





# Participants' Perception of Tasks in an Informatics Contest

Jiří Vaníček<sup>(✉)</sup>  and Václav Šimandl 

University of South Bohemia, České Budějovice, Czech Republic  
{vanicek, simandl}@pf.jcu.cz

**Abstract.** Bebras is a contest organized for schools to promote computational thinking and new tasks are developed for it every year. The contest is held online in over 50 countries around the world and its results provide a great deal of information regarding difficulty levels of particular tasks and factors influencing participants' success in the contest. Our paper brings an analysis of such hard data in relation to statements of those participants who expressed their opinion on the difficulty of particular tasks and whose statements could be matched with their actual academic performance in the contest. Questionnaires were voluntarily completed by contestants participating in the Senior category (for the oldest participants aged over 16) from the Czech Republic.

Statistical analysis of contest results and perceived task difficulty provided new findings. There is no relationship between the proportion of participants with wrong answers to tasks and participants' subjective perception of task difficulty. Subjective difficulty is more aptly expressed in terms of the proportion of participants who did not answer the task. The contest was perceived as being easier by those participants who achieved higher scores in it. Male participants demonstrate a higher level of self-esteem in terms of IT skills than female participants. However, female participants' self-perceptions of their IT skills are slightly more accurate than males'. Results presented in the article could help the authors of the contest to improve compilation of contest tasks and have them included in the Informatics curriculum.

**Keywords:** Bebras · Informatics contest · Task difficulty · Participant perception · Self-esteem

## 1 Introduction

The situational informatics tasks used in the Bebras contest [1, 2] can be viewed from various angles. One view sees them as practical problems, developing various elements of computational thinking such as algorithmization, abstraction, decomposition and evaluation. From another angle, they can be seen as tasks involving various parts of computer science like programming, optimization, data representation, structures, processes, hardware, coding, cryptography, robotics and social aspects [3–6].

Bebras tasks are examined in a large number of studies. Some of them deal with creating good informatics tasks [7–9] or criteria influencing task difficulty [10–13]. Other

papers deal with the use of Bebras tasks - in assessing levels of computational thinking [14], bringing change into the classroom [6, 15] or preparing teachers as curriculum makers [16, 17]. School course books include Bebras tasks [18] or similar situational tasks that are adapted to a particular method of teaching [19].

A considerable proportion of research focuses on the difficulty of particular tasks or their types [13, 20–22]. As stated by Lonati et al. [13], predicting task difficulty is far from being an exact science and understanding difficulties and mistakes afterwards is not that easy either. Bellettini et al. [22] point out that in one third of the cases the tasks were either easier or more difficult than expected. According to van der Vegt and Schrijvers [20], category data structures and representations seem to be easier than algorithmization and programming tasks, while combining these categories increases the difficulty even further. Moreover, open ended tasks turned out to be the hardest as opposed to interactive or multiple choice tasks [20].

Other studies explore the issue of age suitability [22] and gender balance of contest tasks [23–26]. Hubwieser et al. [24] and Izu et al. [26] inquire into motivating younger girls to solve informatics tasks. Stupurienė et al. [25] point out that girls from the 3rd to 8th grades are as interested in solving informatics problems as boys. However, the proportion of 11th and 12th grade girls in the contest drops. Izu et al. [26] state that a declining performance trend of girls vs. boys can be recognized from primary to high school level, with boys outperforming girls in all countries in the Senior category. Other studies focus on international and long-term comparisons of tasks [25, 27, 28].

Given the online nature of the contest, it is not difficult to acquire data that measures achievement. As long as personal data protection rules are complied [29], an online contest can provide a large amount of data on how participants coped with various tasks. Statistical analysis of such data can indeed enable us to ascertain which type of task is difficult.

However, such data cannot ascertain how participants perceive contest tasks, their difficulty, and the motivational effect or “discouraging effect” of task settings. Very few studies actually inquire into this issue. One exception is Lonati et al. [13], who used interviews to uncover difficulties that participants encountered while completing tasks. Our paper is focused on participants’ feedback, providing their retrospective self-perceptions of contest tasks and a comparison of that feedback with their actual results. The aim of our study is to shed light on participants’ perceptions of contest tasks. The following research questions (RQ) were formulated on the basis of the research aim:

- Is there a relationship between perceived difficulty of a task and the proportion of participants who did not answer (RQ1a), gave the correct answer (RQ1b) or gave the wrong answer (RQ1c)?
- Is there a relationship between perceived difficulty of the test as a whole and performance in it (RQ2)?
- Is there a relationship between respondents’ self-perceptions of IT ability and their performance in the test (RQ3) or their perception of tasks difficulty (RQ4)?
- Are self-perceptions of IT ability higher (RQ5) and more accurate (RQ6) for male or for female participants?

## 2 Methodology and Design

### 2.1 Bebras Contest and Questionnaire Survey

In the Bebras contest, participants take an online test, comprising of situational informatics tasks. In a situational task, solvers emerge into a described situation in which they must grasp, get to understand the used concepts and terms, find an informatics principle the task is based on, solve the problem using cognitive and thinking skills and select the right answer from the options [12]. The Senior category for students aged 16 to 19 consists of 15 tasks. These are completed by selecting the correct answer from a number of options or by typing a number or some text. In interactive tasks, objects are to be checked or moved with a mouse. Participants have 40 min to complete the test. For correct answers, participants are awarded a varying number of points according to predefined task difficulty levels. For incorrect answers, a proportional number of points are deducted. Points are not deducted for unanswered tasks (this rule has an impact on deciding whether to answer a task if a participant is not sure about the correct answer). As soon as the test is over, participants can see their points total but cannot see which tasks they got right and where they made mistakes. Such feedback could have an impact on their perception of the difficulty of the test as a whole.

In order to determine how participants perceive the difficulty of contest tasks, we compiled an anonymous online questionnaire. This comprised of questions concerning issues such as pupils' self-perceptions of their IT ability and their opinion on the difficulty of particular tasks in the contest. Where participants agreed, their questionnaire responses were correlated with their test score.

In the research, we used the term IT ability instead of informatics/computer science ability because a school subject known by research participants is called IT. This subject is based on digital technology handling and rarely contains topics from informatics, such as algorithmization. Since most Czech pupils might not properly understand the term informatics, asking a question about their informatics ability could bring unclear results.

The *self-perceptions of IT ability* and *perceived difficulty* of contest task variables express participants' subjective opinion in the six-point Likert scales, whose extreme ends were "I know nothing" and "I am very good at it", or "Easy" and "Hard". Participants also had the possibility of indicating that they are "unable to say".

### 2.2 Research Sample

All Senior category participants in the 2018 contest were notified with a request to complete a questionnaire. In that year 5898 participants took part in that category designated for pupils aged over 16 and our questionnaire was completed by 595 of them. 565 of the questionnaires were sent shortly after completing the test at a time when participants had already been informed of the scores they had achieved in the test but had not yet been given the correct answers to particular tasks and did not know whether their answers were right or wrong. The remaining 30 questionnaires were sent at a time when this information had been made available to them. 294 participants gave their consent to having their questionnaire answers correlated with their contest test scores (for details see Table 1).

**Table 1.** Numbers of participants and received questionnaires according to gender. Some respondents did not state their gender in the questionnaire.

	Men	Women	Total
Total number of participants	4093	1805	5898
Number of questionnaires received	385	131	595
Number of participants agreeing to have their responses correlated with their test scores	220	74	294

### 2.3 Data Analysis

To answer the above mentioned research questions, we developed research hypotheses, as presented in the Appendix to the paper available at <https://www.ibobr.cz/papers/ISSEP2020.pdf>. Data analysis was carried out using the following statistical methods: Pearson's correlation coefficient, Spearman's rank correlation coefficient and Mann–Whitney U test.

Pearson's correlation coefficient was used for linear expression of the relationship between two variables as per Cohen [30]. According to Chráska [31], a test value of 0 indicates a statistical independence between the two variables, while a value of +1 (or –1) indicates a perfect correlation between the two variables.

As per King and Eckersley [32], we could not use Pearson's correlation coefficient for data that are not discrete or continuous and are not normally distributed. For such cases we used Spearman's rank correlation coefficient, which is used to calculate a measure of correlation and works on the ranked values of the data.

As per Chráska [31], we used the nonparametric Mann–Whitney U test to verify whether two samples can come from the same basic set. This test is used to analyze two-sample unpaired data [33] and is based on pooling all original sample values and then ranking them. The following null hypothesis is used: The two populations have identical distributions. The following alternative hypothesis is stated: The two populations have different medians, but otherwise are identical.

Some items acquired from the questionnaire survey were unsuitable for analysis in certain respects as they did not include all the required information. Consequently, they were eliminated, as detailed in the Appendix.

The statistical software R was used for data analysis.

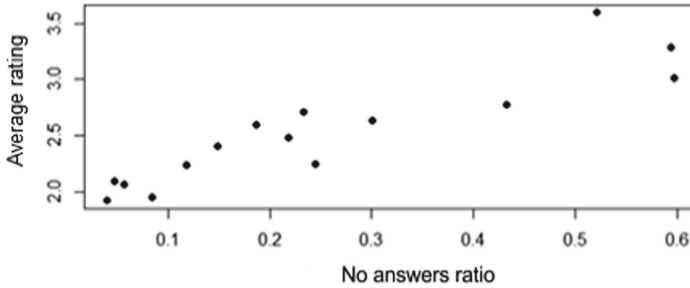
## 3 Results

### 3.1 Relationship Between Perceived Difficulty of a Task and Type of Answer Participants Gave

We used the variable *average perceived difficulty* of a particular task to answer research questions RQ1a to RQ1c. We also used the proportion of a certain type of answer (did not answer, gave the correct answer, gave the wrong answer) within the total number of participants. For the observed variables, the Anderson-Darling test and the Shapiro-Wilk

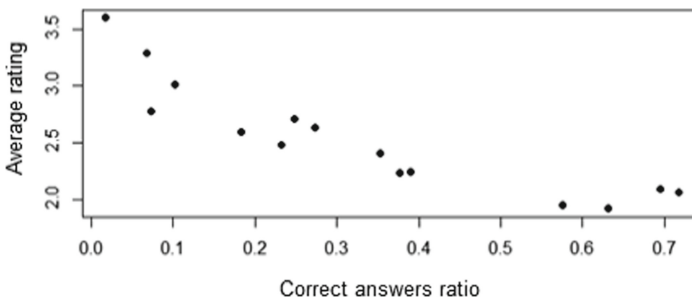
test for normality failed to reject the null hypothesis of normal sample distribution. Consequently, we treated that data as data from a normal distribution. Pearson's correlation coefficient was used for testing.

To answer RQ1a, we verified the hypothesis whether the *average perceived difficulty* and *proportion of no answers* variables are independent. Pearson's correlation coefficient is equal to  $R = 0.91$ . There is a relationship between the variables at a significance level of 0.05. It can therefore be said that the proportion of no answers does express participants' perceptions of their difficulty (Fig. 1).



**Fig. 1.** Relationship between *average perceived difficulty* of a task and *proportion of respondents* who did not answer

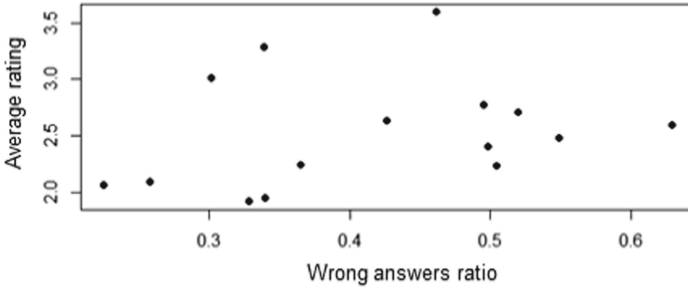
To answer RQ1b, we tested whether the *average perceived difficulty* and *proportion of correct answers* variables are independent. Pearson's coefficient is equal to  $R = -0.89$  and there is negative relationship between the variables at a significance level of 0.05. It can therefore be said that the proportion of correct answers also expresses participants' perceptions of the difficulty of contest tasks – tasks which are perceived by participants as more difficult have fewer correct answers (Fig. 2).



**Fig. 2.** Relationship between *average perceived difficulty* of a task and *proportion of correct answers*

To answer RQ1c, we tested whether the *average perceived difficulty* and *proportion of incorrect answers* variables are independent. Pearson's coefficient is equal to  $R = 0.27$ . We cannot reject the null hypothesis at a significance level of 0.05 so we cannot

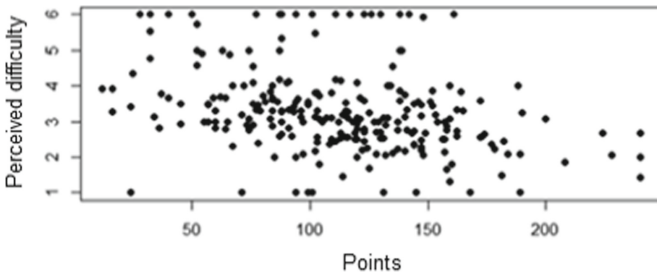
claim to have evidence of any relationship between these two variables. This can lead us to infer that the proportion of wrong answers does not express participants' perceptions of the difficulty of contest tasks (Fig. 3).



**Fig. 3.** Relationship between *average perceived difficulty* of a task and *proportion of wrong answers*

### 3.2 Relationship Between Perceived Difficulty of the Test as a Whole and Performance in It

In research question RQ2, we use the *points scored* and *perceived difficulty of the test* variables, for which all implemented tests (Shapiro-Wilk, Anderson-Darling, Kruskal-Wallis) rejected normality. We used Spearman's coefficient to calculate correlation, which is equal to  $\rho_s = -0.37$ . Its result at a significance level of 0.05 clearly rejects the null hypothesis of no relationship between the variables. There is indirect rank correlation between number of points scored in the test and perception of its difficulty. This can be interpreted to mean that the test was perceived as being more difficult by participants with lower scores in the test (Fig. 4).



**Fig. 4.** Comparison of *points scored* in the test and *perceived difficulty of the test*

### 3.3 Self-perceptions of IT Ability and Performance

The following two research questions concern respondents' self-perceptions of IT ability, their test score and perception of test difficulty:

- RQ3: Is there a relationship between respondents' self-perceptions of IT ability and their performance in the test?
- RQ4: Is there a relationship between respondents' self-perceptions of IT ability and their perception of tasks difficulty?

To answer these questions, we used the *self-perception* variable, which is an ordinal variable that cannot have normal distribution, even asymptotically. For that reason, we used Spearman's coefficient of rank correlation.

To answer RQ3, we used *points scored* as the second variable, Spearman's coefficient being equal to  $\rho_s = 0.26$ . The null hypothesis of no rank correlation between the variables was rejected at a significance level of 0.05. A certain weak correlation between self-perceptions of IT ability and number of points scored is identified.

To answer RQ4, we used the variable *perceived difficulty of the test*. Spearman's coefficient being equal to  $\rho_s = -0.27$ , identifying almost the same weak negative correlation as in the previous research question RQ3. The test again rejected the null hypothesis of no correlation, indicating a certain weak negative correlation between these variables.

To a certain extent, this confirms expectations that participants who perceive themselves as having above-average IT ability actually achieve higher scores in the test and perceive the test as being rather easy. However, this is not a particularly strong corroboration.

### 3.4 Gender Differences in Self-perceptions of IT Ability

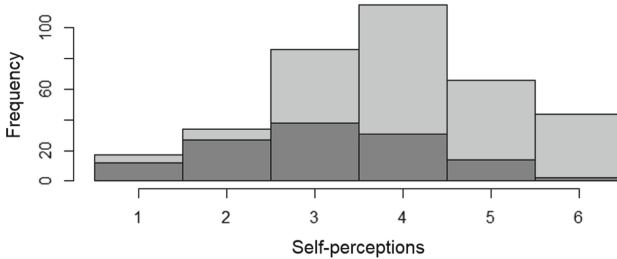
Participants provided data regarding their gender both in the contest and the questionnaire, the collected data enabling us to answer research questions concerning male and female participants' self-perceptions:

- RQ5: Do male or female participants have higher self-perceptions of IT ability?
- RQ6: Are self-perceptions of IT ability more accurate for male or for female participants?

To answer RQ5, we used the *gender* and *self-perception* variables. The *self-perception* variable being of an ordinal type, we cannot compare the expected value. We can only compare the median of *self-perception*, the male median equaling  $M_M = 3.86$  and the female median equaling  $M_F = 3.11$ . The result of the nonparametric Mann–Whitney U test rejected the null hypothesis of both genders having the same self-perceptions of IT ability at a significance level of 0.05.

To answer RQ6, we used a comparison of the *self-perception* and *points scored* variables. To be able to test which gender perceives their IT ability more accurately, we transformed the *points scored* continuous variable into an ordinal variable called *categorized points* by dividing the interval of 0–240 points into 6 equal intervals of 40 points. The median of the absolute value of the difference between *self-perception* and *categorized points* was calculated at:  $\rho_{sM} = 1.25$  for men,  $\rho_{sF} = 1.04$  for women. The Mann–Whitney U test does not reject the null hypothesis of both genders self-perceiving their IT ability with equal accuracy at a significance level of 0.05 but does reject it at a significance level of 0.1.

These results can be interpreted to mean that male participants have higher self-perceptions of IT ability than female participants. Contrarily, female participants' self-perceptions of their IT ability are slightly more accurate than male participants', the difference being negligible (Fig. 5).



**Fig. 5.** Stacked chart showing distribution of female (dark grey) and male (light grey) participants' self-perceptions of IT ability.

## 4 Discussion and Conclusion

By comparing participants' statements with their actual score, it was discovered that perceived difficulty of a particular task is not expressed by the proportion of participants who answered incorrectly. Perceived difficulty of a task is expressed by the proportion of those participants who did not answer and also of those who gave the correct answer. This result corresponds to findings of Vaníček [12] that the no answer indicator describes task difficulty very well. The perceived difficulty of contest tasks can be considered an important parameter, as per Keller and Landhäußer [34] a perceived fit of skills and task demands is a prerequisite for emergence of flow, which means the state in which according to Engeser and Schiepe-Tiska [35] an individual is completely immersed in an activity without self-consciousness but with a deep sense of control.

Our research also shows that participants who achieved higher scores in the test do actually perceive it as being easier. Moreover, participants who have higher self-perceptions of IT ability perceive the test as being easier and achieve higher scores in it. These conclusions correspond to findings of Li et al. [36], Mangos and Steele-Johnson [37] that self-perceptions of ability are negatively related to perceptions of task difficulty; self-perceptions of ability have a positive correlation with performance; and perceptions of task difficulty are negatively associated with performance.

In addition, our research identifies findings relating to gender differences. Male participants of this age were found to have higher self-perceptions of their IT ability than female participants but female participants' self-perceptions of their IT ability are slightly more accurate than male participants'. The finding that men have higher self-perceptions of IT ability than women corresponds to earlier research of Cussó-Calabuig et al. [38] and Birol et al. [39] relating to self-perception in IT. As Vekiri claims, perceived teacher expectations are more strongly associated with girls' than with boys' self-efficacy beliefs in IT [40], thus there seems to be a need to encourage and support girls to cope



with informatics problems. This is a particularly important task as self-confidence and perceived self-efficacy seem to play a big role for students' choice of further studies, as Dagienė et al. quote Ashcraft et al. [41].

The limit of our study is that some participants could have based their self-perceptions of IT ability on the score they had achieved in the test that they had just taken. In a certain way, time delay between the test and the questionnaire could have influenced perceived difficulty of contest tasks. While some responded to the questionnaire shortly after completing the test, other responses were delayed or sent after being provided with the correct answers, knowing whether their answers to particular tasks were right or wrong. As only 5% of respondents completed the questionnaire after finding out which questions they had got wrong, this aspect has a negligible influence on the results.

Further studies should focus on identifying factors that make participants perceive contest tasks as being more difficult, features that have a positive impact on task popularity and the relationship between task popularity and perceived task difficulty. Such factors could be connection of the story of the task with the world of a child, including gender aspects, or connection of the theme of the task with a part of informatics or elements of computational thinking. Other factors could be the attractiveness of interactive tasks (in terms of visual aspect or user-friendliness) and potential laboriousness, for example the assumption that a number of variations will have to be tested to determine an answer or the need to fully understand the task instructions before anything else.

If we can understand the factors that cause participants to perceive task difficulty or popularity in different ways, we will be able to improve the way we classify task difficulty. We will also be able to design or modify tasks to increase participant motivation and to include the more useful ones in the school curriculum.

## References

1. Bebras website. <http://bebras.org>. Accessed 21 May 2020
2. Dagienė, V.: Information technology contests – introduction to computer science in an attractive way. *Inf. Educ.* **5**(1), 37–46 (2006)
3. Wing, J.M.: Research notebook: computational thinking: what and why? In: The Link. Carnegie Mellon, Pittsburgh (2010)
4. Selby, C., Woollard, J.: Computational thinking: the developing definition. In: Special Interest Group on Computer Science Education (SIGCSE) 2014. ACM, Atlanta (2014)
5. ISTE, CSTA: Operational Definition of Computational Thinking for K-12 Education (2011). <https://id.iste.org/docs/ct-documents/computational-thinking-operational-definition-flyer.pdf>
6. Dagienė, V., Sentance, S.: It's computational thinking! Bebras tasks in the curriculum. In: Brodnik, A., Tort, F. (eds.) *ISSEP 2016*. LNCS, vol. 9973, pp. 28–39. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46747-4\\_3](https://doi.org/10.1007/978-3-319-46747-4_3)
7. van der Vegt, W.: Predicting the difficulty level of a Bebras task. *Olympiads Inform.* **7**, 132–139 (2013)
8. van der Vegt, W.: How hard will this task be? Developments in analyzing and predicting question difficulty in the Bebras challenge. *Olympiads Inform.* **12**, 119–132 (2018). <https://doi.org/10.15388/oi.2018.10>

9. Pohl, W., Hein, H.-W.: Aspects of quality in the presentation of informatics challenge tasks. In: Jekovec, M. (ed.) *The Proceedings of International Conference on Informatics in Schools: Situation, Evolution and Perspectives, ISSEP 2015*, pp. 21–32. University of Ljubljana, Ljubljana (2015)
10. Dagienė, V., Futschek, G.: Bebras international contest on informatics and computer literacy: criteria for good tasks. In: Mittermeir, R.T., Sysło, M.M. (eds.) *ISSEP 2008. LNCS*, vol. 5090, pp. 19–30. Springer, Heidelberg (2008). [https://doi.org/10.1007/978-3-540-69924-8\\_2](https://doi.org/10.1007/978-3-540-69924-8_2)
11. Vaníček, J.: Bebras informatics contest: criteria for good tasks revised. In: Gülbahar, Y., Karataş, E. (eds.) *ISSEP 2014. LNCS*, vol. 8730, pp. 17–28. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-09958-3\\_3](https://doi.org/10.1007/978-3-319-09958-3_3)
12. Vaníček, J.: What makes situational informatics tasks difficult? In: Brodnik, A., Tort, F. (eds.) *ISSEP 2016. LNCS*, vol. 9973, pp. 90–101. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46747-4\\_8](https://doi.org/10.1007/978-3-319-46747-4_8)
13. Lonati, V., Malchiodi, D., Monga, M., Morpurgo, A.: How presentation affects the difficulty of computational thinking tasks: an IRT analysis. In: *Proceedings of the 17th Koli Calling International Conference on Computing Education Research*, pp. 60–69. ACM, New York (2017). <https://doi.org/10.1145/3141880.3141900>
14. Chiazese, G., Arrigo, M., Chifari, A., Lonati, V., Tosto, C.: Educational robotics in primary school: measuring the development of computational thinking skills with the Bebras tasks. *Informatics* **6**(4), 43 (2019). <https://doi.org/10.3390/informatics6040043>
15. Černochová, M., Vaníček, J.: Informatics education: current state and perspectives of development within the system of field didactics in the Czech Republic. *Int. J. Inform. Commun. Technol. Educ.* **4**(3), 14–31 (2015). <https://doi.org/10.1515/ijicte-2015-0011>
16. Dagienė, V., Futschek, G., Stupurienė, G.: Teachers' constructionist and deconstructionist learning by creating Bebras tasks. In: Sipitakiat, A., Tutiya-phuengprasert, N. (eds.) *Constructionism in Action 2016*, pp. 257–264. Suksapattana Foundation, Bangkok (2016)
17. Manabe, H., Tani, S., Kanemune, S., Manabe, Y.: Creating the original Bebras tasks by high school students. *Olympiads Inform.* **12**, 99–110 (2018). <https://doi.org/10.15388/oi.2018.08>
18. Berki, J., Drábková, J.: *Základy informatiky pro 2. stupeň ZŠ*. Textbook. Technical University of Liberec, Liberec (2020)
19. Hromkovič, J., Lacher, R.: *Einfach informatik. Lösungen finden*. Textbook. Klett und Balmer, Baar (2019)
20. van der Vegt, W., Schrijvers, E.: Analyzing task difficulty in a Bebras contest using cuttle. *Olympiads Inform.* **13**, 145–156 (2019). <https://doi.org/10.15388/oi.2019.09>
21. Yagunova, E., Podznyakov, S., Ryzhova, N., Razumovskaia, E., Korovkin, N.: Tasks classification and age differences in task perception. Case study of international on-line competition “Beaver”. In: Jekovec, M. (ed.) *The Proceedings of International Conference on Informatics in Schools: Situation, Evolution and Perspectives, ISSEP 2015*, pp. 33–43. University of Ljubljana, Ljubljana (2015)
22. Bellettini, C., Lonati, V., Malchiodi, D., Monga, M., Morpurgo, A., Torelli, M.: How challenging are Bebras tasks?: an IRT analysis based on the performance of Italian students. In: *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, pp. 27–32. ACM, New York (2015). <https://doi.org/10.1145/2729094.2742603>
23. Tort, F., Drot-Delange, B., Mano, M.: Filles et informatique: qu'en est-il du concours Castor?. In: Henry, J., Aude, N., Vandeput, E. (eds.) *L'informatique et le numérique en classe. Qui, quoi, comment?*, pp. 69–84. Presses Universitaires Namur, Namur (2017)
24. Hubwieser, P., Hubwieser, E., Graswald, D.: How to attract the girls: gender-specific performance and motivation in the Bebras challenge. In: Brodnik, A., Tort, F. (eds.) *ISSEP 2016. LNCS*, vol. 9973, pp. 40–52. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46747-4\\_4](https://doi.org/10.1007/978-3-319-46747-4_4)

25. Stupurienė, G., Vinikienė, L., Dagienė, V.: Students' success in the Bebras challenge in Lithuania: focus on a long-term participation. In: Brodник, A., Tort, F. (eds.) ISSEP 2016. LNCS, vol. 9973, pp. 78–89. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46747-4\\_7](https://doi.org/10.1007/978-3-319-46747-4_7)
26. Izu, C., Mirolo, C., Settle, A., Mannila, L., Stupurienė, G.: Exploring Bebras tasks content and performance: a multinational study. *Inform. Educ.* **16**(1), 39–59 (2017). <https://doi.org/10.15388/infedu.2017.03>
27. Budinská, L., Mayerová, K., Šimandl, V.: Differences between 9–10 years old pupils' results from Slovak and Czech Bebras contest. In: Pozdniakov, S.N., Dagienė, V. (eds.) ISSEP 2018. LNCS, vol. 11169, pp. 307–318. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-02750-6\\_24](https://doi.org/10.1007/978-3-030-02750-6_24)
28. Tomcsányi, P., Vaníček, J.: International comparison of problems from an informatic contest. In: Mechlová, E., Valchař, A. (eds.) Proceedings of the Information and Communication Technology in Education 2009, pp. 219–223. University of Ostrava, Ostrava (2009)
29. Šimandl, V.: Ochrana osobních údajů podle GDPR v soutěži Bobřík informatiky. In: DidInfo 2019, pp. 146–151. Univerzita Mateja Bela, Banská Bystrica (2019)
30. Cohen, J.: *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Lawrence Erlbaum Associates, New Jersey (1988)
31. Chráška, M.: *Metody pedagogického výzkumu*, 2nd edn. Grada, Praha (2016)
32. King, A.P., Eckersley, R.J.: Descriptive statistics II: bivariate and multivariate statistics. In: *Statistics for Biomedical Engineers and Scientists*, pp. 23–56. Academic Press (2019). <https://doi.org/10.1016/b978-0-08-102939-8.00011-6>
33. King, A.P., Eckersley, R.J.: Inferential statistics III: nonparametric hypothesis testing. In: *Statistics for Biomedical Engineers and Scientists*, pp. 119–145. Academic Press (2019). <https://doi.org/10.1016/b978-0-08-102939-8.00015-3>
34. Keller, J., Landhäuser, A.: The flow model revisited. In: Engeser, S. (ed.) *Advances in Flow Research*. Springer, New York (2012). [https://doi.org/10.1007/978-1-4614-2359-1\\_3](https://doi.org/10.1007/978-1-4614-2359-1_3)
35. Engeser, S., Schiepe-Tiska, A.: Historical Lines and an overview of current research on flow. In: Engeser, S. (ed.) *Advances in Flow Research*. Springer, New York (2012). [https://doi.org/10.1007/978-1-4614-2359-1\\_1](https://doi.org/10.1007/978-1-4614-2359-1_1)
36. Li, W., Lee, A., Solmon, M.: The role of perceptions of task difficulty in relation to self-perceptions of ability, intrinsic value, attainment value, and performance. *Eur. Phys. Educ. Rev.* **13**(3), 301–318 (2007). <https://doi.org/10.1177/1356336X07081797>
37. Mangos, P.M., Steele-Johnson, D.: The role of subjective task complexity in goal orientation, self-efficacy, and performance relations. *Hum. Perform.* **14**(2), 169–185 (2001). [https://doi.org/10.1207/S15327043HUP1402\\_03](https://doi.org/10.1207/S15327043HUP1402_03)
38. Cussó-Calabuig, R., Farran, X.C., Bosch-Capblanch, X.: Are boys and girls still digitally differentiated? The case of Catalanian teenagers. *J. Inf. Technol. Educ. Res.* **16**(1), 411–435 (2017). <https://doi.org/10.28945/3879>
39. Birol, C., Bekirogullari, Z., Etcı, C., Daglı, G.: Gender and computer anxiety, motivation, self-confidence, and computer use. *Eurasian J. Educ. Res.* **34**, 185–198 (2009)
40. Vekiri, I.: Boys' and girls' ICT beliefs: do teachers matter? *Comput. Educ.* **55**(1), 16–23 (2010). <https://doi.org/10.1016/j.compedu.2009.11.013>
41. Dagiene, V., Mannila, L., Poranen, T., Rolandsson, L., Stupuriene, G.: Reasoning on children's cognitive skills in an informatics contest: findings and discoveries from Finland, Lithuania, and Sweden. In: Gülbahar, Y., Karataş, E. (eds.) ISSEP 2014. LNCS, vol. 8730, pp. 66–77. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-09958-3\\_7](https://doi.org/10.1007/978-3-319-09958-3_7)