



Plan For Study

Project Title

Molecular Structure Drawing Prediction

Researcher Name

Jacob Alexander Cameron Rivett

Supervisor Name

Simon Dobson

Participants

We will be looking for Chemistry students or Lectures that are familiar with chemical structure editors. Each participant should be comfortable with the universal operations of a chemical structure editor. This means students in 2nd year or above mostly but competent 1st years could be considered also.

Process

Each participant will be given two lists of molecular structures to draw using a web-based chemical structure editor with the prediction tool and also without the prediction tool.

One list will be larger than the other and contain Lewis structures and another smaller list of simple condensed formulas. Each participant will first draw the list of structures without the tool, this will fill the database. Once all participants have done the first step the database will be frozen in that state. Then The participants will be asked to draw the structures again with the tool.

They will each be assigned a unique user Id and a group Id. The user Id will be unique and the group Ids will be equally spread out so that there is no benefit to being in one group than another.

Measurements

We will be measuring the time it takes a participant to complete a structural drawing. When the prediction tool is introduced we will then measure the number of times the participant uses the tool. We are looking for a pattern in the data connected to the time it takes to complete the drawing of a structure and the number of times the participant uses the prediction tool. We are also looking for a direct difference in speed with the prediction tool and without it.

The number of undos/errors made by the participant while drawing will also be recorded. The reason for measuring errors is to see if there is any difference between the participant making mistakes when they have the tool and without the tool.

The reason for two different list is that a condensed formula means the participant has to maybe think longer than when drawing a lewis structure because they are not blindly copying. The lists of data for each user will be similar if they have the same group id. This is because in the real world, a research group is far more likely to be looking at similar structures. If the data is sourced from the NOMAD system, which is the initial plan, this will become automatic.

What is the Hypothesis?

The hypothesis is that when the prediction tool is introduced there will be a significant reduction in time and number of errors made compared to when it is not there. There should be a direct relationship between the time and the number of times the participant clicks onto one of the predictions.