A researcher was interested in examining how people conform to norms concerning littering. In particular, the researcher wanted to determine whether the tendency of secondary school students to litter depended on the amount of litter already in the area. Students were handed a pamphlet as they entered a room that already had either 0, 4 or 16 pamphlets lying on the ground. Students were then observed to determine whether or not they dropped their pamphlet on the ground. Using the results below, write a report on whether the researcher's hypothesis was supported.

Littering * Amount of litter Crosstabulation

| | | | Amount of litter | | | |
			none	some (4 pieces)	A lot (16 pieces)	Total
Littering	did not drop	Count	102	102	69	273
		% within Amount of litter	85.7%	85.0%	57.5%	76.0%
	dropped pamphlet	Count	17	18	51	86
		% within Amount of litter	14.3%	15.0%	42.5%	24.0%
Total		Count	119	120	120	359
		% within Amount of litter	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	34.045[a]	2	.000
Likelihood Ratio	32.603	2	.000
Linear-by-Linear Association	26.106	1	.000
N of Valid Cases	359		

The following student responses were obtained: *Critique them with your group*

Student 1

What research has found was that overall students were less likely to drop the pamphlet, with 76% not dropping at all under the three circumstances. (Chi-square Equation $\chi^2=(2) = 34$, $p < .001$).

Student 2

Generally there was very little difference between the adults who littered and the amount of litter that was around.

Student 3

The results produced in the crosstabulation only support the hypothesis slightly that the tendency of secondary students to litter depended on the amount of litter already in the area. There was little distribution between the litter dropped in a room with no pamphlets already on the ground and the room with 4 pamphlets already on the ground, reporting a column percentage difference of .07% (one piece of litter). However, whilst 14.3% dropped litter in the room with no pamphlets, 42.5% dropped litter in the room with 16 pamphlets. Using the chi squared test, we can see there is a significant difference ($\chi^2(2) = 34.05$, $p < .01$) between the amount of litter dropped depending on the amount of litter already in the area.

Student 4

The researcher's hypothesis was supported. Dependant upon the amount of litter was around the school children indicated their lack to place their pamphlets in the bin. This is indicated by the crosstabulation table. Correlation is significant as 0.000 (2-tailed). Pearson's Chi-Square value at 34.045.

Clearly this student has no idea what the analysis is about. They are not looking at the relationship between amount of litter present and littering behaviour, and are just blindly reporting the chi-square statistic without any notion of what it conveys.

Not only does this response show some basic problems in the student's understanding, it is also another example of a student writing something which they would not write in a non-statistical setting.

Even when students understand a particular analysis, they can still find it difficult to write a cohesive report. The response from Student 3 seems to show some understanding of the analysis, but an inability to express it clearly.

Part of the difficulty in writing reports relates to poor English expression skills, as the response from student 4 amply demonstrates (and yes, English is their first language)

Table 2 *Process for Exploring the Relationship between Two Categorical Variables*

Process in General Terms	*Process in Practice for Littering Example*
What are the variables involved in this study?	'amount of litter present' (none, 4 pieces or 16 pieces) and 'littering behaviour' (drop pamphlet or do not drop pamphlet).
How are they measured (categorical or metric)?	Both variables are categorical.
Is there an explicit hypothesis here – based on previous research? If so, what is it – for two categorical variables, phrase the hypothesis in terms of which group is more (or less) likely to - - -	There is no explicit hypothesis in the question.
If there is no explicit hypothesis – say what you expect to find. This could be a wild guess – use the words more (or less) likely to - - -	We might guess that people are more likely to drop the pamphlet if there is a lot of litter present than if there is no litter present.
Who are you comparing? This is the IV, place in columns and request column %	We are comparing participants in the 'no litter' environment to participants in the 'lots of litter' environment, so amount of litter is the independent variable. This is placed in the columns of the crosstabulation and column percentages calculated.
In your initial sentence of expectations which level of the dependent variable were you focussing on? Look at all of the percentages in this row – which group is actually more or less likely to?	In our hypothesis we focussed on how likely people were to drop the pamphlet, so this is the row of the table we'll focus on. Looking at the percentages in this row, we can see that the percentages look pretty much as we expected – people were more likely to drop the pamphlet when there were 16 pieces of litter present than when there was no litter present.
Give an overview statement that comments on this (don't use any percentages, just say more or less likely)	"In our sample of 359 school students, people were more likely to drop their pamphlet when the room had a lot of litter on the floor than when there was no litter on the floor."
Give the percentages to back this up.	"While 42.5% of students dropped their pamphlet when there was a lot of litter on the floor, only 14.3% of students dropped their pamphlet when there was no litter on the floor and 15.0% of students dropped their pamphlet when their was some litter on the floor.
Is there a consistent pattern in other rows? No – then describe other rows as well	Because there are only two categories of the dependent variable here, a consistent pattern emerges for both rows of the table, so there is no need to mention the percentages from the 'non-littering' row – they add nothing to the story.
Use chi-square to comment on the significance of the relationship.	"The relationship between amount of litter and littering behaviour is significant ($\chi^2(2) = 34.05$, $p < .001$)."