CPSC 410 - Assignment 3

Exercise 1

1.
   - Type of study: Survey
   - Threats: The first threat belongs to `internal validity`. Developers may have biases towards the IDE, so we may not say the result is solely based on our manipulations. The second threat belongs to `external validity`. There are many kinds of developers, so our sample is too small and not repersentative. The third threat belongs to `construct validity` as the study would like to measure the real productivity instead how the programmers feel about their productivity.
2.
   - Type of study: Historical Data Collection and Analysis
   - Threats: The first threat belongs to `internal validity`. Although there is strong correlation between number of bugs and number of commits showed in our analysis, we are unable to conclued is as a causal relationshiop. The second threat belongs to `external validity`. We may not find this correlation in other projects, so our findings can not be generalized.
3.
   - Type of study: Observational / Exploratory Studies
   - Threats: The threat belongs to `construct validity`. The DSL language's aim users are kids. The concepts and effictiveness of the DSL could not be correctly measured if we ask developers to use the DSL.

Exercise 2

1. The potential risk here is that my 410 classmates are willing to be recorded in the video and being showed to other audience. Also, the intended user groups are kids, so the responses from the classmates may not truly evaluate the DSL. Participants would not truly voluntarily participate. I am their friends may affect their decisions even if they do not want to participate.
2. The potential risk here is that students who participate the experiment may acadmically misconduct in the exam. They have been promised by the professor that thet will not get penalized if they cheat. They may not truly volunarily participate. They may participate the experiment in order to get advantage in the exam.
3. The potential risk here is that running software background may violate the privacy of the employees. Also, if their real productivities are lower than their real ratings, they may feel bad about themselves. THe employees may not participate voluntarily as I am their employer and they do not want to annoy me.

Exercise 3

1. The author talked about two options in the introduction. The first option is maximizing internal validity which gives full explanation but is at cost of generlizability. The second option is maximizing external validity which could be applied to a variety of cirumstances, but can't fully explain the underlying factors and relationships.

2. I think the section 7 , Further Insights, contains the most important take-aways for me as a reader. With the rising awareness and usage of empirical studies, as a software developer, it is important to learn from experts how they thought the emprical studies would develop in the field of software engineering.

3. If we want to maximize internal validity, we can only have one chaning factor. The result would give sufficient explanation to the underlying relationships but is at cost of generality. if we want to maximize externality validity, then we would have many chaning factors. Therefore, we could not maximaize them at the same time.

4. In my opinion, the most interesting and suprising part is that removing humans from experiments improve the internal validity. The researchers state that human would lie or act differently in differnt contexts. Therefore, I think we may not apply same criteria for human and non-hunab studies.