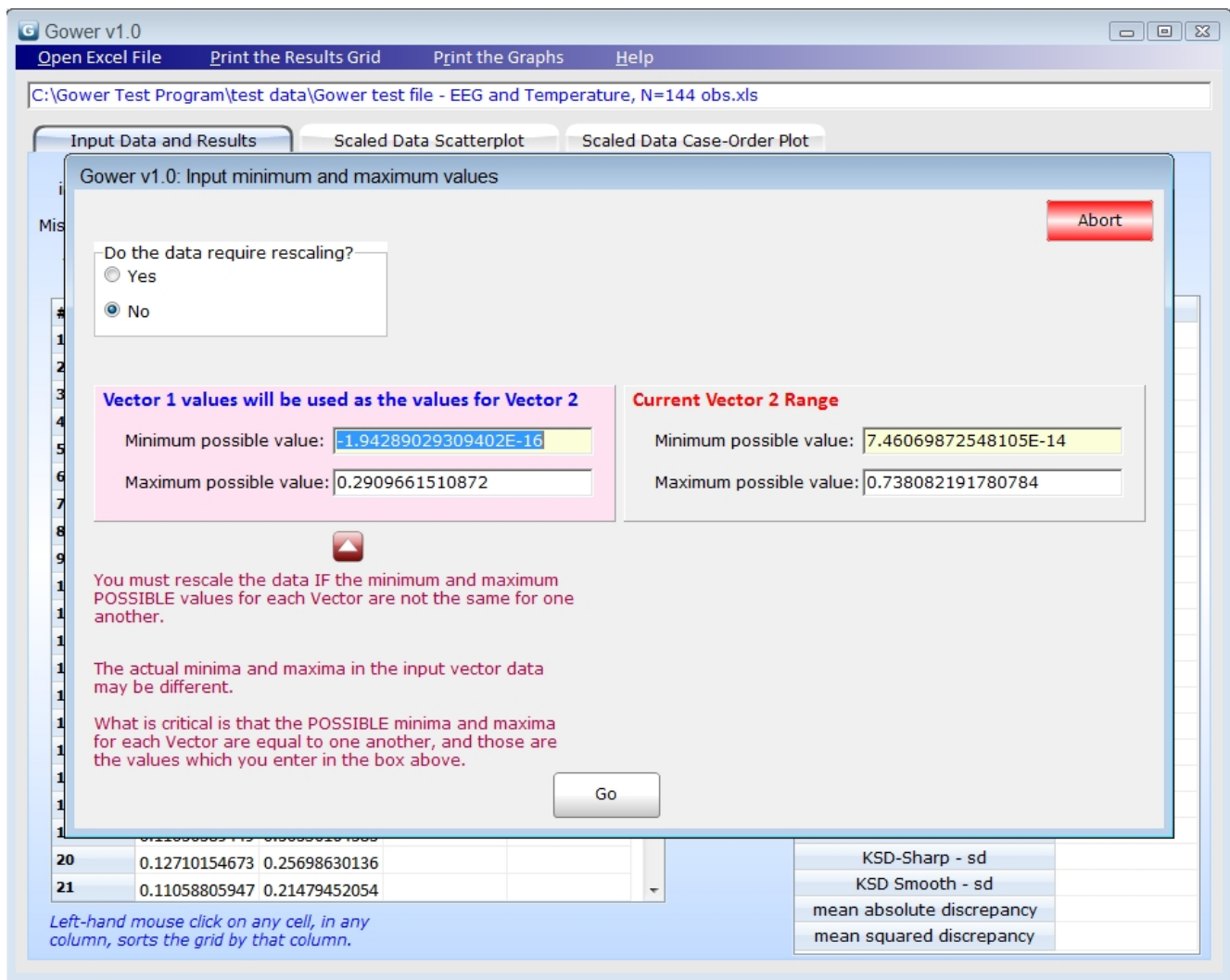
However, a quick overview follows below ..

When you click on the program icon, all you can do is either open a file or look at this help.



So, you open the file. The program has a "memory". The first time you search for and open a file, the program remembers where it was so that next time you run the program (or search for a new input file), it goes to the most recent subdirectory you used previously.

OK - now you have your file ..



There is a lot to take in here ...

The [Gower](#) (and [DSEs](#)) both express discrepancy as [relative](#) to the maximum possible discrepancy which could be observed with these two attributes. Now, that can either be the range of values within the actual vectors of data at hand, or a "could actually have been these values".

For example, if using a rating scale which varies between 0 and 10, two raters who rate 20 people (the two vectors of data, one per rater) might only have used values between 6 and 10, with Rater 1 using values between 8 and 10, and Rater 2, 6 and 8. So, the data-based minimum and maximum values would be 8 and 10 respectively for Rater 1's data, and 6 and 8 for Rater 2.

The program computes the minima and maxima for each input vector of data and displays this initially (as the above image shows).

In the case of our hypothetical rata data, knowing that ratings could have been given between 0 and 10, we would change the minimum and maximum for vector 1 to 0 and 10 respectively.

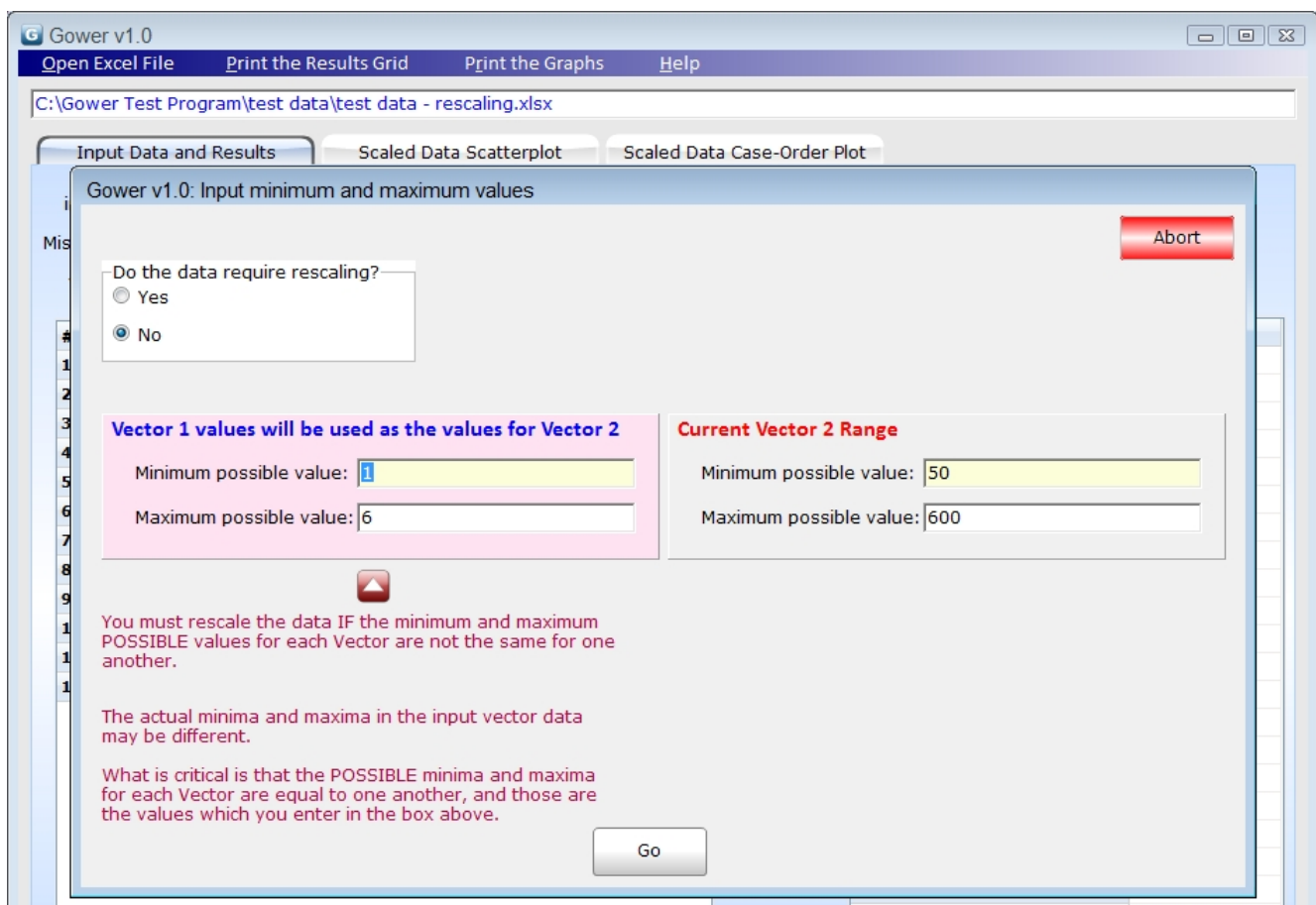
This is fine when both vectors use a common range of measurement .

In the case of the [common-metric](#) scaled EEG and Temperature data example above, all we know is that the minimum and maximum possible values which could be observed are between 0 and 1, so we would set the minimum and maximum for vector 1 accordingly. This assumes that Vector 2's values also range between these values.

Rather than go into pages of written description and "nuance" - just go look at the relevant [AV helpfile](#) on this issue.

VERY IMPORTANT! What if Vector 2 data is on a completely different measurement scale from Vector 1?

Well, look at this file ...



Here we see that Vector 1 values range between 1 and 6, while Vector ranges between 50 and 600.

Using two variable vectors with differing minimum and maximum values produces incorrect agreement measures .. because the relative change in one variable vector is not equal to the relative change in another (because of the inequality which exists between the minima and maxima).

In short, the [Gower](#) (and DSEs, and KSD coefficients - i.e. all agreement coefficients) **must be computed using attribute/variable vectors which share exactly the same measurement range.**

The program advises you about this as you can see above ... clearly in this case, the metric ranges are unequal. So .. we need to [rescale](#) the data.

Rescaling the Data

We now click on the option "Yes" in the box asking the question "Do the data need rescaling?"

What we see is this ...

C:\Gower Test Program\test data\test data - rescaling.xlsx

Input Data and Results Scaled Data Scatterplot Scaled Data Case-Order Plot

Gower v1.0: Input minimum and maximum values

Do the data require rescaling?

☒ Yes

☐ No

Which vector is to be the reference vector?

☒ Vector 1

☐ Vector 2

Vector 1 values will be used as the values for Vector 2

Minimum possible value: 1

Maximum possible value: 6

Current Vector 2 Range

Minimum possible value: 50

Maximum possible value: 600

If Vector 1 is to be the "Reference", that means Vector 2s current values will be rescaled into the range of the minimum and maximum values above (you can change them).

You might need to change the Vector 2 current minimum and maximum values, -IF- these are not the POSSIBLE minimum and maximum for that Vector. Otherwise, the current data (actual observed) minimum and maximum values that you see for Vector 2 will be used.

Go

Abort

We could use Vector 1 as the "reference" vector. That is, it's minimum and maximum values will be used to rescale Vector 2's values into the same range as Vector 1.

Whatever you type in for Vector 1 (*if you need to*) will be used to set the [common-metric](#) range for both Vectors.

However, also check Vector 2 now .. as you need to make sure that the POSSIBLE minimum and maximum values are used for Vector 2, not just the ones you see on display - UNLESS these really are the possible minimum and maximum values for that Vector variable.

Alternatively, we might choose to use Vector 2 as the "reference" vector. That is, it's minimum and maximum values will be used to rescale Vector 1's values into the same range as Vector 2.

C:\Gower Test Program\test data\test data - rescaling.xlsx

Input Data and Results Scaled Data Scatterplot Scaled Data Case-Order Plot

Gower v1.0: Input minimum and maximum values

Do the data require rescaling?
☒ Yes
☐ No

Which vector is to be the reference vector?
☐ Vector 1
☒ Vector 2

Current Vector 1 Range

Minimum possible value: 1
Maximum possible value: 6

Set Vector 2 Range

Minimum possible value: 50
Maximum possible value: 600

If Vector 2 is to be the "Reference", that means Vector 1's current values will be rescaled into the range of the minimum and maximum values above (you can change them).

You might need to change the Vector 1 current minimum and maximum values, -IF- these are not the POSSIBLE minimum and maximum for that Vector. Otherwise, the current data (actual observed) minimum and maximum values that you see for Vector 1 will be used.

Go

Abort

Whatever you type in for Vector 2 (*if you need to*) will be used to set the [common-metric](#) range for both Vectors.

However, also check Vector 1 now .. as you need to make sure that the POSSIBLE minimum and maximum values are used for Vector 1, not just the ones you see on display - UNLESS these really are the possible minimum and maximum values for that Vector variable.

The relevant [AV helpfile](#) on this issue helps explain things even further. This is tricky!

Anyway, let's do it both ways for now .. first we rescale Vector 2s values into the same metric as Variable Vector 1. We know Vector 1s possible minimum and maximum values are indeed 1 to 6. So no change there. But, we need to change Vector 2's minimum to 0, because these measurements could have varied between 0 and 600; we just happened to observe a minimum value of 50.

The screen looks like:

The screenshot shows a software window titled "Gower v1.0: Input minimum and maximum values". The window has three tabs: "Input Data and Results" (selected), "Scaled Data Scatterplot", and "Scaled Data Case-Order Plot". The dialog contains the following elements:

- Do the data require rescaling?** with radio buttons for "Yes" (selected) and "No".
- Which vector is to be the reference vector?** with radio buttons for "Vector 1" (selected) and "Vector 2".
- Set Vector 1 Range** (pink background):
 - Minimum possible value: 1
 - Maximum possible value: 6
- Current Vector 2 Range** (white background):
 - Minimum possible value: 0
 - Maximum possible value: 600
- A red square button with a white upward-pointing arrow.
- Red text instructions: "If Vector 1 is to be the 'Reference', that means Vector 2s current values will be rescaled into the range of the minimum and maximum values above (you can change them)."
- Red text instructions: "You might need to change the Vector 2 current minimum and maximum values, -IF- these are not the POSSIBLE minimum and maximum for that Vector. Otherwise, the current data (actual observed) minimum and maximum values that you see for Vector 2 will be used."
- An "Abort" button in the top right corner.
- A "Go" button at the bottom center.

click on the Go button and we see ..

The screenshot shows the Gower v1.0 software interface. The main window displays the 'Input Data and Results' tab. On the left, there are input fields for 'input data N: 12', 'Missing Data N: 0', and 'valid case N: 12'. Below these is a table with 5 columns: '#', 'var 1', 'var 2', 'Scaled Col#1', and 'Scaled Col#2'. The table contains 12 rows of data. To the right of the table, there is a red banner that says 'RESCALING IMPLEMENTED'. Below this, there is a text box that says '<Vector 2: var 2> has been rescaled into the metric of <Vector 1: var 1> using the minimum and maximum values specified for Vector 1.' To the right of the table, there is a red text box that says 'Right-hand mouse click anywhere on either grid copies it to the clipboard'. Below the table, there is a blue text box that says 'Left-hand mouse click on any cell, in any column, sorts the grid by that column.' To the right of the table, there is a table with 2 columns: 'COEFFICIENT' and 'Value'. This table contains the following data:

COEFFICIENT	Value
Pearson (Monotonicity)	-0.7940
Gower	0.4528
DSE-s	0.3775
KSD-s Sharp (range/6)	0.1467
KSD-s Smooth (range/3)	0.3840
ICC-1	-0.7770
ICC-2	-0.8949
ICC-3	-0.7901
DIAGNOSTICS	
No. of valid cases	12
Original Minimum Vector 1	1.00
Original Maximum Vector 1	6.00
Original Minimum Vector 2	50.00
Original Maximum Vector 2	600.00
Uses Minimum Vector 1	1.00
Uses Maximum Vector 1	6.00
KSD-Sharp - sd	0.8333
KSD Smooth - sd	1.6667
mean absolute discrepancy	2.7361
mean squared discrepancy	9.6887

The [Results Grid](#) help section defines the various entries.

Here we are more concerned with the two columns in blue text in the data grid, "Scaled Col #1" and "Scaled Col #2". These show that Vector 1 (**var 1**) is not rescaled at all (as it is the reference vector), but Vector 2's data (**var 2**) is now rescaled into the metric of Vector 1. Note the maximum value for Vector 2 (600) is now (6) - which is the maximum value for Vector 1.

Now we will rescale Vector 1's values into the same metric as Variable Vector 2. We know Vector 2's possible minimum and maximum values are 0 to 600, so we make that change. But, we there is no need to need to change Vector 1's minimum or maximum as these actually are the possible values which we happened to observe in the input data.

What we do to run this analysis is click on Open Excel File menu button, open the file again, and set the new constraints ... (don't forget to set Vector 2's minimum to 0):

The screen looks like:

C:\Gower Test Program\test data\test data - rescaling.xlsx

Input Data and Results Scaled Data Scatterplot Scaled Data Case-Order Plot

Gower v1.0: Input minimum and maximum values

Do the data require rescaling?
☒ Yes
☐ No


Which vector is to be the reference vector?
☐ Vector 1
☒ Vector 2

Current Vector 1 Range

Minimum possible value: 1
Maximum possible value: 6

Set Vector 2 Range

Minimum possible value: 0
Maximum possible value: 600



If Vector 2 is to be the "Reference", that means Vector 1s current values will be rescaled into the range of the minimum and maximum values above (you can change them).

You might need to change the Vector 1 current minimum and maximum values, -IF- these are not the POSSIBLE minimum and maximum for that Vector. Otherwise, the current data (actual observed) minimum and maximum values that you see for Vector 1 will be used.

Go

Abort

Clicking on the Go button ...

Gower v1.0
 Open Excel File Print the Results Grid Print the Graphs Help

File: J:\Gower Test Program\test data\test data - rescaling.xlsx

Input Data and Results Scaled Data Scatterplot Scaled Data Case-Order Plot

input data N: 12
 Missing Data N: 0
 valid case N: 12

#	var 1	var 2	Scaled Col#1	Scaled Col#2
1	1	600	0.0000	600.0000
2	2	500	120.0000	500.0000
3	3	400	240.0000	400.0000
4	4	300	360.0000	300.0000
5	5	300	480.0000	300.0000
6	6	400	600.0000	400.0000
7	1	600	0.0000	600.0000
8	2	500	120.0000	500.0000
9	3	400	240.0000	400.0000
10	4	100	360.0000	100.0000
11	5	70	480.0000	70.0000
12	6	50	600.0000	50.0000

Left-hand mouse click on any cell, in any column, sorts the grid by that column.

RESCALING IMPLEMENTED
 <Vector 1: var 1>
 has been rescaled into the metric of
 <Vector 2: var 2>
 using the minimum and maximum values
 specified for Vector 2.

COEFFICIENT	Value
Pearson (Monotonicity)	-0.7940
Gower	0.4528
DSE-s	0.3775
KSD-s Sharp (range/6)	0.1467
KSD-s Smooth (range/3)	0.3840
ICC-1	-0.7770
ICC-2	-0.8949
ICC-3	-0.7901

Right-hand mouse click anywhere on either grid copies it to the clipboard

DIAGNOSTICS	
No. of valid cases	12
Original Minimum Vector 1	1.00
Original Maximum Vector 1	6.00
Original Minimum Vector 2	50.00
Original Maximum Vector 2	600.00
Uses Minimum Vector 2	0.00
Uses Maximum Vector 2	600.00
KSD-Sharp - sd	100.0000
KSD Smooth - sd	200.0000
mean absolute discrepancy	328.3330
mean squared discrepancy	139,517.0000

The results for all coefficients are the same (Gower, DSE-s, KSD etc.) because the rescaling is linear- but you can see the data are now expressed in a common metric defined by Vector 2s minimum and maximum, ranging between 0 and 600.

You can print the results grid, and look at/print the scatterplot of the data. Just click on the relevant tabs or menu options. Again, checkout one of the video helpfiles available at the [webpage](#) from which you downloaded this software.

The [Results Grid](#) help section defines the various entries.

 Paul Barrett, (paul@pbarrett.net)

Part

IV

4 The Results Grid

Here I define the entries in the results grid ...

The best way to familiarize yourself with the program is to look at accompanying video help available at the [webpage](#) from which you downloaded this software.

I'll use the file: test data - 640 cases, 2-raters, range (0-50).xls

The data is taken from 2 raters who used a 0-50 range behavioral rating scale, rating 640 job applicants over a whole year, in an assessment center. What we want to do is evaluate the interrater reliability of these raters over the whole year. The cases are "block-ordered" over time. That is, cases were entered into the file in blocks, corresponding to each months' throughput of candidates in the assessment center. So, the data do correspond to a kind of case-ordering, which might be useful for a [case-order plot](#).

Inputting the data file ...

Gower v1.0

Open Excel File Print the Results Grid Print the Graphs Help

C:\Gower Test Program\test data\test data - 640 cases, 2-raters, range (0-50).xls

Input Data and Results Scaled Data Scatterplot Scaled Data Case-Order Plot

Gower v1.0: Input minimum and maximum values

Do the data require rescaling?

☐ Yes

☒ No

Vector 1 values will be used as the values for Vector 2

Minimum possible value: 2

Maximum possible value: 50

Current Vector 2 Range

Minimum possible value: 2

Maximum possible value: 46

You must rescale the data IF the minimum and maximum POSSIBLE values for each Vector are not the same for one another.

The actual minima and maxima in the input vector data may be different.

What is critical is that the POSSIBLE minima and maxima for each Vector are equal to one another, and those are the values which you enter in the box above.

Go

Left-hand mouse click on any cell, in any column, sorts the grid by that column.

20	32	22			
21	35	28			

KSD-Sharp - sd

KSD Smooth - sd

mean absolute discrepancy

mean squared discrepancy

the first thing we have to do is change the range to the maximum possible rating range which could have been used by the raters (0 to 50) ... Given both vectors use the same-metric rating scale, we just change the minimum value from 2 to 0 - without any rescaling ... as shown below ...

Vector 1 values will be used as the values for Vector 2

Minimum possible value: 0

Maximum possible value: 50

Current Vector 2 Range

Minimum possible value: 2

Maximum possible value: 46

You must rescale the data IF the minimum and maximum POSSIBLE values for each Vector are not the same for one another.

The results are:

Gower v1.0

Open Excel File Print the Results Grid Print the Graphs Help

C:\Gower Test Program\test data\test data - 640 cases, 2-raters, range (0-50).xls

Input Data and Results Scaled Data Scatterplot Scaled Data Case-Order Plot

input data N: 640
Missing Data N: 0
valid case N: 640

NO RESCALING

uses the specified minimum and maximum for
<Vector 1: Rater1>
for the maximum possible discrepancy range.

#	Rater1	Rater2	Scaled Col#1	Scaled Col#2
1	30	30	30.0000	30.0000
2	30	30	30.0000	30.0000
3	30	18	30.0000	18.0000
4	35	22	35.0000	22.0000
5	30	18	30.0000	18.0000
6	32	25	32.0000	25.0000
7	32	25	32.0000	25.0000
8	26	25	26.0000	25.0000
9	27	25	27.0000	25.0000
10	33	20	33.0000	20.0000
11	32	16	32.0000	16.0000
12	35	34	35.0000	34.0000
13	35	25	35.0000	25.0000
14	37	20	37.0000	20.0000
15	31	27	31.0000	27.0000
16	31	29	31.0000	29.0000
17	28	24	28.0000	24.0000
18	31	23	31.0000	23.0000
19	33	25	33.0000	25.0000
20	32	22	32.0000	22.0000
21	35	28	35.0000	28.0000

Left-hand mouse click on any cell, in any column, sorts the grid by that column.

Right-hand mouse click anywhere on either grid copies it to the clipboard

COEFFICIENT	Value
Pearson (Monotonicity)	0.0136
Gower	0.8686
DSE-s	0.8332
KSD-s Sharp (range/6)	0.7093
KSD-s Smooth (range/3)	0.8944
ICC-1	-0.0509
ICC-2	0.0120
ICC-3	0.0136
DIAGNOSTICS	
No. of valid cases	640
Original Minimum Vector 1	2.00
Original Maximum Vector 1	50.00
Original Minimum Vector 2	2.00
Original Maximum Vector 2	46.00
Uses Minimum Vector 1	0.00
Uses Maximum Vector 1	50.00
KSD-Sharp - sd	8.3333
KSD Smooth - sd	16.6667
mean absolute discrepancy	6.5687
mean squared discrepancy	69.5250

Let's take each area in turn ..

COEFFICIENT	Value
Pearson (Monotonicity)	0.0136
Gower	0.8686

These are the [Pearson](#) and [Gower](#) coefficients respectively. Note that if you only used/interpreted a Pearson correlation to measure rater reliability, you'd conclude there was none. Yet the Gower indicates 87% agreement. What these two coefficients are telling you is that the raters rate similarly to one another, but their ratings are not [monotonically related](#) (they don't co-vary).

DSE-s	0.8332
-------	--------

This is the double-scaled euclidean distance, expressed as a similarity coefficient ([DSE-s](#)). Slightly lower than the Gower, but still pretty high.

KSD-s Sharp (range/6)	0.7093
KSD-s Smooth (range/3)	0.8944

The kernel smoothed distance coefficients, expressed as similarities. The top one uses a much sharper criterion to express similarity than bottom one. Remember, the [KSD-s](#) coefficient "shapes" "discrepancy" - by increasing the magnitude of a linear discrepancy non-linearly, so that a larger discrepancies are penalized more than smaller ones.

ICC-1	-0.0509
ICC-2	0.0120
ICC-3	0.0136

The intraclass coefficient, [ICC-1](#), [ICC-2](#), [ICC-3](#) for Models 1, 2, and 3. Normally the chosen coefficient for rater reliability. You would conclude there is none if you relied upon these as your indicators. But they are totally misleading. I really can't stress enough that the use of the Pearson or ICC for any kind of reliability work is really, really inadvisable (including test-retest).

DIAGNOSTICS	
No. of valid cases	640
Original Minimum Vector 1	2.00
Original Maximum Vector 1	50.00
Original Minimum Vector 2	2.00
Original Maximum Vector 2	46.00
Uses Minimum Vector 1	0.00
Uses Maximum Vector 1	50.00

These are just "information" indices associated with the analysis - the actual minimum and maximum values for each input vector are presented along with the minimum and maximum used for the analysis (reflecting the change we made).

Note also the "status" report about rescaling displayed over the top of the results grid ...

NO RESCALING	
uses the specified minimum and maximum for	
<Vector 1: Rater1>	
for the maximum possible discrepancy range.	

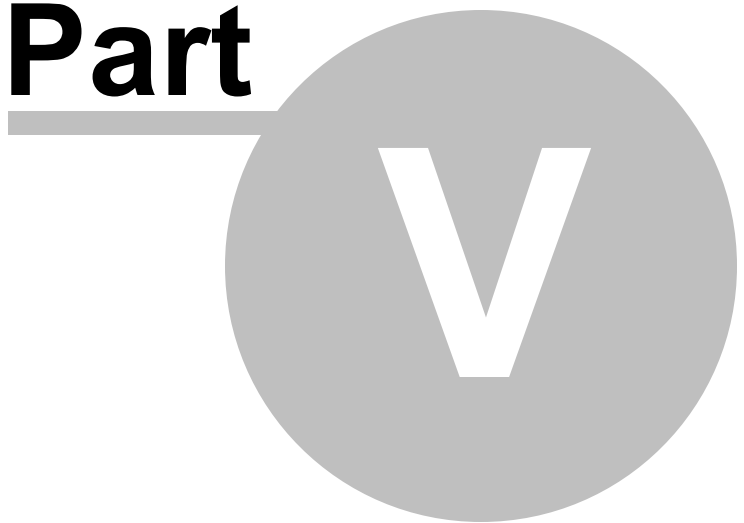
this will remind you whether or not you requested scaling, and which variable range is taken as the "referent".

KSD-Sharp - sd	8.3333
KSD Smooth - sd	16.6667
mean absolute discrepancy	6.5687
mean squared discrepancy	69.5250

These are some diagnostics associated with the Kernel smoothed distance coefficient and the DSE-s and Gower. The [KSD](#) values are the size of the smoothing standard deviation parameter. The discrepancies are raw (absolute) and squared values, without being scaled by the maximum possible discrepancy. See the help section entitled [Relative v Simple Discrepancy](#).

Paul Barrett, (paul@pbarrett.net)

Part



5 The data plots

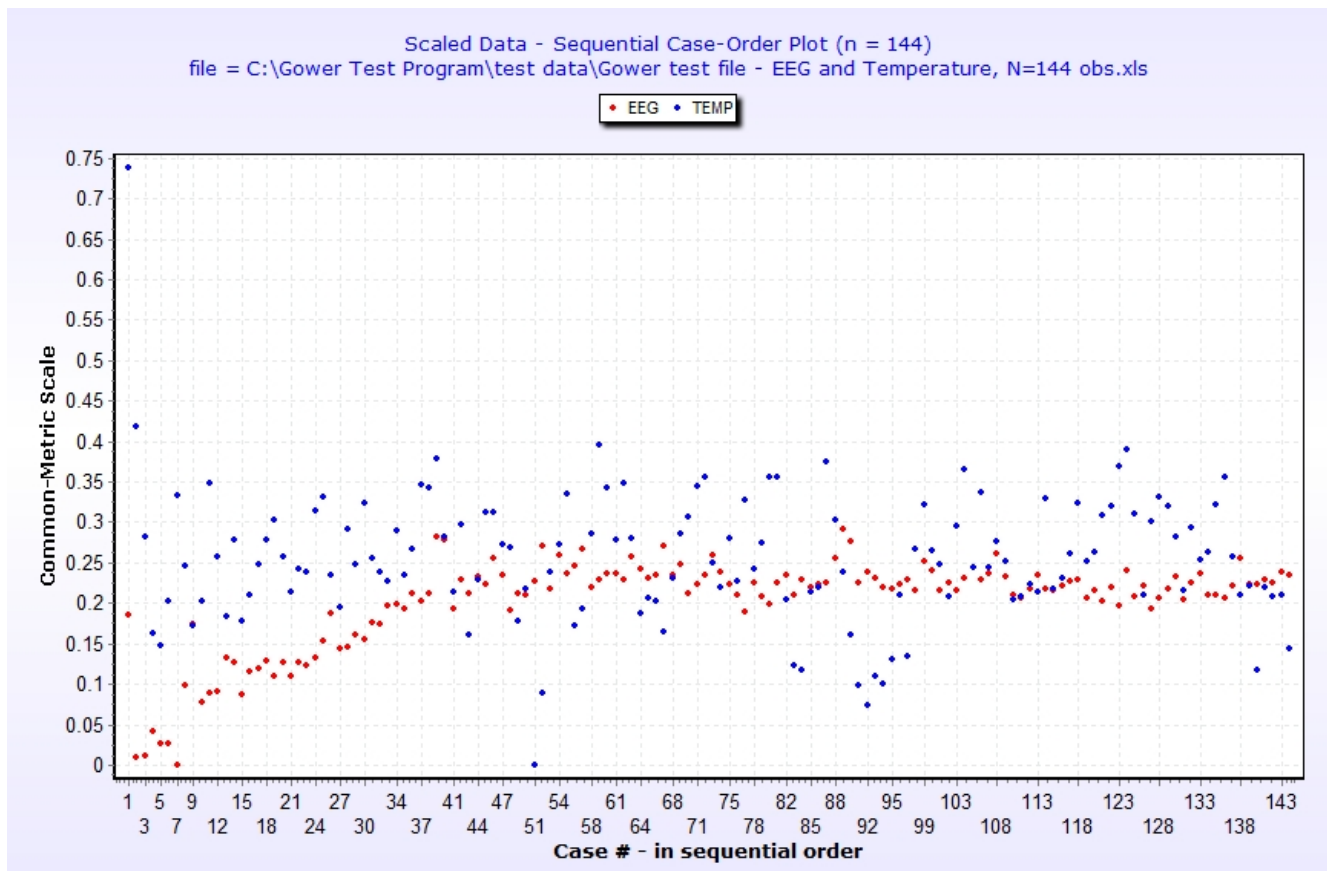
Two kinds of data plot are provided with this routine, a straightforward scatterplot of the two vector variables, and a case-order plot of the two.

The scatterplot plots the data given their respective values on both variables.

The case-order plot presents the data in "case-order". That is, it plots both vector values in sequential case order, each pair of values plotted against case # in the file. This is useful where the case order is systematic (as in observations taken over time for example).

The plot is only useful when the variables are expressed in a [common-metric](#), because then you get to see the shape/trends within the data without the axis scale distorting the visual perception (where the same scale has to account for completely different metric EEG and Temperature values).

If we look at the EEG and Temperature file, where these pairs of values were observed over time, and the order of cases in the file is by time-point, we see ..



You can see 'by eye' the agreement between EEG and Temperature .. the Gower is 0.91 for these data.

Part

VI

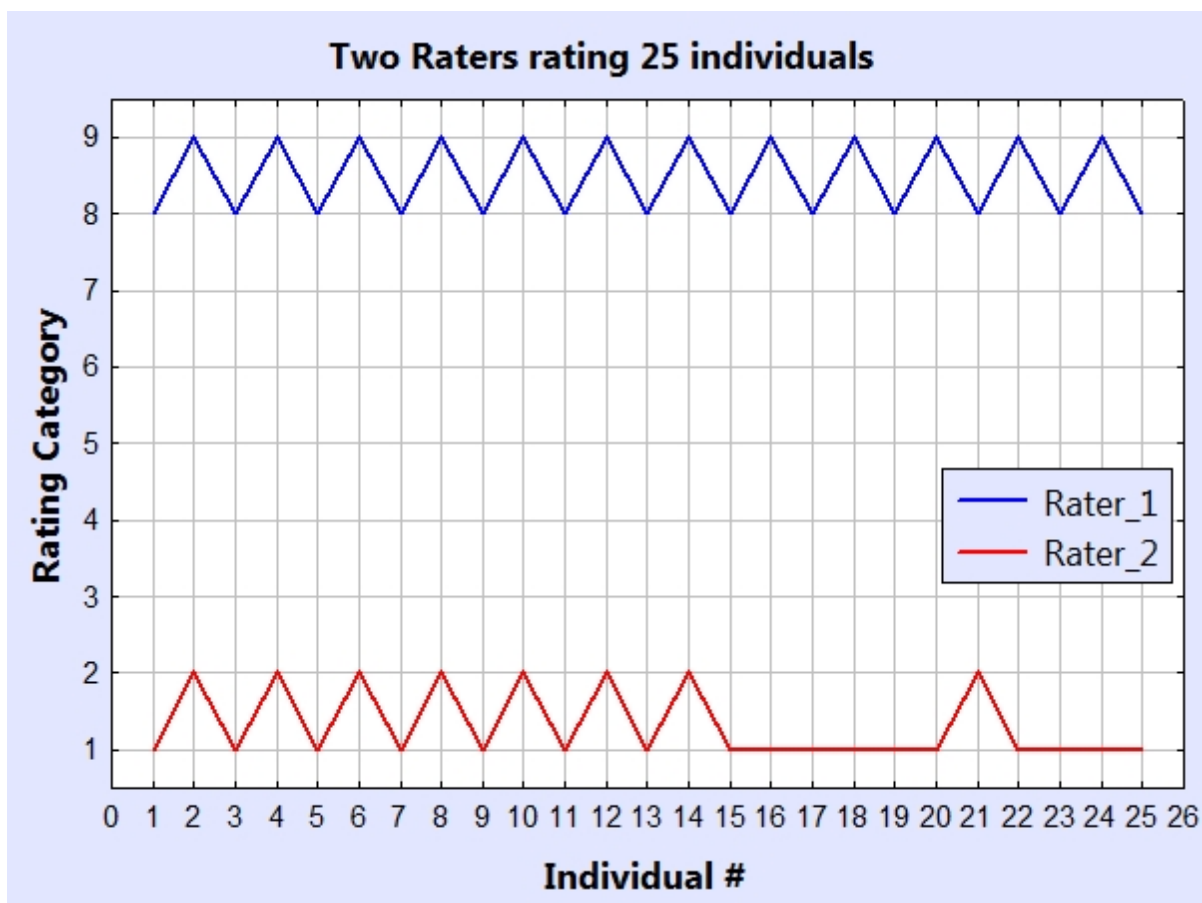
6 The Coefficients

In terms of "classes" of coefficients, you think of them in terms of their capacity to estimate [agreement vs monotonicity](#).

The [Pearson](#) is an estimate of monotonicity, not magnitude agreement. Likewise [ICC-1](#) and [ICC-3](#). [ICC-2](#) tends to show more sensitivity to magnitude but it's not always the case.

In short, ICCs really are not very useful **IF** you want to make a "clean" assessment of monotonicity and magnitude agreement. ICC-2 just confounds the two.

For example, look at the two vectors of rating data below:



The rating range is between 1 and 9. The two raters are giving completely discrepant ratings - you really couldn't get much more disagreement! Yet, look at the coefficients:

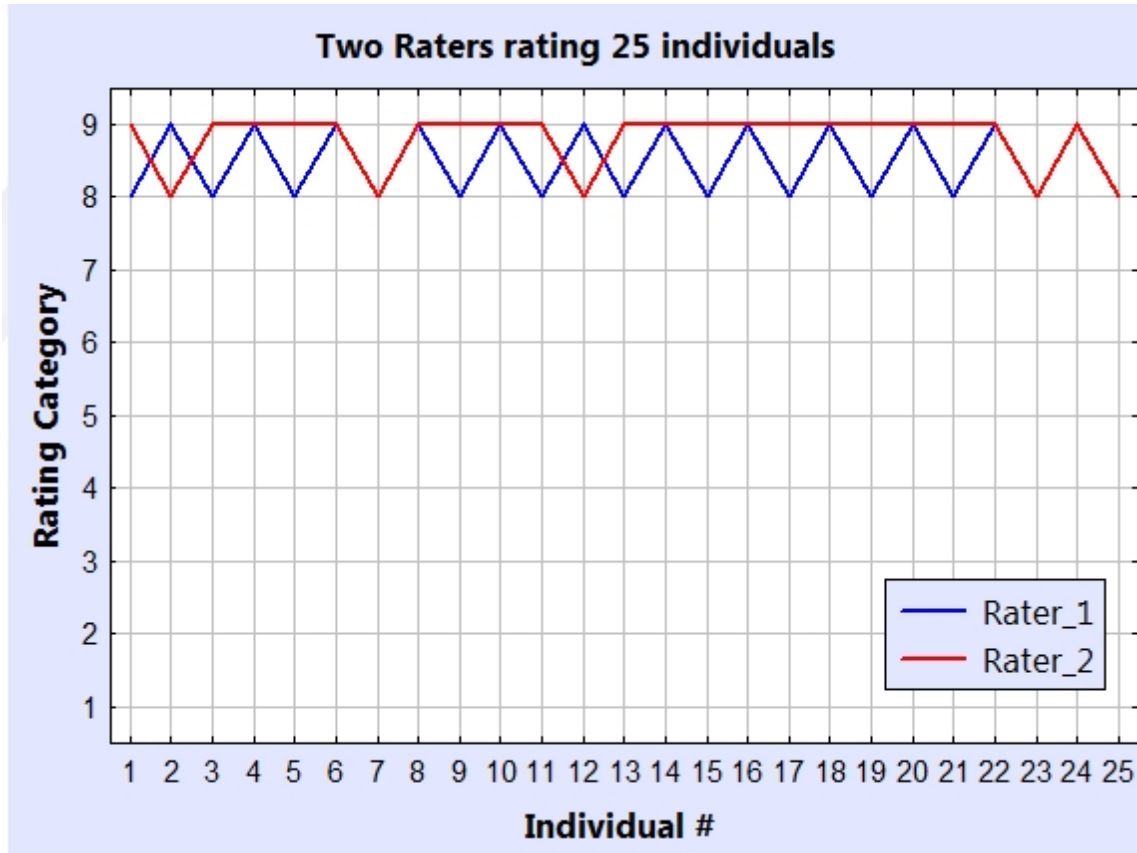
Pearson $r = .54$

icc1 = $-.97$

icc2 = $.00$

icc3 = .54
 Gower = .10
 DSEs = .10
 KSDs = .00

You might be tempted into thinking ICC-2 is working just fine ... **don't**. Look at the next graph below ..



The rating range is between 1 and 9. Our two raters are giving ratings in almost perfect agreement with one another - you really couldn't get much more agreement! Yet. look at the coefficients:

Pearson $r = .08$
 icc1 = $-.02$
icc2 = .07
 icc3 = $.08$
 Gow = $.94$
 DSEs = $.91$
 KSDs = $.88$

If you want a measure of monotonicity (how well observations co-vary with one another), use the Pearson correlation. It's perfect for this.

If you want a measure of magnitude agreement (how well the actual magnitudes of observation agree with one another, *irrespective of whether they "co-vary"*), use the [DSE-s](#), [KSD-s](#), or [Gower](#) ...

An investigator needs both sources of information to come to a complete decision about the association between variables.

Paul Barrett, (paul@pbarrett.net)

6.1 Pearson monotonicity

The formula for the Pearson correlation coefficient is:

$$r = \frac{\text{COV}_{xy}}{s_x \cdot s_y} \quad \text{where } N = \text{the number of paired observations}$$

COV_{xy} is the covariance between variables x and y calculated as:

$$\text{COV}_{xy} = \frac{\sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})}{N-1} \equiv \frac{\sum_{i=1}^N x_i y_i - \left(\frac{\sum_{i=1}^N x_i \sum_{i=1}^N y_i}{N} \right)}{N-1}$$

$$s_x = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}} \equiv \sqrt{\frac{\sum_{i=1}^N x_i^2 - N \cdot \bar{x}^2}{N-1}}$$

and likewise for s_y

The coefficient varies between +1.0 and -1.0, with +1.0 indicating maximum similarity, -1.0 = maximum inverse similarity, and 0.0 indicating no relationship whatsoever. The Pearson correlation indexes this relationship by first transforming the raw data, standardizing each variable by its

respective mean and standard deviation, then computing the covariance between the two transformed variables. Absolute magnitude information is lost in order to enable this scale-free computation of relationship (monotonicity).

6.2 ICC-1 (intraclass)

Normally applied to ratings as in an interrater reliability estimation procedure; each person/case to be rated is rated by a unique rater, with each rater randomly selected from a larger population (a one-way ANOVA random effects model). Specifically, for every case or item to be rated, there is a unique rater. Each rater makes only one rating decision. This model assumes you have a large pool of raters, who are randomly assigned to make one rating per case per variable. So, for a study in which we rate 10 cases on 5 variables, we would need 50 raters. The ANOVA formula is:

$$r_{icc1}^2 = \frac{MS_p - WMS}{MS_p + (n_r - 1) * WMS}$$

where MS_p = Between Cases mean square

n_r = number of Vectors and n_p = number of cases

MS_{res} = Residual mean square

WMS = Within cases mean square

MS_r = Between Vectors ("measures") mean square

with

$$WMS = \frac{[(MS_r * (n_r - 1)) + (MS_{res} * (n_p - 1) * (n_r - 1))]}{n_p * (n_r - 1)}$$

The formula can be applied to Vector data, where the vectors of data are considered random samples from some hypothetical population, with the vector attributes also independently sampled.

The Intraclass coefficients vary between 0 and +1 (where 0 = no relationship/reliability at all). They can under certain circumstances assume negative values but this is very rare with anything but random rating data.

6.3 ICC-2 (intraclass)

Normally applied to ratings as in an interrater reliability estimation procedure; every case is rated by each rater. We assume the raters are randomly selected from some population of raters (a two-way random effects model). In essence, each rater rates all cases on all variables. This is the default model that covers most rating situations. For example, for a study in which we rate 10 cases on 5 variables, we would need at least 2 raters in order to assess interrater reliability. Each rater would make $(10 \times 5) = 50$ rating judgments. The ANOVA formula is:

$$r_{icc2} = \frac{MS_p - MS_{res}}{MS_p + (n_r - 1) * MS_{res} + \left(\frac{n_r * (MS_r - MS_{res})}{n_p} \right)}$$

where MS_p = Between Cases mean square

MS_{res} = Residual (interaction) mean square

MS_r = Between Vectors (measures) mean square

n_r = The number of Vectors

n_p = The number of cases

The formula can be applied to Vector data, where the vectors of data consist of independent observations on each case, with the vector attributes also independently sampled.

The Intraclass coefficients vary between 0 and +1 (where 0 = no relationship/reliability at all). They can under certain circumstances assume negative values but this is very rare with anything but random rating data.

6.4 ICC-3 (intraclass)

Normally applied to ratings as in an interrater reliability estimation procedure; every case is rated by each rater; every case is rated by each rater, BUT, in contrast to [ICC Model 2](#), we assume the raters are THE population of raters (a two-way, fixed-effects model). In essence, each rater rates all patients on all variables. For example, for a study in which we rate 10 patients on 5 variables, we would select say 2 raters in order to assess interrater reliability. Each rater would make (10*5)=50 rating judgments. However, it is assumed that these are the only two raters who will ever make ratings – no generalizability assumed to other raters. The ANOVA formula is:

$$r_{icc3} = \frac{MS_p - MS_{res}}{MS_p + (n_r - 1) * MS_{res}}$$

where MS_p = Between Cases mean square

MS_{res} = Residual (interaction) mean square

n_r = The number of Vectors

Here, each vector would be considered a unique vector of data, maybe from two conditions or two unique sources.

The Intraclass coefficients vary between 0 and +1 (where 0 = no relationship/reliability at all). They can under certain circumstances assume negative values but this is very rare with anything but random rating data.

6.5 DSE-s (double-scaled euclidean)

This index computes the squared discrepancy between two Vector's values, then divides this squared discrepancy by the maximum possible squared discrepancy for that case/variable. Summing and taking the square root of these "scaled" discrepancies across cases/variables yields a scaled Euclidean distance. But, the metric of this scaled and cumulatively summed variable discrepancy distance varies between 0 and some value greater than 1.0. In order to scale this coefficient into a unit (0 to 1) metric, a further scaling operation takes place. That is, the initially scaled Euclidean distance is divided by the square root of the number of variables comprising the distance computation. This second scaling now produces a coefficient which always varies between 0 (no

distance between variables) to 1 (maximum possible distance between variables given the designated maximum and minimum values for each variable).

This dual scaling ensures that the coefficient is comparable between studies and samples where different variable magnitudes might otherwise distort a conventional Euclidean distance. Further, because the initial scaling of distance is linear (rather than the non-linear operation used within the more common solution of converting data to normalized z-scores prior to any distance calculation), the linear distance relations between magnitudes on the variables remains unchanged. Finally, in order to complete the process, the double-scaled distance is expressed as a similarity index by subtracting it from 1, thus yielding the double-scaled Euclidean similarity (DSE-S) measure, where 0 for this coefficient indicates maximum possible disagreement, and 1 indicates that all cases/variables possess identical rating magnitudes.

The formula for this coefficient is:

$$dse_s = 1 - \left[\frac{\sqrt{\sum_{i=1}^n \left(\frac{(case_{1i} - case_{2i})^2}{md_i} \right)}}{\sqrt{n}} \right] = 1 - \left[\frac{\sum_{i=1}^n \left(\frac{(case_{1i} - case_{2i})^2}{md_i} \right)}{n} \right]$$

where

n = the number of cases

md_i = (maximum-minimum possible value for case i)²

$case_{1i}$ = the value for case i of n from the first vector

$case_{2i}$ = the value for case i of n from the second vector

The similarity to the [Gower](#) coefficient is obvious, but these will produce different coefficient sizes and distribution densities as the Gower is based upon absolute value discrepancy while the double-scaled Euclidean is based upon squared discrepancies.

The DSE-S coefficient computes a scaled similarity coefficient, utilizing scaled discrepancies. It varies between 0 and +1, where +1 is equal to identity between the two vectors being compared.

6.6 KSD-s (kernel-smooth distance)

This coefficient is based upon a very simple idea that a distance function should be shaped in such a way that if the simple arithmetic unsigned difference between a person's attribute value and a target value is computed to be within a certain range, then the computed distance should reflect a very small distance, almost regardless of the actual distance. But, as that distance grows larger, then the computed distance should be accelerated in size. In short, an "inertial" effect was aimed for – translated into a distance metric. The coefficient itself is scaled as a measure of similarity, varying between 0 (maximal dissimilarity) to 100 (identity).

$$KSD = \frac{\sum_{i=1}^n \left[\frac{1}{s\sqrt{2\pi}} e^{-\left[\frac{(case_{1i} - case_{2i})^2}{2s^2} \right]} \right] \cdot (100 \cdot (s \cdot \sqrt{2\pi}))}{n}$$

where

$$s = \frac{\text{range}_i}{\text{smoother}}$$

smoother = the "smoothing" parameter

range_i = (maximum-minimum value) for case/variable *i*

case_{1i} = the rating value for case *i* of *n* from the first vector

case_{2i} = the rating value for case *i* of *n* from the second vector

n = the number of cases

The key to using this coefficient is selecting an appropriate value of the smoother constant which produces the desired inertial effect. The selection choice is application-specific; the function in fact needs to be calibrated for each specific application, taking into account the costs and benefits of a sharper or smoother distance/discrepancy function. Person-target profiling applications are the most readily understood in this regard, where profile similarity can be adjusted to reflect only very close matches, with even "nearly similar" is reduced to "no-match" with even tiny departures of a person's profile from a target. Likewise, if a broad screen is required, then the smoothing function

can be more gradual - providing a kind of "plateau" effect around the target value, before the discrepancy between person and target is accelerated by the non-linear function.

In essence, this coefficient needs to be "calibrated", to match the "by eye" judgment of the user. That is, when plotting two profiles, or when deciding whether two values are to be adjudged similar to one another given the range of the rating scale being used, it is the user who has to decide when two values are to be adjusted "similar", and not the "mathematics". The KSD coefficient is sensitive only to magnitude discrepancy (not monotonicity). It also takes into account the range of ratings or score values from which the person-target, or rater reliability ratings are drawn.

For example, consider the comparison of two sets of scores from two Raters ...

Rater 1	Rater 2
3	4
4	3
3	4
4	3
3	4
4	3
3	4
4	3
4	4
4	4

If we apply the coefficient formula to the data in the table, with a KSD smoother value of 5,

Rater 1	Rater 2	KSD
3	4	45.783
4	3	45.783
3	4	45.783
4	3	45.783
3	4	45.783
4	3	45.783
3	4	45.783
4	3	45.783
4	4	100
4	4	100

the average of these 10 rater pairs is 56.63% similarity, or using a 0-1 coefficient scaling, 0.566.

Incidentally, the Pearson correlation for these data is -0.67 , ICC model 2 = -0.80 , and model 3 = -0.67 , while the Gower is $+0.80$, and DSE-S = $+0.78$.

Such data will produce highly negative Pearson and Intraclass correlations. Assuming a rating scale range of 1-5, the KSD coefficient with smoother factor of 5 is: 0.57. If we assume the scores are drawn from a range 1-20, then the KSD coefficient is 0.97. Clearly, the design of the coefficient introduces the element of relativity; a score discrepancy of 1 looks reasonably important when the range is just 1 to 5, but trivial when the range is 1-20.

In order to look at the effect of various smoother values over the measurement range, a interactive utility program is available from my [website](#) which allows a user to select various values and see the achieved smoothing/plateau effect. This file also contains further help information about the coefficient.

6.7 Gower agreement

Relative to the maximum possible absolute (unsigned) discrepancy between the two pairs of observations, the gower discrepancy coefficient indicates the % average absolute discrepancy between all pairs of observations. When expressed as a similarity coefficient (by subtracting it from 1), it indicates the % average similarity between all pairs of observations.

So, a Gower **similarity** coefficient of say 0.90 indicates that relative to the maximum possible absolute (unsigned) discrepancy between them, the observations agree *on average* to within 90% of each other's values.

If you change the value of that maximum possible discrepancy, then the Gower coefficient will change to reflect this, as the discrepancies between pairs of observations are divided (scaled) by that maximum possible discrepancy value.

E.g. if two observations differ by 5, and the measurement range of each observation is 10, then the relative discrepancy is 0.5. However, if the measurement range for each observation was say 100, then the relative discrepancy would be just 0.1.

But that's the whole point of the Gower, it tells you how discrepant (or similar) observations are, **RELATIVE** to how discrepant they could have been. A 5-point difference in a 10-point maximum measurement range is not very good. A 5-point difference between observations within a 100-point measurement range is pretty accurate.

The equation for the gower similarity index is:

$$Gower_{similarity} = 1 - \left[\frac{\sum_{i=1}^n \left(\frac{|obs_{1i} - obs_{2i}|}{range} \right)}{n} \right]$$

n = the number of cases

$range$ = the maximum possible discrepancy between the two variables

obs_{1i} = the observed value for case i of n in the first set of observations

obs_{2i} = the observed value for case i of n in the second set of observations

The Gower coefficient computes a scaled similarity coefficient, utilizing scaled discrepancies. It varies between 0 and +1, where +1 is equal to identity between the two vectors being compared.

For example, look at the following scores on two tests, with the minimum and maximum possible scores for each test between 0 and 50 ..

	Var1	Var2
1	30	5
2	30	5
3	30	5
4	30	5
5	30	5
6	30	5
7	30	5
8	30	5
9	30	5
10	30	5

The discrepancy between each pair of observations is 25, which is exactly half the maximum possible discrepancy range of 50. So each paired scaled discrepancy is exactly 0.5, with a resulting Gower and DSE-s of 0.5.

That really is the logic of a discrepancy-based coefficient in a nutshell. A value of 0.5 tells you that the similarity between observations is 50% of maximum (which = identity).

Another way of expressing this is that the average discrepancy between observations is one half of the maximum possible discrepancy.

However, we have to be careful here. For the Gower, "the maximum possible discrepancy" is the range of the data (the difference between the maximum possible and minimum possible values). The Gower is the average of the absolute discrepancies, divided through by the range, subtracted from 1 to provide the measure of similarity.

A Statistical Significance of a Gower Coefficient?

If you want a significance test for the Gower, use the Bootstrap program which is available for download from my [website](#). This computes the empirical sampling distribution for a Gower coefficient, taking into account the maximum possible discrepancy range, the kind of data (integer or reals) and the number of observations over which a Gower has been computed. It also allows you to test the significance of a difference between two Gowers, by constructing the empirical sampling distribution of the difference between two Gower coefficients, given their respective data characteristics (sample size, number type, and maximum possible discrepancy of observations in each sample).

There is a lot more information about this coefficient and the consequences of its practical application (and comparisons with other coefficients) in:

Barrett, P.T. (2010) [Test reliability and validity: The inappropriate use of the Pearson and other variance ratio coefficients for indexing reliability and validity](#). Technical Whitepaper #9, available at: www.pbarrett.net, from the [Whitepapers](#) link.

The coefficient itself was first presented to the research community in 1971 ...

Gower, J. C., 1971. [A general coefficient of similarity and some of its properties](#). Biometrics 27: 857-874.

Paul Barrett, (paul@pbarrett.net)

Part

VII

7 Agreement and Monotonicity

Sometimes, it is important to know whether observations are not only in agreement to one another, but whether they share a directional relation, and how strong such a relation is. That is, observations on one attribute tend to increase as observations on another increase, or whether as observations on one attribute increase, they decrease on the other (the monotonicity principle). The [Pearson](#) correlation and [ICC](#) coefficients do provide such information. Coefficients of agreements based upon absolute or squared discrepancies do not.

As [Barrett \(2010\)](#) has shown, monotonicity and agreement are independent from one another. Sometimes it is theoretically important that two variables' magnitudes covary; this is the essence of the definition of monotonicity. The Pearson correlation indexes this monotonicity by first transforming the raw data, standardizing each variable by its respective mean and standard deviation, then computing the covariance between the two transformed variables. Absolute magnitude information is lost in order to enable this scale-free computation of monotonicity.

But, sometimes, it is as important to know how similar the magnitudes of two variables are to one another, independent of whether the values covary. For example, the recorded temperature for two animals over time may be almost exactly equal to one another (to within $\pm .1$ °C) across say 20 time periods (allowing for random observation and instrument error). Because the error is random around the same temperature, the correlation between the two sets of observations would be expected to be zero (no monotonicity). But, an agreement index which is sensitive only to magnitude discrepancy would show that the similarity of observations is almost identical, given the range of variation in which temperature would have been able to fluctuate.

Barrett, P.T. (2010) *Test reliability and validity: The inappropriate use of the Pearson and other variance ratio coefficients for indexing reliability and validity*. Technical Whitepaper #9, available at: www.pbarrett.net, from the Whitepapers link.

Paul Barrett, (paul@pbarrett.net)

Part



8 Common-Metric Rescaling

Linear Rescaling into a Common Metric

When wanting to display and compare data from variables whose measurement metric is not the same, it is useful to rescale each variable's values into a convenient common metric. This is especially the case when displaying data with different ranges in multiple line-graphs and scatterplots, and for comparing multiple variables with coefficients such as the Gower agreement index.

Also, with specific regard to the [Gower](#) and other relative magnitude agreement indices, using two variable vectors with differing minimum and maximum values produces incorrect agreement measures .. because the relative change in one variable vector is not equal to the relative change in another (because of the inequality which exists between the respective minima and maxima).

Unlike conventional standardization, which transforms a variable's values by subtracting each observed value from the mean of all observed values and dividing this difference by the standard deviation of the values, the rescaling implemented here preserves the relativity between each variables' observations while rescaling the magnitudes into a common metric.

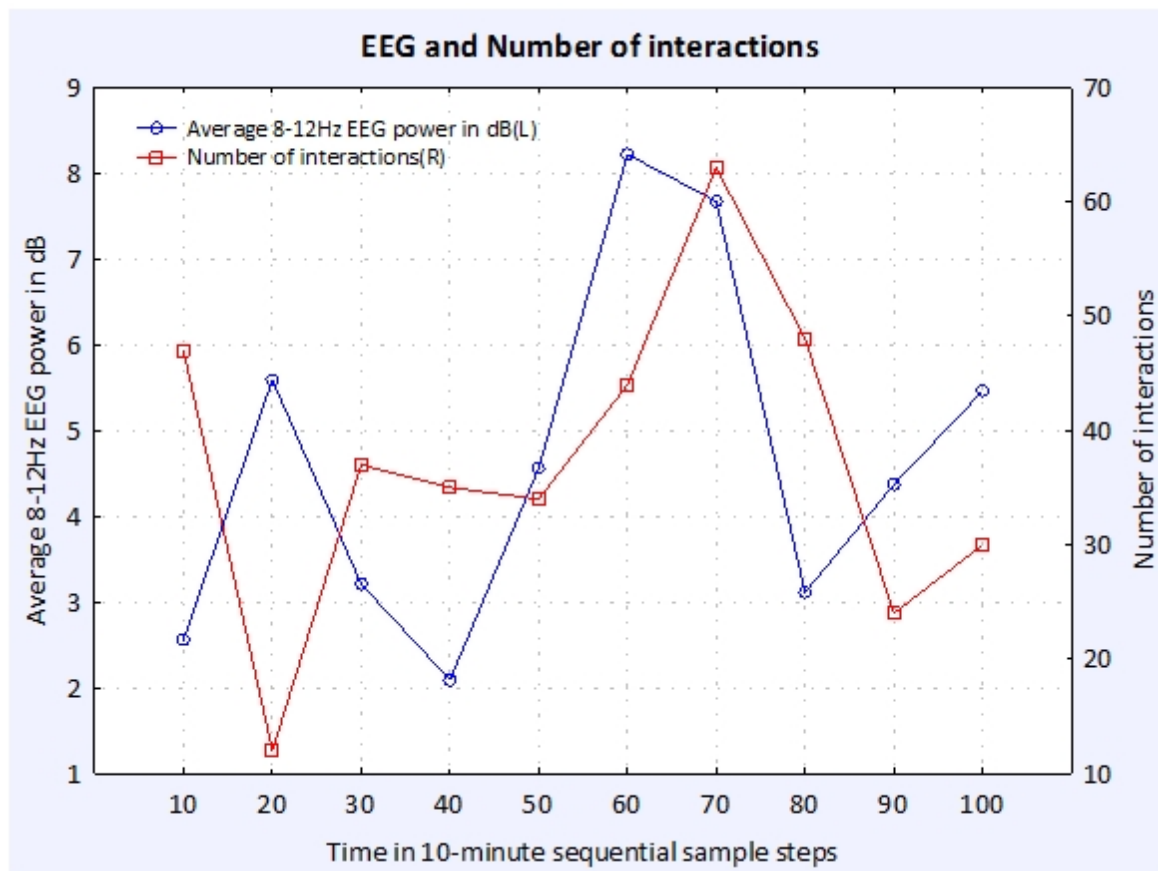
For example, displayed in Table 1 are data recorded from an animal over 10 x 10-minute intervals: average EEG output (dB) over 10 minutes within the 8-12Hz frequency band and the number of interactions with other animals observed over each 10-minute period.

Table 1

10 minute time points	Test spreadsheet for rescaling example	
	1 Average 8-12Hz EEG power in dB	2 Number of interactions
10	2.56	47
20	5.59	12
30	3.21	37
40	2.09	35
50	4.56	34
60	8.22	44
70	7.67	63
80	3.11	48
90	4.37	24
100	5.46	30

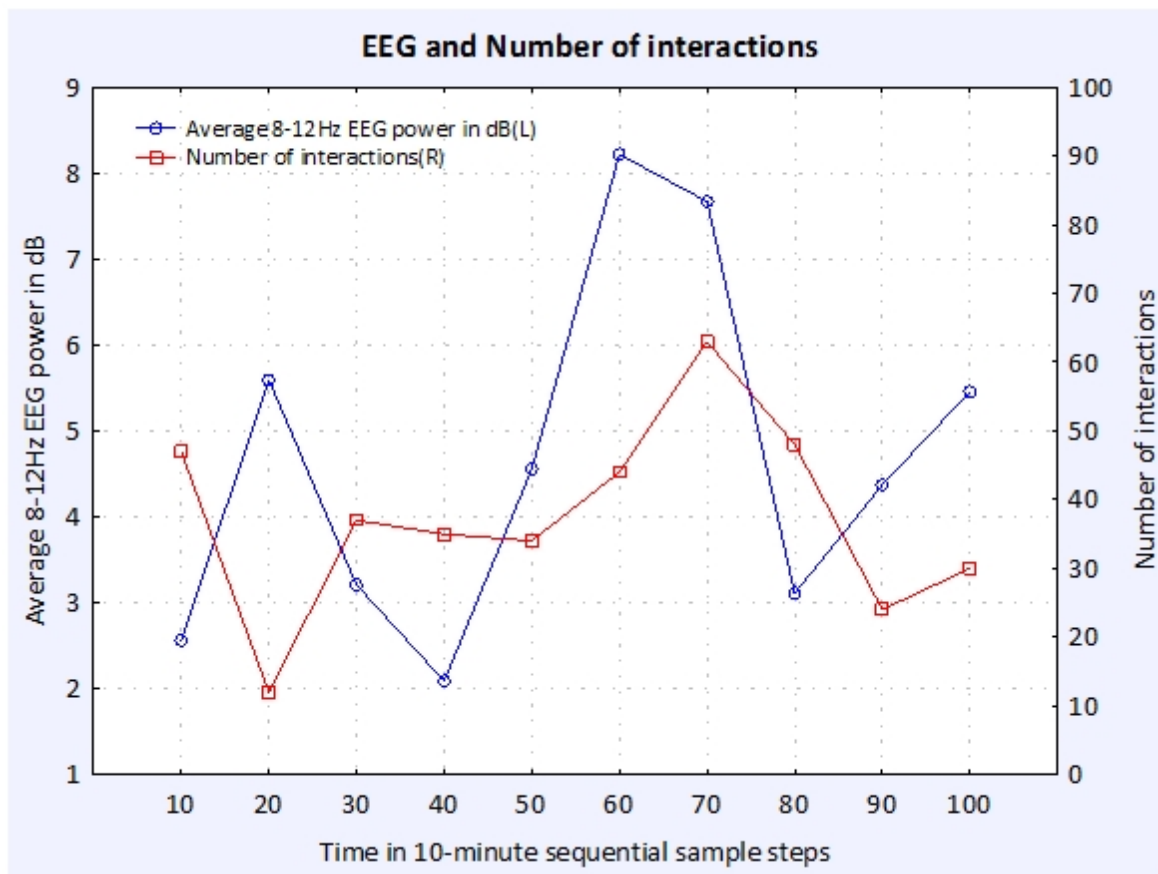
If we plot the data using a double-Y line plot, we have ..

Figure 1



However, the plot might also look like just by changing the right-Y axis values.

Figure 2



So if you have many such plots to generate, it is useful to express them all in a common-metric, so that visual comparison is not influenced by the particular range of values used within each plot. In this way, the shape and relative magnitudes within each series of values are maintained, but are now directly comparable across variables.

The rescaling is achieved by linearly regressing the original magnitude range for a variable onto the new range, using the original magnitude minimum and maximum for a variable, and the chosen metric minimum and maximum values (in essence, fitting a straight line between two points in 2-dimensional space). The result is that the conversion is linear over the range for all values.

The equations are:

$$b = \frac{\left(\sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i \right)}{\left(n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right)} \quad a = \frac{\left(\sum_{i=1}^n y_i - b \sum_{i=1}^n x_i \right)}{n}$$

where

$n = 2$ (always)

$x_1 =$ the original variable minimum

$x_2 =$ the original variable maximum

$y_1 =$ the minimum value of the new metric

$y_2 =$ the maximum value of the new metric

which, for just two values for each variable simplifies into:

$$x_{rescale} = \left\{ \left[\frac{(x_{old} - x_{min})}{(x_{max} - x_{min})} \right] * (y_{max} - y_{min}) \right\} + y_{min}$$

where

$x_{min} =$ minimum value of x_{old}

$x_{max} =$ maximum value of x_{old}

$y_{min} =$ minimum rescaled value of $x_{rescale}$

$y_{max} =$ maximum rescaled value of $x_{rescale}$

So, transforming the data in Table 1 into a new common metric whose values range between 0 and 1, we convert each variable using its own minimum and maximum.

An important consideration here is what values are to be used as representative of the minimum and maximum for a variable, the values from the sample of data, or values which represent the entire possible measurement range for a variable. For example, for EEG, do we use the observed minimum and maximum values in the sample of data we have observed, or the theoretically possible range of values (the resolution accuracy of the EEG preamplifier at 1-microvolt, through to some theoretical maximum value for EEG within that frequency band)? In general, with more or less

open-ended -range measurement, using the observed minimum and maximum values in the dataset is the most straightforward choice.

So, for Table 1 EEG, using the simple formula:

$$EEG_{rescale} = \left\{ \left[\frac{(EEG_{old} - EEG_{old\min})}{(EEG_{old\max} - EEG_{old\min})} \right] * (EEG_{new\max} - EEG_{new\min}) \right\} + EEG_{new\min}$$

where

$$EEG_{old\min} = 2.09, \quad EEG_{old\max} = 8.22, \quad EEG_{new\min} = 0, \quad EEG_{new\max} = 1$$

Likewise for number of interactions, whose minimum and maximum is specified to be between 12 and 63:

$$INTER_{rescale} = \left\{ \left[\frac{(INTER_{old} - INTER_{old\min})}{(INTER_{old\max} - INTER_{old\min})} \right] * (INTER_{new\max} - INTER_{new\min}) \right\} + INTER_{new\min}$$

where

$$INTER_{old\min} = 12, \quad INTER_{old\max} = 63, \quad INTER_{new\min} = 0, \quad INTER_{new\max} = 1$$

so, for example, the 1st value for EEG in Table 1 is 2.56, which converts into:

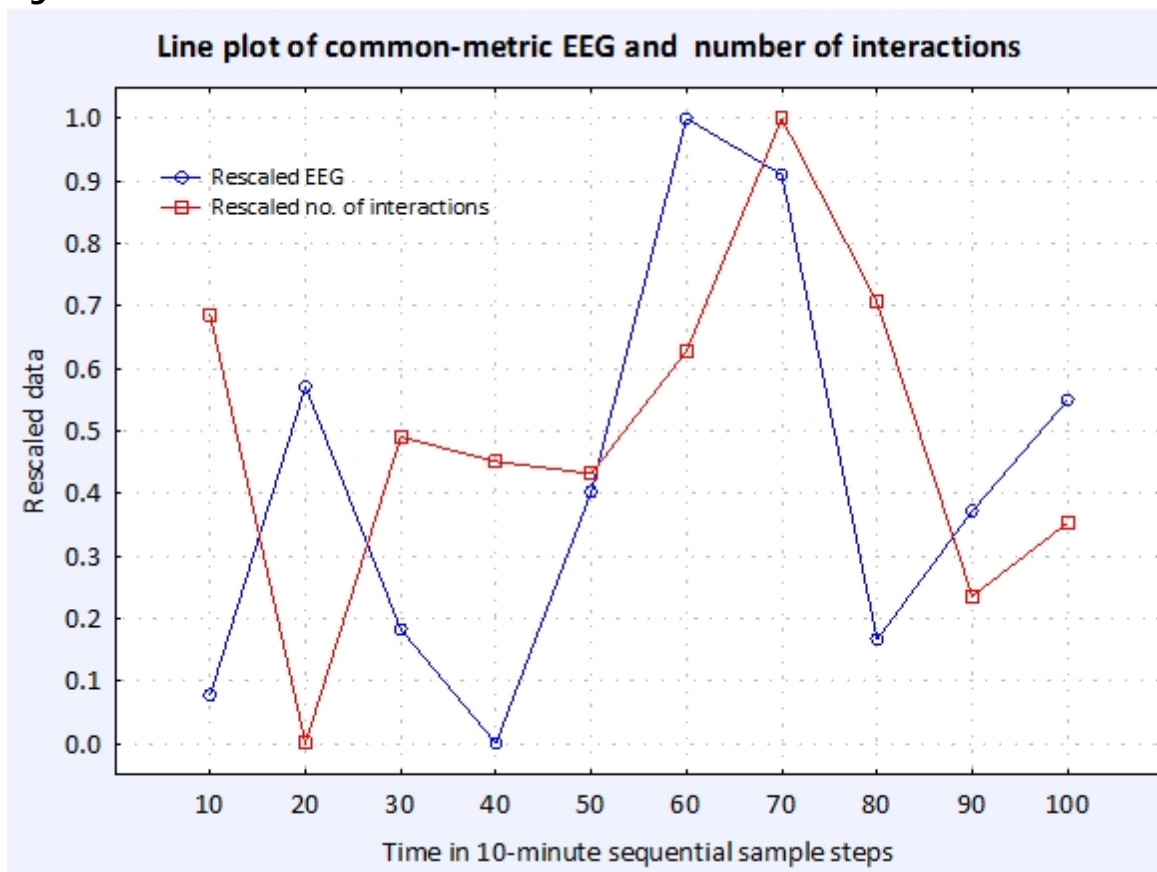
$$EEG_{rescale} = \left\{ \left[\frac{(2.56 - 2.09)}{(8.22 - 2.09)} \right] * (1 - 0) \right\} + 0 = .0767$$

The result of the scaling (rounded to 2 decimal places) is shown in Table 2, and Figure 3:

Table 2

Data: Test Spreadsheet for rescaling.sta (10v by 10c)				
10 minute time points	Test spreadsheet for rescaling example			
	1 Average 8-12Hz EEG power in dB	2 Number of interactions	3 Rescaled EEG	4 Rescaled no. of interactions
10	2.56	47	0.08	0.69
20	5.59	12	0.57	0.00
30	3.21	37	0.18	0.49
40	2.09	35	0.00	0.45
50	4.56	34	0.40	0.43
60	8.22	44	1.00	0.63
70	7.67	63	0.91	1.00
80	3.11	48	0.17	0.71
90	4.37	24	0.37	0.24
100	5.46	30	0.55	0.35

Figure 3

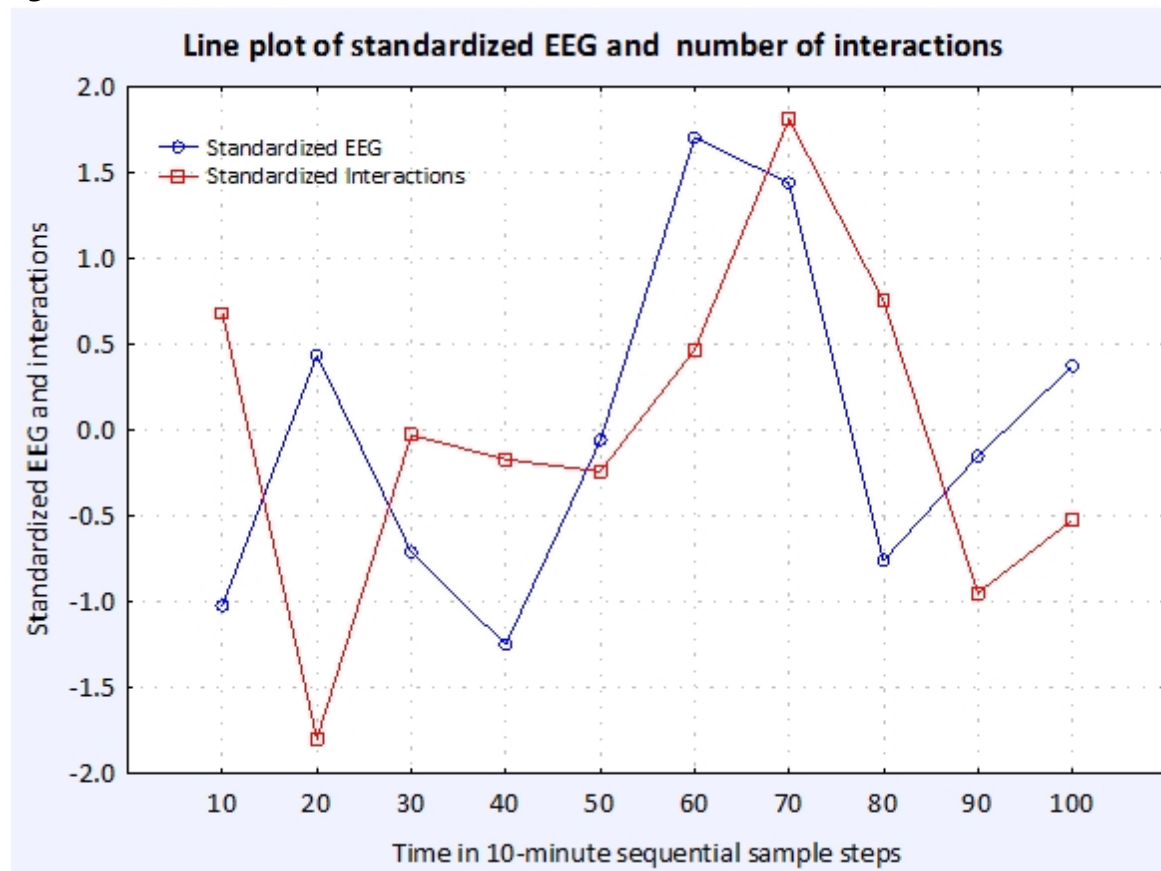


Now the variation in both variables is now expressed in a common linear metric, allowing direct visual comparison of lines, trends etc.

If, instead of linearly rescaling the data, we had standardized them instead (by subtracting each observed value from the mean of all observed values for a variable, and dividing this difference by the standard deviation of the values), the result would be that shown in the last two columns of Table 3, and Figure 4.

Table 3

10 minute time points	Test spreadsheet for rescaling example					
	1 Average 8-12Hz EEG power in dB	2 Number of interactions	3 Rescaled EEG	4 Rescaled no. of interactions	5 Standardized EEG	6 Standardized Interactions
10	2.56	47	0.08	0.69	-1.02	0.68
20	5.59	12	0.57	0.00	0.44	-1.80
30	3.21	37	0.18	0.49	-0.71	-0.03
40	2.09	35	0.00	0.45	-1.25	-0.17
50	4.56	34	0.40	0.43	-0.06	-0.24
60	8.22	44	1.00	0.63	1.70	0.47
70	7.67	63	0.91	1.00	1.44	1.81
80	3.11	48	0.17	0.71	-0.76	0.75
90	4.37	24	0.37	0.24	-0.15	-0.95
100	5.46	30	0.55	0.35	0.37	-0.52

Figure 4

If you look at Figure 3 and Figure 4 closely - you can see there is a slight perturbation in the magnitudes caused by the standardization transformation (because the divisor of each of the observation minus mean difference, the standard deviation, is not the same for both variables) - which is what causes the observed difference in standardized value metric.

Paul Barrett, (paul@pbarrett.net)

Part

IX

9 Relative v Simple Discrepancy

The key to fully understanding the [Gower](#) and [DSEs](#) logic is to understand the difference between relative and simple discrepancy.

Simple discrepancy is literally the subtraction of one value from another. It can be expressed as a signed or unsigned (absolute) difference.

Relative discrepancy expresses a simple discrepancy relative to the maximum possible discrepancy which could ever be observed between the two attributes whose magnitudes are being compared.

The problem with simple discrepancy is that, by itself, you have no idea whether the discrepancy is trivial or substantive. It is not until you compare it to what you could have observed given the measurement range of each variable /attribute that you can make such a judgment.

For example, a set of simple signed, unsigned (absolute), and relative discrepancies between two attribute magnitudes might be:

	vector scores and differences					
	1 vector 1	2 vector 2	3 Unsigned difference	4 absolute difference	5 Range 0-7	6 Range 0-20
1	6	7	-1	1	0.14	0.05
2	4	7	-3	3	0.43	0.15
3	3	1	2	2	0.29	0.1
4	5	4	1	1	0.14	0.05
5	7	5	2	2	0.29	0.1
6	2	3	-1	1	0.14	0.05

Consider the unsigned or signed differences, without knowing what the maximum possible discrepancy could have been, these might be trivial or substantive differences. And you'll note, it's pointless using unsigned differences in any averaging calculation, as the sum of these is not useful as an indicator of overall discrepancy (in the table above the sum of the signed discrepancies is zero).

But, when we express the absolute discrepancies relative to a maximum possible discrepancy (the range) of 7, and compute the average of these, we obtain a value of 0.24 (an average discrepancy of 24% of the maximum possible discrepancy between observations of 7). To be precise, $0.238095238095238 * 7 = 1.6667$ average error over a 7-point range.

But, if the maximum possible discrepancy was in fact 20, then the average relative discrepancy is now just 8% of the total possible discrepancy of 20). To be precise, $0.0833333333333333 * 20 = 1.6667$ average error over a 10-point range.

The magnitude of error (1.6667) is the same, but 1.6777 error within a 7-point range is meaningfully larger than a 1.6667 error within a 20-point range.

The relative discrepancy calculation not only achieves a common-metric for discrepancy (always between 0 and 1) regardless of the metric of the observations, but, it allows an investigator to make judgments about the trivial or substantive size of the discrepancies, bearing in mind just how discrepant they could have been.

