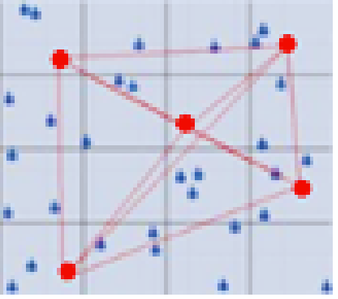

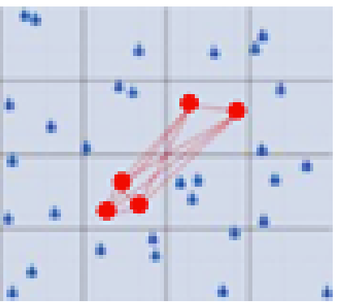

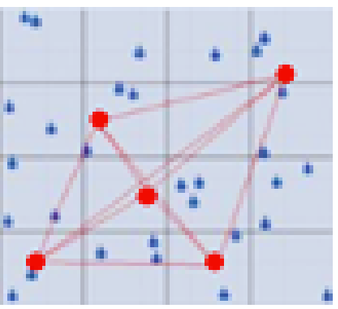

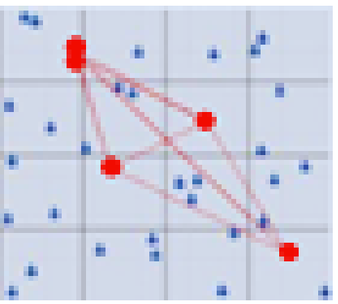

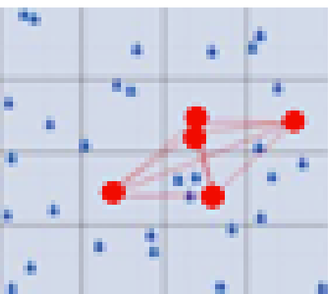

RANDOM RAINFALL

AMY RENELLE
DEPARTMENT OF STATISTICS
SUPERVISORS: STEPHANIE BUDGETT & RHYS JONES

Can we tell the difference between spatial randomness plots that were created by human participants and randomly generated coordinates?
You be the judge! Which plots on the left do you think were made by human participants?

SURVEY QUESTION

We asked participants the following question in an online, anonymous questionnaire:

You are watching the rain on one of these days as it falls on the pave stones outside. There is nothing blocking the pavement and the rain is falling straight down. The last 30 raindrops were recorded below (blue spots)**. On the grid below, indicate where you think the next five raindrops might fall** (red spots)**.**

WHY?

Spatial randomness is one way we can explore participants' perceptions of randomness; in this case, relating independent events to distributions. Engel and Sedlmeier (2005), using a similar snowflake example, explored middle school students' perceptions of "...the unforeseeable behaviour of single flakes on the one hand and the well-experienced fact that on the other hand after some time of snowfall the snow is approximately on even height, for example, on a flat rooftop" (p. 169).
Randomness is a fundamental concept across numerous fields but is often riddled with misconceptions (Tversky & Kahneman, 1974) - my research explores the presence of this within our secondary school teachers' community!

METHODOLOGY

A set of randomly generated plots, along with the participant responses, had their labels removed, their snowflakes connected (to better see the relationship between the points), and then they were shuffled together. Five people then arranged these into groups according to whether they felt the plot was a participant response or not according to some characteristic. We can then test whether the allocation by the judges could've been just by guessing. We can also see if some participant responses were commonly selected by the judges.

It is hypothesised that, if the judges could not tell between the participant responses and randomly generated plots, the proportion correct would be 50%. So, Null hypothesis: $p = 0.5$ and Alternative hypothesis: $p \neq 0.5$.

FINDINGS

The characteristics used by three of the judges led to correctly selecting more human responses than under chance alone, with a p-value of 0.0177 (24/35 identified). Another set of characteristics were on the border with a p-value of 0.0504 (23/35), while the last set of characteristics were less successful and not significantly different from the null hypothesis, with a p-value of 0.3961 (20/35). So, for the most part, there is something identifiable about the human-generated plots and this could suggest a poor sense of spatial randomness.

EXAMPLE CONCLUSION

Did you spot the participant-generated plots?
While the study collected 35 human-produced plots and 35 randomly generated plots, only a small sample of five was shown here.
And, in fact, all five are examples of participants' plots! All five had been correctly selected by the judges as being human-derived - maybe there is something about the location participants placed their snowflakes that seems non-random? Notice the empty spaces that were filled by the participants? Or the cluster of three original blue snowflakes that participants gravitated towards? Did you spot the centrality of the placement of points? Or the relationship between the coordinates?

WHERE WOULD YOU PLACE YOUR SNOWFLAKES? THINK YOU COULD TRICK MY JUDGES?

Engel, J., & Sedlmeier, P. (2005). On middle-school students' comprehension of randomness and chance variability in data. Zentralblatt Für Didaktik Der Mathematik, 37(3), 168–177. https://doi.org/10.1007/s11858-005-0006-4
Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124