

Example: Rater 1 and Rater 2 evaluate 500 sampling units and code each into one of three classifications: Academic, Emotional, or Physical.

#		Rater 1		
		<b>Academic</b>	<b>Emotional</b>	<b>Physical</b>
Rater 2	<b>Academic</b>	210	50	25
	<b>Emotional</b>	35	125	15
	<b>Physical</b>	5	10	25

%		Rater 1		
		<b>Academic</b>	<b>Emotional</b>	<b>Physical</b>
Rater 2	<b>Academic</b>	0.42	0.10	0.05
	<b>Emotional</b>	0.07	0.25	0.03
	<b>Physical</b>	0.01	0.02	0.05

% w/marginals		Rater 1			
		<b>Academic</b>	<b>Emotional</b>	<b>Physical</b>	
Rater 2	<b>Academic</b>	0.42	0.10	0.05	<i>0.57</i>
	<b>Emotional</b>	0.07	0.25	0.03	<i>0.35</i>
	<b>Physical</b>	0.01	0.02	0.05	<i>0.08</i>
		<i>0.50</i>	<i>0.37</i>	<i>0.13</i>	

Observed Agreement (OA) = 0.42 + 0.25 + 0.05 = **0.72**

Expected Agreement (EA)

$$\begin{aligned}
 EA &= (0.50 \cdot 0.57) + (0.37 \cdot 0.35) + (0.13 \cdot 0.08) \\
 &= 0.285 + 0.1295 + 0.0104 = 0.4249 = \mathbf{0.42}
 \end{aligned}$$

$\kappa = OA - EA / 1 - EA$

$$\begin{aligned}
 &= (0.72 - 0.42) / (1 - 0.42) \\
 &= 0.3 / 0.58 = 0.517 = \mathbf{0.52}
 \end{aligned}$$