

**Table S1.** Interobserver reliability of four evaluators for indication of RA, unidimensional scales (VAS, NRS, and SDS), and composite pain scales in dogs (CMPS-SF), obtained by analyzing 208 videos from four perioperative moments (M1, M2, M3, and M4) of dogs undergoing orthopedic (n = 20) or soft tissue surgeries (n = 22) and negative control group (n = 10), between two phases, 21 days apart.

Statistical test	Variables	PhD student x Ethologist	PhD student x MSc student	PhD student x Student	Ethologist x MSc student	Ethologist x Student	MSc student x Student
Kw (CI95%)	Analgesic rescue	<b>0.6</b> <b>(0.52-0.68)</b>	<b>0.53</b> <b>(0.45-0.61)</b>	<b>0.51</b> <b>(0.42-0.6)</b>	0.42 (0.34-0.51)	<b>0.54</b> <b>(0.46-0.62)</b>	0.4 (0.31-0.5)
	Numerical rating scale	0.45 (0.38-0.52)	<b>0.6</b> <b>(0.54-0.66)</b>	0.41 (0.34-0.49)	0.29 (0.22-0.36)	<b>0.68</b> <b>(0.52-0.67)</b>	0.29 (0.22-0.36)
	Simple descriptive scale	<b>0.52</b> <b>(0.43-0.6)</b>	<b>0.54</b> <b>(0.47-0.61)</b>	<b>0.53</b> <b>(0.45-0.61)</b>	0.33 (0.24-0.41)	<b>0.6</b> <b>(0.52-0.67)</b>	0.39 (0.31-0.46)
ICC (CI95%)	Visual analogue scale	<b>0.71</b> <b>(0.65-0.76)</b>	<b>0.77</b> <b>(0.72-0.81)</b>	<b>0.7</b> <b>(0.58-0.72)</b>	0.36 (0.3-0.45)	<b>0.75</b> <b>(0.69-0.79)</b>	<b>0.64</b> <b>(0.56-0.7)</b>
Kw (CI95%)	A(i) – Spontaneous behavior	<b>0.58</b> <b>(0.46-0.69)</b>	<b>0.53</b> <b>(0.41-0.63)</b>	<b>0.52</b> <b>(0.4-0.63)</b>	<b>0.52</b> <b>(0.43-0.59)</b>	<b>0.56</b> <b>(0.43-0.68)</b>	<b>0.56</b> <b>(0.47-0.64)</b>
	A(ii) – Attention to wound	<b>0.59</b> <b>(0.4-0.74)</b>	0.37 (0.21-0.55)	0.32 (0.11-0.53)	0.46 (0.28-0.6)	0.56 (0.27-0.77)	0.3 (0.13-0.47)
	B (iii) – Mobility	<b>0.6</b> <b>(0.5-0.68)</b>	<b>0.62</b> <b>(0.51-0.71)</b>	0.46 (0.34-0.57)	<b>0.52</b> <b>(0.4-0.63)</b>	<b>0.71</b> <b>(0.58-0.81)</b>	0.46 (0.33-0.59)
	C (iv) – Response to touch	<b>0.65</b> <b>(0.55-0.73)</b>	<b>0.77</b> <b>(0.69-0.83)</b>	<b>0.76</b> <b>(0.69-0.84)</b>	<b>0.72</b> <b>(0.63-0.79)</b>	<b>0.74</b> <b>(0.66-0.82)</b>	<b>0.83</b> <b>(0.76-0.88)</b>
	D (v) – Demeanor	0.28 (0.2-0.36)	0.48 (0.39-0.56)	<b>0.54</b> <b>(0.47-0.6)</b>	0.34 (0.26-0.43)	0.27 (0.19-0.35)	<b>0.52</b> <b>(0.43-0.6)</b>
	D (vi) – Posture	0.41 (0.32-0.49)	0.37 (0.29-0.46)	0.29 (0.19-0.4)	0.31 (0.21-0.4)	0.39 (0.28-0.49)	0.33 (0.22-0.43)
ICC (CI95%)	<b>Total CMPS-SF</b>	<b>0.8</b> <b>(0.76-0.83)</b>	<b>0.83</b> <b>(0.8-0.86)</b>	<b>0.8</b> <b>(0.77-0.85)</b>	<b>0.78</b> <b>(0.73-0.81)</b>	<b>0.82</b> <b>(0.78-0.85)</b>	<b>0.83</b> <b>(0.79-0.86)</b>

Interpretation: 0.81-1.0 very good; 0.61-0.80 good; 0.41-0.60 moderate; 0.21-0.40 reasonable; <0.20 poor [35]. CMPS-SF is Short-Form Composite Pain Score [10], Kw is Weighted Kappa Coefficient, ICC is Intraclass Correlation Coefficient, [CI95% (Min-Max): 95% confidence interval (minimum-maximum values)]. In bold are the values considered acceptable by the adapted GRADE criteria [14,15].