

Introduction/motivation/goal clearly explained (10)		Appropriate statistical and visualization tools are used (20)	
FALSE	Is the motivation of the study stated?	FALSE	Are preliminary comparisons presented, e.g. as box plots with standard deviations?
FALSE	Is the goal of the study stated?	FALSE	Is an argument or proof for normal distribution of values in each category presented?
FALSE	Is it stated in a prominent location?	FALSE	If ANOVA was used, are the variations due to effects and the variations due to error reported?
FALSE	Are hypotheses stated?	FALSE	If ANOVA or Chi-square were used, are degrees of freedom reported?
FALSE	Are the algorithms introduced?		
FALSE	Are the used tools introduced?		
Setup and parameter selection clearly explained (10)		Conclusions are justified and sound (20)	
FALSE	Are parameters presented initially?	FALSE	Are there any conclusions?
FALSE	Are the reasons for parameter choices obvious or explained?	FALSE	Are the conclusions sound?
FALSE	Are the used metrics introduced?	FALSE	Are the conclusions based on the simulation, taken from elsewhere, or only speculation?
FALSE	Are categories and their properties introduced?	FALSE	Are there reflections of the own work in the conclusion?
FALSE	Is the discussion of categorical, continuous, interval and ratio terms for categories	FALSE	Are there future work questions?
Results are well explained (40)		SCORING	
FALSE	Are preliminary comparisons correctly presented without overstatement?	0	Terrible. Nothing in this characteristic was answered sensibly.
FALSE	Are their labels and legends complete?	1	Bad: Nearly nothing achieved. Apparently nothing was learned from the lecture or presentation round.
FALSE	Are figures readable?	2	Not good: Very few boxes are checked.
FALSE	Are all figures well-explained?	3	Satisfactory. There are mistakes, some of the boxes are not checked.
FALSE	Have Friedman's Chi-square or ANOVA been applied to ≥ 1 question?	4	Very good. No mistake have been made, all the boxes are checked.
FALSE	Have Friedman's Chi-square or ANOVA been applied in a context with >2 categories.	5	Extraordinary. All boxes are ticked, either perfectly solved or nice/innovative ideas are added.
FALSE	Was the choice of method (Chi-square vs ANOVA) explained well?		
FALSE	Has a correct Post-hoc analysis (e.g. Tukey's HSD) been performed?		
FALSE	Are the p values interpreted correctly?		
FALSE	Are the confidence levels interpreted correctly?		