I2, F2) changed the strategy of having hints always available or available only after an incorrect attempt. The exercises of type S (S1 and S2) were of MC type, the exercises of type M (M1 and M2) were of MR type, the exercises of type I and F (I1, F1, I2 and F2) were of FIB type. We did not take into account the interaction of students with the hints of the MC exercises, because we detected that some students did not understand well the meaning of the hints for these exercises, so their behavior can not be reliable.

All the exercises of the experiment can be done in a maximum of 3 attempts, and have associated 2 hints (a first hint and a second hint). These two hints for each exercise give different information. We took into account the difficulty of the exercises. Some exercises were easy (S1, M1, I1, F1) and others were difficult (S2, M2, I2, F2). We made this distinction of difficulty based on the success percentage of students in the different exercises.

3   Analysis of the Experiment

This section makes an analysis of the experiment and draws conclusions based on the taken data. We have defined some behavior indicators to measure different student behaviors and the conclusions are based on these indicators.

3.1   Penalty Effect in the Hints

We define indicator I_1 as the percentage of students that view the first hint of an exercise with respect to the total number of students that did the exercise. We define indicator I_2 in a similar way but for the second hint. These indicators measure the tendency of students to view hints.

The values of I_1 and I_2 for the different exercises that change the penalty strategy are shown in Table 1. We can compare the values of these indicators when there is penalty and when there is no penalty.

We can see that for all the exercises, the number of students that view the hints is greater when there is no penalty than when there is penalty. This is the expected behavior, because the penalty discourages students to view hints. The difference is more significant for the first hint than for the second hint.

We can also see that the difference between having penalty or not is more significant for the difficult exercises than for the easy exercises. This can be explained because when an exercise is difficult, the student needs more help, and the penalty is a more important factor to decide whether to view the hint or not.

I2, F2) changed the strategy of having penalties for viewing hints, or not but with a maximum limit of hints to select. Finally, four exercises (S4, M4, I4, F4) changed the strategy of hint availability. Each student interacted only with one of these hinting techniques for a specific problem, and this is selected randomly at the beginning but in a way that there is a balance in the number of total presented hinting techniques.

3 Results and Analysis

Tables 1 and 2 show the statistics for the 12 exercises, comparing the hinting techniques regarding penalties for selecting hints (table 1) and hint availability (table 2). The tables distinguish the type of problem (MC, MR or FIB), their difficulty (in a scale from 0 to 4), the total number of interactions with each one of the techniques compared, the number of times a user selected at least one hint for each one of the techniques compared, and the probability of requesting at least one hint (P) provided in a confidence interval for each one of the hinting techniques (applying the binomial non-parametric test with a 95% probability). If there is not an intersection of the intervals of the two techniques compared for a specific exercise, then there is a statistically significant difference between both techniques for this specific exercise. For each specific problem, samples are independent. We are planning to analyze more hinting techniques and integrate the results with a logistic regression model taking into account the random effect of the same student answering different exercises.

Table 1. Statistics for penalties or not for selecting hints, for the eight exercises

Problem	Type	Diff.	Total Interact. with Penalties	Total Interact. without Penalties	Selected Hints with Penalties	Selected Hints without Penalties	P (%) Hint with Penalties	P (%) Hint without Penalties
S1	MC	1	48	46	1	3	[0, 10]	[1, 16]
M1	FIB	1	43	40	27	37	[49, 75]	[82, 98]
I1	FIB	3	43	43	35	35	[69, 91]	[69, 91]
F1	MC	3	42	47	7	13	[8, 29]	[17, 40]
S2	MR	4	47	45	30	42	[51, 75]	[84, 98]
M2	MR	3	42	42	25	34	[45, 73]	[68, 90]
I2	MR	2	44	41	19	32	[31, 57]	[65, 88]
F2	MR	1	37	36	11	11	[18, 45]	[18, 45]

Table 2. Statistics for hints directly available or not for the four exercises

Problem	Type	Diff.	Total Interact. no available	Total Interact. directly available	Selected Hints no available	Selected Hints directly available	P (%) Hint no available	P (%) Hint directly available
S4	FIB	3	39	46	23	34	[45, 72]	[61, 84]
M4	FIB	4	36	40	27	37	[60, 86]	[82, 98]
I4	FIB	3	28	41	19	29	[51, 82]	[57, 82]
F4	FIB	1	30	37	6	22	[9, 36]	[44, 73]

From table 1, there is a significant difference between the number of hints selected, comparing the cases of having penalties or not, when the amount of selected hints is not too high nor too low. As MC problems make the number of requested hints to be low then none of the cases resulted in a significant difference. As FIB problems make the number of requested hints to be high, then if the problem is difficult the amount of requested hints will be high, and there will not be significant difference; if it is easy or medium, then the number of requested hints will be medium and there will be significant difference. Finally, the MR problems have also certain effect that makes students to do a lot of attempts until they reach the correct answer, but they need much more attempts than for MC problems. So if the problem is easy then there will be few requested hints and the difference will not be significant, but if the problem is medium or difficult then the amount of requested hints will not be too high nor too low, so this can make the difference significant. A similar effect happens with the hinting technique of having hints directly available or not. In this case, it was only tested with FIB problems. A similar analysis can be performed from table 2.

4 Conclusions

There were several statistically significant differences in the number of hint requests between the hinting technique of having penalties for selecting hints or not, depending on the type and difficulty of the problem. In a similar way, there was a statistically significant difference in the number of hint requests comparing hints directly available or not, depending on the problem difficulty. The probabilities for selecting hints were provided for each case, as confidence intervals.

References

1. Aleven, V., Mclaren, B., Roll, I., Koedinger, K.R.: Toward Meta-cognitive Tutoring: A Model of Help Seeking with a Cognitive Tutor. Int. J. Artif. Intell. 16, 101–128 (2006)
2. Feng, M., Heffernan, N.T., Koedinger, K.R.: Predicting state test scores better with intelligent tutoring systems: developing metrics to measure assistance required. In: Ikeda, M., Ashley, K.D., Chan, T.-W. (eds.) ITS 2006. LNCS, vol. 4053, pp. 31–40. Springer, Heidelberg (2006)
3. Beck, J.E., Chang, K., Mostow, J., Corbett, A.: Does Help Help? Introducing the Bayesian Evaluation. In: Woolf, B.P., Aïmeur, E., Nkambou, R., Lajoie, S. (eds.) ITS 2008. LNCS, vol. 5091, pp. 540–550. Springer, Heidelberg (2008)
4. Conejo, R., Guzmán, E., Pérez de-la Cruz, J.L., Millán, E.: An Empirical Study About Calibration of Adaptive Hints in Web-Based Adaptive Testing Environments. In: Wade, V.P., Ashman, H., Smyth, B. (eds.) AH 2006. LNCS, vol. 4018, pp. 71–80. Springer, Heidelberg (2006)
5. Muñoz-Merino, P.J., Delgado Kloos, C.: A software player for providing hints in problem-based learning according to a new specification. Computer Applications in Engineering Education 17, 272–284 (2009)