Exploring the Need for Explainable Artificial Intelligence (XAI) in Intelligent Tutoring Systems (ITS)

Vanessa Putnam
Computer Science
University of British Columbia
Vancouver, British Columbia, Canada
vputnam@cs.ubc.ca

Cristina Conati
Computer Science
University of British Columbia
Vancouver, British Columbia, Canada
conati@cs.ubc.ca

ABSTRACT

This work is the first step towards understanding when and if it is necessary for an Intelligent Tutoring System (ITS) to explain its underlying user modeling techniques to students. We conduct an initial pilot study to explore student attitudes towards incorporating explanations to an ITS, by asking participants for suggestions on the type of explanations, if any, that they would like to see. Our results indicate an overall positive sentiment towards wanting explanation and suggest a few design directions for incorporating explanation into an existing ITS.

CCS CONCEPTS

Human Centered Computing → User Studies; Laboratory experiment

KEYWORDS

Explainable Artificial Intelligence (XAI); Intelligent Tutoring Systems (ITS); User Modeling

ACM Reference format:
Vanessa Putnam and Cristina Conati. 2019. Exploring the Need for Explainable Artificial Intelligence (XAI) in Intelligent Tutoring Systems (ITS). In Joint Proceedings of the ACM IUI 2019 Workshops, Los Angeles, USA, March 20, 2019, 7 pages.

1 Introduction

Lack of transparency in many AI techniques has created a growing interest in incorporating explanation in artificial intelligence systems, in order to express an intelligent system's behavior in a way that is interpretable and understandable. Existing work in XAI aims to make AI techniques more transparent in hope of increasing user trust and providing users with information that can develop their understanding of an intelligent system's learning mechanism.

Some existing research suggests that explanations may be useful to express an intelligent system's modeling technique to users [6,8,14]. Such work [6,8,14] points to the promise of XAI. However, other works uncover that there are circumstances when added explanation is not always beneficial [3,4,15]. Thus, existing literature suggests that the capability of explaining themselves may be important for AI systems, however there is still work to be done in understanding when and if explanations are necessary or useful.

Existing literature in XAI focuses on various domains (i.e. recommender systems, GUI customization, etc.), however our work is to the best of our knowledge the first to investigate the value of enabling an intelligent system to explain its behavior in the context of a specific subfield of AI namely, intelligent tutoring systems (ITS). In this field, AI algorithms are utilized to create educational systems that can monitor relevant needs, states and properties of their students during interaction, and provide personalized instruction accordingly [1]. These systems may have a potentially long lasting impact on a student's learning, therefore increasing the need for interpretability and transparency of the algorithms driving an ITS's behavior [6]. The long-term goal of the work presented in this paper is to investigate whether having an ITS that can explain its behavior helps fulfill this need, and how the explanations affect the ITS effectiveness, with specific focus on if/how individual student differences influence the effect of explanations.

This paper presents an initial step toward our long-term goal, consisting of a pilot study to elicit student opinions on the explanations they would like to see when interacting with an existing ITS known as the Adaptive CSP (ACSP) applet, an exploratory learning environment that provides tools for students to explore an algorithm that solves constraint satisfaction problems in an interactive simulation. The ACSP adapts in real-time to provide personalized instruction to students using interventions in the form of textual hint messages. The ACSP was evaluated in a prior study [10] and we follow the same procedure as [9] to guide our study procedure.

In the following sections we describe the ACSP user modeling mechanisms, followed by the modifications we made to the ACSP in order to gather suggestions on the type of explanations students would like to see. Lastly, we discuss our results on the explanations students suggested, and the reasons why students want the explanations they suggested.

2 Related Work

Initial evidence suggests that explanations may be useful for increasing an intelligent systems overall effectiveness (i.e. personalization, perceived value, and increased compliance [6,8,14]). Many positive results on explanation have been found in the field of recommender systems [8,11]. For example, explanations have been shown to successfully explain the positive and negative features of a recommended item compared to alternative items [11]. Additionally other domains such as personalizable machine learning have also found that explanations are helpful for an intelligent system to explain its reasoning to an end user, who in turn explains corrections back to the system [14].

On the other hand, other works present findings that indicate XAI is not always wanted or necessary [3,4,15]. Results in [4] question the need for explanation in intelligent interactive systems that assist users in making low-cost decisions, concluding that explanations are not critical in certain systems namely, low cost intelligent interactive systems (e.g. Google suggestions). Similarly, results in [3] express uncertainty regarding the importance and usage of explanations in a mixed-initiative system for GUI customization, showing initial evidence that some users find less utility in explanations than others and conclude that explanations may not be crucial for all systems and all users. Additionally, results in [15] point to potential individual differences with respect to preferences and attitudes towards the utility of social versus task-based recommender systems. Contrasting evidence in existing literature suggests that there is still work to be done in understanding when explanation is necessary or useful for different AI systems, users, and domains.

3 The AI Space ACSP Applet

A Constraint Satisfaction Problem (CSP) consists of a set of variables, variable domains and a set of constraints on legal variable-value assignments. Solving a CSP requires finding an assignment that satisfies all constraints. The CSP applet illustrates the Arc Consistency 3 (AC-3 [13]) algorithm for solving CSPs represented as networks of variable nodes and constraint arcs (see Figure 1). AC-3 iteratively makes individual arcs consistent by removing variable domain values inconsistent with a given constraint, until all arcs have been considered and the network is consistent. Then, if there remains a variable with more than one domain value, a procedure called domain splitting is applied to that variable in order to split the CSP into disjoint cases so that AC-3 can recursively solve each case.

The ACSP applet demonstrates the AC-3 algorithm dynamics through interactive visualizations on graphs using color and highlighting (Figure 1). The applet provides several mechanisms for the interactive execution of the AC-3 algorithm, including: (1) Fine Step: use the fine step button to see how AC-3 goes through its three basic steps (selecting an arc, testing it for consistency, removing domain values to make the arc consistent); (2) Direct Arc Click: directly click on an arc to apply all these steps at once; (3) Auto AC: automatically fine step on all arcs one by one using the auto arc consistency button; (4) Stop: pause auto arc consistency; (5) Domain Split: select a variable to split on, and specify a subset of its values for further application of AC-3 (see pop-up box in the left side of Figure 1); (6) Backtrack: recover alternative sub-networks during domain splitting; (7) Reset: return the graph to its initial status.

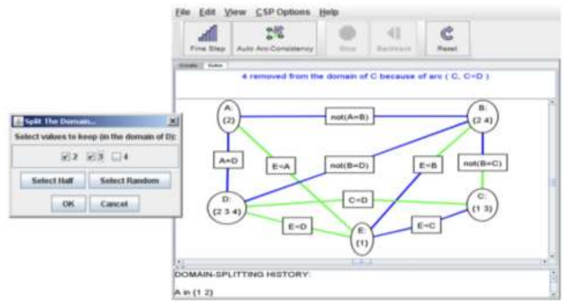

Figure 1. The CSP applet with an example CSP problem

In addition to the applets mechanisms for solving CSP's, the ACSP also includes the ability to adapt in real-time to provide personalized instruction to users using interventions. These interventions are provided to students in order to encourage effective learning behaviors and discourage ineffective ones. In the next section, we briefly summarize the user modeling approach used to determine in real-time during user interaction if and why a user needs an intervention [9,10]. The two following sections introduce the mechanism designed to generate and deliver these interventions in the CSP applet [9,10]. The final section describes the 'Explain Hint Feature' that was added for the purpose of this work to solicit student suggestions on explanations they would like to see in the applet.

3.1 Modeling Student Learning

The user modeling approach used in the adaptive-CSP applet consists of two phases: Behavior Discovery and User Classification (Figure 2). In Behavior Discovery (Figure 2- top), data from existing interaction logs is preprocessed into feature vectors where features consist of statistical measures that summarize the user's actions in the interface (e.g., action frequencies, time interval between actions). Each vector summarizes the behaviors of one user. A clustering algorithm then groups these vectors according to their similarities, thus identifying users who interact similarly with the interface. Next, association rule mining is applied to each cluster to extract its common behavior patterns. (Table 1 shows examples of these rules).

Clusters are then analyzed to identify how they relate to student learning performance. Thus, the Behavior Discovery phase generates groups of users who are associated with different levels of learning performance, as well as sets of interaction behaviors typical of each group.

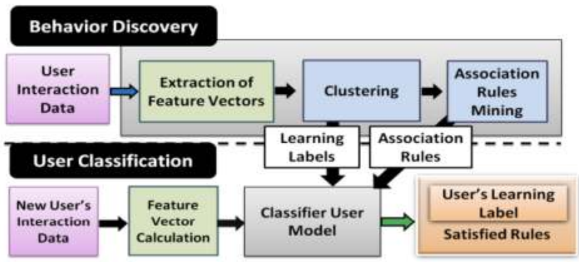

Figure 2. User Modeling Framework

The user model for both of the described phases was built on a dataset of 110 users obtained from two previous studies on the simulation [10]. From this dataset, the Behavior Discovery generated two clusters of users that achieved significantly different learning levels, labeled as High Learning Gain (HLG) and Low Learning Gain (LLG) groups. Table 1 shows a subset of behavior patterns (association rules) discovered for each cluster from this dataset. A total of four and fifteen rules were found for the HLG and LLG respectively.

The *User Classification* (Figure 2-bottom) phase uses the clusters and class association rules extracted in the Behavior Discovery phase to build an online classifier user model. This classifier assesses in real-time the (possibly evolving) learning performance of a new user by (i) incrementally building a feature vector based on the interface actions seen so far; (ii) classifying this vector in one of the available clusters (HLG or LLG). Note that the classification can change over time depending on the evolution of the user's interaction behaviors.

3.2 Adaptive Interventions

In addition to classifying a user as HLG or LLG, the ACSP's user model also returns the satisfied association rules causing that classification decision. These rules represent the distinctive interaction behaviors of a specific user so far, including a subset of behaviors satisfied for HLG (effective), as well as a subset of those satisfied for LLG (ineffective).

The process of providing adaptive interventions starts by identifying which of the intervention items in Table 2 are relevant at any given point of a user's interaction with the ACSP. More specifically, when a user is classified as LLG, the applet identifies which detrimental behaviors a user should stop performing and/or which beneficial behaviors they should adopt, based on the association rules that caused the classification.

It is possible for the system to have several intervention items that are relevant given a user's behaviors at any given point of their interaction with the ACSP. To avoid confusing or overwhelming the user, the applet only delivers one intervention at a time, chosen based on a ranking that reflects the relevance of the behaviors being targeted for learning. In other words, each intervention item is assigned a score that is calculated as the sum of the weights of the association rules which triggered that item. The weight of a rule in turn indicates its importance in classifying a user as a high or low learner. At each hinting opportunity, the intervention item with the highest score is chosen among the relevant intervention items.

Rules for HLG cluster:
Rule1: Infrequently auto solving the CSP
└ **Rule 2:** Infrequently auto solving the CSP and infrequently stepping through the problem
Rule 3: Pausing for reflection after clicking CSP arcs
Rules of LLG cluster:
Rule 4: Frequently backtracking through the CSP and not pausing for reflection after clicking CSP arcs
Rule 8: Frequently auto solving the CSP and infrequently clicking on CSP arcs
Rule 10: Frequently resetting the CSP

Table 1. A subset of representative rules for HLG and LLG clusters

3.3 Delivering Adaptive Interventions

The ACSP delivers adaptive interventions incrementally. Thus, each selected intervention item is first delivered with a textual hint that prompts or discourages a target behavior, followed when needed by a textual hint that reiterates the same advice, accompanied by a related interface adaptation that can help the user follow the advice (e.g., highlighting relevant interface items). The general mechanism to deliver incremental adaptive interventions in the CSP applet works as follows:

(i) Each intervention item selected for delivery (target item in the rest of this section) is first presented as a textual hint phrased subtly as a suggestion for behaviors to be adopted or avoided (level-1). For instance, a level-1 textual hint for the *Using Direct Arc Click more often* intervention item in Table 2 is "Do you know that you can tell AC-3 which arc to make consistent by clicking on that arc?" which aims to promote the Direct Arc Click action.

(ii) After receiving a level-1 hint on the target item, the student is given some time to change their behavior accordingly (a reaction window equal to 40 actions). During this time, the user model will keep updating the feature vector describing the user interaction behavior, its classification and the ranked list of relevant intervention items (excluding the target item), delivering any of them as needed.

(iii) At the end of the time window, based on the updated feature vector the user model determines whether the user has followed the hint for the target item or not. If at this point the rules that generated the level-1 hint for the target item are still satisfied, then this means the user has not followed the hint and the target item is selected for delivery again.

(iv) In this case, a level-2 hint is delivered using a more forceful approach than a level-1 hint containing a textual message that reiterates the same advice along with highlighting relevant interface elements to help the user follow the hint. For instance, a level-2 textual hint for the *Using Direct Arc Click more often* intervention item in Table 2 is "As I suggested earlier, you can choose which arc to make consistent next by clicking on it. I have highlighted the relevant arcs for you." Simultaneously, the applet highlights the relevant CSP arcs.

Intervention Description
Using Direct Arc Click more often
Spending more time after performing Direct Arc Clicks
Using Reset less frequently
Using Auto Arc-consistency less frequently
Using Domain Splitting less frequently (only when appropriate)
Spending more time after performing Fine Steps
Using Back Track less frequently (only when appropriate)
Using Fine Step less frequently
Spending more time after performing rest for planning

Table 2. Description of hints

3.4 Soliciting Student Suggestions for Explanation

In order to gather suggestions on the type of explanations that students would like to have in the ACSP, we incorporated an 'explain hint' feature into the ACSP interventions. Once a student is delivered a level-1 or level-2 hint, the student can select to have the hint explained (Figure 4A). Once the "Explain Hint" button is selected, the tools to input the desired explanation become visible (Figure 4B). Here, students select from the following options : why the system gave this hint, how the system chose this hint, some other explanation about this hint (with a field for user input), or no explanation for this hint. We use [2] to establish a 'why' type and 'how' type explanation. A why explanation will provide a chain of causal reasons (i.e. facts) about the system. Furthermore, a how explanation will provide information on the systems process that allows the system to establish these facts/causes. We define a response from a student (i.e. "Submit Response" in Figure 4b) to be a submission from a student.

4 User Study

In the following sections we discuss the participants recruited for our study, the procedure followed, and the material that was used to gather suggestions on explanations students would want to see incorporated in the applet.

4.1 Study Participants

Nine university students (two female) participated in the experiment. Participants were recruited from an introductory AI course to ensure the requirement that they had enough computer science knowledge (e.g. basic graph theory and algebra) to learn the concept of CSP's. One participant did not receive any hints during their interaction, and as a result, this participant's responses were excluded from the rest of our analysis.

4.2 Procedure

The procedure for our study was as follows: (1) students studied a textbook chapter on the AC-3 algorithm; (2) wrote a pre-test on the concepts covered in the chapter; (3) watched an introductory video on how to use the main functionalities of the ACSP applet; (4) used the ACSP applet[1] to solve two CSPs; (5) took a post-test analogous to the pre-test; and (6) answered a post-questionnaire and a follow up interview [2] that solicited feedback on the explanations they selected in the dialogue box in Figure 4.

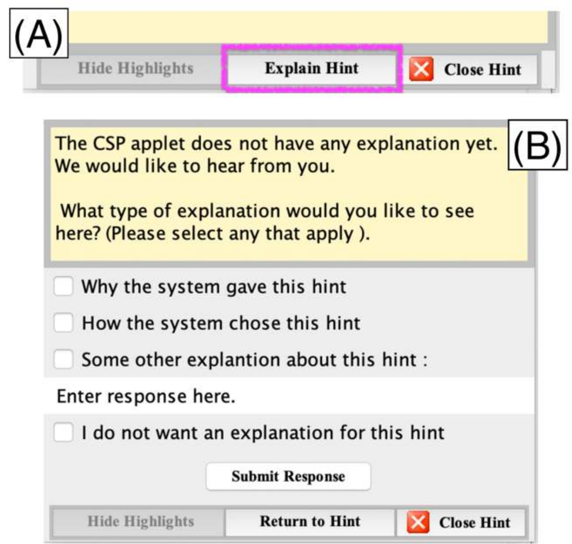

Figure 4. Explain Hint Feature

4.3 Study Material

The conceptual materials for this study include a pre-test, post-test, and two sample CSP problems to be solved with the ACSP. These materials were selected from homework questions used in an introductory AI course at a university in North America.

A questionnaire and follow up interview were also given to participants in order to corroborate the feedback gathered from the Explain Hint feature. Unlike the real time responses in the Explain Hint feature, the questionnaire was given to get a general understanding of when students want to know why or how. The follow up interview was given to get a more detailed understanding of the reasons why students want or do not want the explanations they suggested in the explain hint feature.

[1] When participants were exposed to the ACSP applet we told to to look at the explain hint feature at least once during their interaction but did not tell them that there was no explanation added to the systems hints. All participants viewed and responded to the explain hint feature at least once during their interaction.

[2] In our interview we asked participants if it was clear that we are looking for their suggestions on explanations they would like to see for system hints. All participants answered yes to this question.

5 Results on Type of Explanations Wanted

Our results show that each participant received an average of 8.5 hints per session, and that participants responded to 51% of hints. Of these responses, 86% were in response to level-1 hints, and 14% were in response to level-2 hints. It could be the case that the remining 49% hints participants did not select to explained is because they have already selected the explain hint and do not want to give feedback again or already know that there are no explanation for hints. In other words, we cannot conclude that all 49% of hints not responded too were due to the participant not wanting explanation.

In order to ensure participants understood the aforementioned explanation types we asked participants to give a verbal explanation, to the best of their ability, on how the system provides hints and why the system provides hints in our interview. From this feedback we determined that 5 of the 8 participants understood the distinction between a how type and why type explanation.

We encode the participants suggestions from the "Explain Hint" (Figure 4) feature as a Why, How, Other, or None response. Figure 5 gives a detailed distribution of participant responses. Our findings show that participants are generally interested in wanting explanations, and only report not wanting explanation less than 20% of the time in the Explain Hint feature. This is also true when looking at participant responses for level-1 and level-2 hints independently (Figure 5B-C). Additionally, Figure 5A shows participants report wanting to know why most often during their interaction, closely followed by wanting to know how. These findings are also true across all figures in Figure 5, indicating explanations for ACSP hints should be designed for both level-1 and level-2 hints.

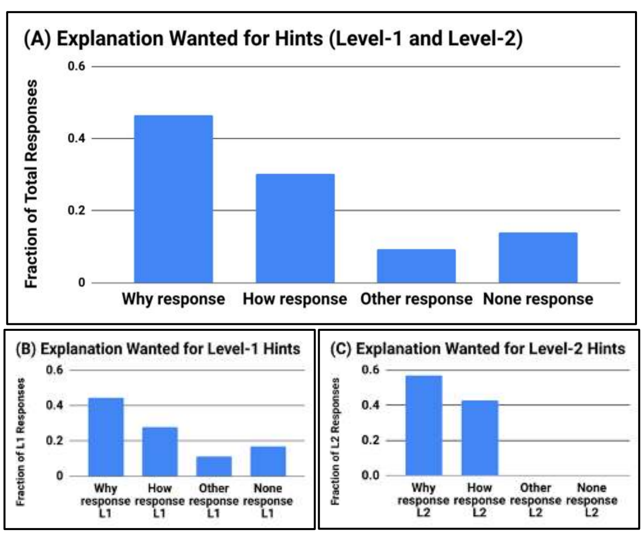

Figure 5. (A) histogram of the type of explanations wanted for both level-1 and level-2 hints. (B) histogram of the type of explanation wanted for level-1 hints (C) histogram of the type if explanation wanted for level-2 hints.

Findings from the explain hint feature are also mirrored in results from the follow-up questionnaire. When asked to rate statements on wanting to know why or how the system gives hints after interaction, participants voted positively overall. Using a Likert scale from 1: strongly disagree to 5: strongly agree, participants voted with an overall average of 4.4 wanting to know why, and an overall average of 3.9 wanting to know how (Table 3). There is also a larger standard deviation for wanting to know how, indicating that there is a larger range of student responses (i.e. students want to know why more consistently than they want to know how).

In addition to wanting to know why and how, we also received a few suggestions for other types of explanations participants would like to see in the dialogue box. One participant reported, "The hint needs to be more transparent so that I can trust it." Although this is not a type of explanation, we still consider this feedback to be useful, because it illustrates the students desire for explanation in order to trust the system. Additionally, for a hint that states, "You have reset the problem over and over again. Why don't you try using the other available actions instead of resetting the problem?" One participant suggested an explanation that specifies which options the applet would like him to try. This is an explanation that we will consider when developing explanations for the applet in the future.

	Mean	Std Dev	Max	Min
I want to know why.	4.375	0.75	5	3
I want to know how.	3.875	1.35	5	2

Table 3. Descriptive statistics for wanting to know why and how from Likert scale results

In addition to the type if explanations participants would like to see, we also analyzed the temporal pattern of students responses regarding explanations, to uncover when students are more likely to need or not need explanation during interaction. Specifically, we normalized the session time for all students and broke the session up into quartiles (i.e. ¼, ½, ¾, etc.). Then we took a ratio of the responses for why how and none over the total number of responses at that time. The results of this analysis are summarized in Figure 6, indicating that students want to know why during the first half of their interaction (i.e. first and second quartiles). A possible design direction for explanation would mean explaining why earlier in the student's interaction with the applet. Figure 6 also shows a tendency for participants to not want explanation increasing with time. We do not attribute this decrease in wanting explanation to the fact that participants know they won't be getting any explanation. This is because at this point in the interaction (i.e. third and fourth quartiles) there is an increase of students willingly responding 'no explanation' to the explain hint feature. This indicates that participants are more interested in explanations earlier in their interaction and may find them less necessary closer to the end of their interaction.

6 Results on Reasons for Explanations Wanted

In this section we combine the open-ended feedback from our follow up questionnaire and interview to understand the reasons participants want explanations for system hints. In the questionnaire participants were asked if there was a situation when explanation was needed for the systems hints. Additionally, in the interview participants were asked why they found the explanations they suggested valuable.

Across these sources two major themes emerge: participants need explanations, and consider them valuable when they are curious, and when they disagree with the systems decision making. One student stated they wanted explanations because they were curious to know how the hint was created. Other students expressed that they wanted explanation when they disagreed with the system. One student expressed value in an explanation that would justify a student's decision to ignore system hints, claiming:

Some students learn at different paces. So for a student that learns quickly, a reasoning behind a slow down hint may allow the student to see the reasoning and know 'oh this does not apply to me' and they can take note of that

This is an interesting finding since it suggests that explanations may be needed when students feel the systems hints are not useful, justified, or do not apply to them.

Participants also expressed other reasons why explanations would be valuable to them. One participant expressed that wanting to know the systems decision making was important to him. This implies that explanations for system hints would be necessary for participants who want to know the systems reasoning. Additionally, the same participant that suggested that a type of explanation (one that specifies which options the applet would like him to try) would be valuable because the explanation could guide him to trying different ways of solving the problem. This response expresses why explanations may be specifically valuable to ITS's, and reinforces the idea that incorporating explanation to the ACSP may be a feature students would like to see.

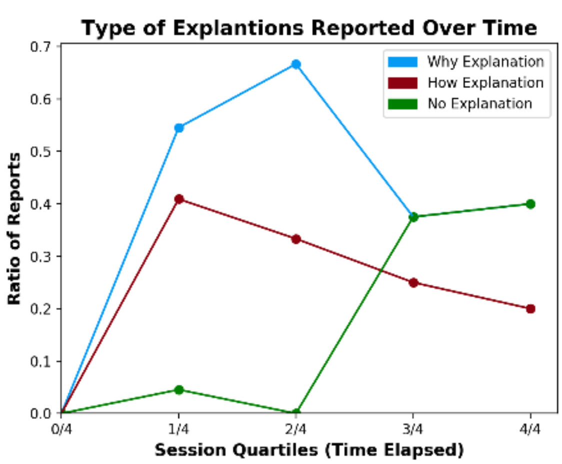

Figure 6. Explanation responses over time. X-axis normalizes session time for all students, broken up into quartiles. Y-axis is a ratio of reports for each type of explanation response over the total number of responses for each session quartile.

Other participants addressed situations when they needed an explanation during their interaction. One participant claimed explanation was needed for a hint when it was first delivered. This is useful information because it may be a reason for the large number of reports wanting to know why in the first quartile of Figure 6. Another participant expressed that they wanted an explanation for why a hint was delivered at that moment in time and not earlier, indicating that explaining the timing of a hint is also important to students.

Apart from participants who expressed positive feelings toward explanation, three out of the eight participants responded that they did not experience a situation when explanation was necessary. This indicates that not all participants needed or wanted explanation during their interaction and supports our future work investigating if/how individual student differences influence the effect of explanations. The remaining participants not accounted for in this section either did not answer the question or gave an answer that was not interpretable. For these participants we did not add their responses to our analysis.

7 Discussion and Future Work

Our findings suggest that incorporating explanation in the ACSP is a feature most students want to see. We find that participants express wanting some form of explanation for 54% of hints that were delivered. This is and underestimate if we account for the ambiguity of a participant not responding the explain hit feature (i.e. this could mean they do not want to give feedback or they do not want explanation). Of these 54% of responses, participants report that they would like to know why most often during their interaction with the system. Additionally, we uncover that explanations explaining why may be more necessary at the beginning of a student's interaction, and that students report not wanting explanations more the longer they interact with the system. We also find that curiosity and disagreeing with the systems decision making are common reasons for needing explanations amongst students.

We believe our results pave the way for further studies to incorporate added explanation to the ACSP. As a first step, we will use the suggestions obtained in this study to create explanations that make the described user model more transparent for students. We will also investigate potential individual differences that may impact explanation (i.e. personality traits, expertise, need for cognition [5]). Next, we will test the following research questions : Q1: Does incorporating explanation increase the ACSP applet effectiveness? Q2: Do individual differences impact the effectiveness of incorporating explanations? Here we evaluate *effectiveness* in terms of measures such as follow rate, student learning, trust, and acceptance. We will answer these questions in future studies to further investigate the need for explanation in ITSs and if/how individual student differences influence the effect of explanations.

REFERENCES

[1] Anderson, J. R., Boyle, C. F., & Reiser, B. J. (1985). Intelligent tutoring systems. Science, 228(4698), 456-462.

[2] Brian Y. Lim, Anind K. Dey, and Daniel Avrahami. 2009. Why and why not explanations improve the intelligibility of context-aware intelligent systems. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, New York, NY, USA, 2119-2128. DOI: https://doi.org/10.1145/1518701.1519023

[3] Bunt A., McGrenere J., Conati C. (2007) Understanding the Utility of Rationale in a Mixed-Initiative System for GUI Customization. In: Conati C., McCoy K., Paliouras G. (eds) User Modeling 2007. UM 2007. Lecture Notes in Computer Science, vol 4511. Springer, Berlin, Heidelberg

[4] Bunt, A., Lount, M., & Lauzon, C. (2012, February). Are explanations always important?: a study of deployed, low-cost intelligent interactive systems. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*(pp. 169-178). ACM.

[5] Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition. Journal of personality assessment, 48(3), 306-307.

[6] Conati, C., Porayska-Pomsta, K., & Mavrikis, M. (2018). AI in Education needs interpretable machine learning: Lessons from Open Learner Modelling. arXiv preprint arXiv:1807.00154.

[7] Ehrlich, K., Kirk, S. E., Patterson, J., Rasmussen, J. C., Ross, S. I., & Gruen, D. M. (2011, February). Taking advice from intelligent systems: the double-edged sword of explanations. In Proceedings of the 16th international conference on Intelligent user interfaces (pp. 125-134). ACM.

[8] Herlocker, J. L., Konstan, J. A., & Riedl, J. (2000, December). Explaining collaborative filtering recommendations. In Proceedings of the 2000 ACM conference on Computer supported cooperative work (pp. 241-250). ACM.

[9] Kardan, S., & Conati, C. (2015, April). Providing learning: An experimental evaluation. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems(pp. 3671-3680). ACM.

[10] Kardan, S. (2017). A data mining approach for adding adaptive interventions to exploratory learning environments (T). University of British Columbia. Retrieved from https://open.library.ubc.ca/cIRcle/collections/24/items/1.0348694

[11] Lawlor A., Muhammad K., Rafter R., Smyth B. (2015) Opinionated Explanations for Recommendation Systems. In: Bramer M., Petridis M. (eds) Research and Development in Intelligent Systems XXXII. SGAI 2015. Springer, Cham

[12] Miller, T. (2017). Explanation in Artificial Intelligence: Insights from the Social Sciences. [online] Arxiv.org. Available at: https://arxiv.org/abs/1706.07269 [Accessed 4 Apr. 2018].

[13] Poole, D.L. and Mackworth, A.K. Artificial Intelligence: Foundations of Computational Agents. Cambridge University Press, 2010.

[14] Todd Kulesza, Margaret Burnett, Weng-Keen Wong, and Simone Stumpf. 2015. Principles of Explanatory Debugging to Personalize Interactive Machine Learning. In Proceedings of the 20th International Conference on Intelligent User Interfaces (IUI '15). ACM, New York, NY, USA, 126-137. DOI: https://doi.org/10.1145/2678025.2701399

[15] Wiebe, M., Geiskkovitch, D. Y., & Bunt, A. (2016, March). Exploring user attitudes towards different approaches to command recommendation in feature-rich software. In Proceedings of the 21st International Conference on Intelligent User Interfaces (pp. 43-47). ACM.