

Participants' perception of tasks in an informatics contest

Appendix to the paper

1 Research hypotheses

To answer the given research questions, we formulated several research hypotheses:

- Based on RQ1 we developed the hypothesis H1: There is no relationship between average perceived difficulty of a task and the proportion of participants who did not answer it.
- Based on RQ1b we developed the hypothesis H2: There is no relationship between average perceived difficulty of a task and the proportion of participants who got it right.
- Based on RQ1c we developed the hypothesis H3: There is no relationship between average perceived difficulty of a task and the proportion of participants who got it wrong.
- Based on RQ2 we developed the hypothesis H4: There is no relationship between perceived difficulty of the test and participants' test score.
- Based on RQ3 we developed the hypothesis H5: There is no relationship between contestants' self-perceptions of IT ability and their achievement in the test.
- Based on RQ4 we developed the hypothesis H6: There is no relationship between participants' self-perception of IT ability and their perception of tasks difficulty.
- Based on RQ5 we developed the hypothesis H7: There is no relationship between gender and self-perception of IT ability.
- Based on RQ6 we developed the hypothesis H8: There is no relationship between gender and accuracy of self-perception of IT ability.

2 Variables

To verify the given hypotheses, we made following variables:

- *Average perceived difficulty of a particular task* – arithmetic mean of difficulty of a task perceived by participants
- *Perceived difficulty of the test* – arithmetic mean of difficulty of tasks in the test perceived by a participant
- *Self-perception* – participant's self-perception of IT ability
- *Gender* – participant's gender
- *Points scored* – number of points scored by a participant
- *Proportion of no answers* – proportion of contestants who did not answer a task
- *Proportion of correct answers* – proportion of contestants who answered a task correctly
- *Proportion of incorrect answers* – proportion of contestants who answered a task incorrectly

3 Sets of data used and data set cleaning

To answer the research questions, we used data samples that would enable a particular sample to be as large as possible whilst ensuring each item included all the required information. Table 1 shows which samples of participants we used to validate each particular research question.

Table 1. Samples of data used to validate the hypotheses. The X-value denotes the first variable used in the hypothesis, the Y-value denotes the second variable used in a particular hypothesis.

	All participants	Participants who responded to the questionnaire	Participants who responded to the questionnaire and agreed to have their responses correlated with their test score
RQ1a-RQ1c	Y	X	
RQ2			X, Y
RQ3			X, Y
RQ4		X, Y	
RQ5		X, Y	
RQ6			X, Y

Some items acquired from the questionnaire survey were unsuitable in certain respects for further analysis as they did not include all the required information. We did not completely eliminate these items from the study but only from a particular research question. The following is a list of the items which we eliminated:

- RQ1a-RQ1c: Those who did not rate the difficulty of any of the contest tasks or rated all the tasks the same, making their rating unreliable, were eliminated from the X variable (leaving 381 records)
- RQ2: Those who did not rate the difficulty of any of the contest tasks were eliminated from the X and Y variables (leaving 243 records)
- RQ3: Those who did not rate self-perception of their IT ability were eliminated from the X and Y variables (leaving 280 records)
- RQ4: Those who did not rate the difficulty of any of the contest tasks or rated all the tasks the same, making their rating unreliable, were eliminated from the X and Y variables (leaving 381 records)
- RQ5: Those who did not give their gender or did not rate self-perception of their IT ability were eliminated from the X and Y variables (leaving 486 records)
- RQ6: Those who did not give their gender were eliminated from the X and Y variables (leaving 280 records)