

Reputation Systems' Impact on Help Seeking in MOOC Discussion Forums

Iris Howley, Gaurav Tomar, Oliver Ferschke, and Carolyn Penstein Rosé

Abstract—Many interactional archetypes from outside of learning contexts are being adapted and widely used for online learning environments without consideration for some of the side effects relevant to learner outcomes. Of particular concern is the effectiveness of help exchange in these learning environments. To address this need, this article explores how the reputation system features of up/downvoting, badges, and displayed expertise impact student helper selection in a peer help exchange system within a MOOC discussion forum. We draw from Expectancy Value Theory for Help Sources as a theoretical framework for positioning the work. Results from our field experiment show that up/downvoting has a negative impact on help seeking which is mitigated by the positive effect of Help Giver badges. The mechanism behind these results are then explored in a survey experiment investigating reputation systems' impact on students' expectancies, values, and costs for a help source.

Index Terms—Discussion forums, education, social computing, computer-aided instruction

1 INTRODUCTION

MASSIVE Open Online Courses (MOOCs) arose in popularity due to the potential to offer high quality instruction to tens of thousands of students for minimal cost. Along with the growing number of MOOC students, there is also a growing demand for supporting those students' learning in a scalable manner. Discussion forums are viewed as a way to increase engagement, learning, and a sense of community in online courses. However, participation in discussion forums has been sparse, and the quality of interaction that has occurred frequently leaves much room for improvement. To address these issues, MOOC designers incorporated features of reputation systems, especially up/downvoting, into discussion forums to encourage student engagement with the discussion board and the course in general. However, reputation systems were designed for contexts outside of learning and so it is a serious question whether the use of reputation system features is truly beneficial to students in online learning environments where the set of relevant concerns is broader. In particular, the most important goal of technologically enhanced learning environments is to support learning and learning-relevant behaviors, such as help seeking, which may be negatively impacted by interactional archetypes such as up/downvoting.

Appropriate help seeking is a necessary skill in becoming a successful self-regulated learner. Help seeking has been demonstrated to highly correlate with student achievement in classroom learning [1]. Students who do not seek help with difficult concepts, or who fail to consult with instructors, or who request inappropriate help are not as

likely to experience success as students who seek help effectively [1]. Without seeking necessary help, students may fail to understand complex concepts that they do not understand or are unable to comprehend on their own.

An alternative to a reputation system for influencing participation in a MOOC discussion forum is an application specifically designed to facilitate help exchange. In this article, we introduce such an application, which we refer to as Quick Helper. Quick Helper is a tool for students who ask questions in a MOOC discussion forum. It is designed to connect help seekers with appropriate help providers who are student peers. We use Quick Helper as a platform for gaining insights needed for the design of effective support for help exchange in MOOC settings, rather than as the ultimate solution to the help seeking problem.

In this paper, we show that up/downvoting of posts in a Quick Helper discussion forum leads to a negative impact on student help seeking. This negative effect can be eliminated with the use of badges, another common feature of reputation systems. Through the use of an associated survey, the causal mechanism for this interaction effect is shown to be evaluation anxiety. Evaluation anxiety [2] is a negative cost of an expectancy value motivation theory. We leverage other parts of this theory to understand the impact of additional discussion forum design features on help seeking behaviors. The main purpose of this paper is to illustrate how insights from Expectancy Value Theory for Help Sources [3] can predict and explain student motivations for seeking help in peer help exchange tools. Our results show that online courses can provide intelligent peer help exchange tools, but the effectiveness of these tools can be impacted by common design features.

2 PRIOR WORK

The Massive Open Online Course (MOOC) revolution began with a course on connectivism and connective learning in 2008 [4], but has grown to include a variety of different MOOC platforms, including edX, Coursera, and NovoEd

among others. Initial research on these platforms revealed surprisingly high dropout rates that have attracted the attention of the media. This revelation inspired considerable interest in preventing MOOC student dropout.

Forums are a common means of developing communication and community within MOOCs, but the large scale of these courses introduces several issues. Forums commonly lose participation due to poor thread management and what is perceived by students as an overwhelming number of discussion forum threads [5]. When these forums fail to properly sustain a sense of community, high rates of student dropout often follow.

Prior work has explored how student confusion in discussion forums is related to dropout rates in Algebra and Microeconomics MOOC courses [5]. The linguistic measure of confusion was estimated by a machine learning classification algorithm that was trained on Mechanical Turk participants' ratings of levels of confusion. Student confusion was associated with increased dropout from the MOOCs. However, resolution of posts expressing confusion mitigated this effect, decreasing the likelihood of dropping out by 22%. Receiving responses from peers likewise was associated with a reduction in the likelihood of drop out by 14%. Only a small minority of threads expressing confusion included any sort of instructor involvement in this study (i.e., 13-18%). This work suggests that connecting students on discussion forums to the help they need can have a positive impact on student dropout.

Based on these findings, Yang and colleagues [7] approached the problem by proposing a thread recommendation algorithm designed to operate in a batch processing mode. This algorithm would recommend discussion forum posts to students at the end of a time period based upon common patterns of past participation, and content of previous posts. The corpus-based evaluation of that general approach did not include necessary adaptations for a help exchange support tool that would operate in real time. The Quick Helper tool featured in this article fills that void as a first step towards broadening understanding of the design space for help exchange support in MOOCs.

Quick Helper is comparable in some ways to existing peer help exchange systems. In particular, the I-Help system unites discussion forums and one-on-one peer help exchange networks in a multi-agent architecture. I-Help incorporates student knowledge, interests, eagerness, helpfulness, interaction preferences, opinions of peers, and user actions when matching pairs of students for private discussion [8]. Matchmaker agents rely upon Bayesian modeling in their negotiations for helper-helpee matching.

Another system, the Intelligent IntraNet Help-Desk, uses student modeling to unite online course materials, a Cooperative Peer Response (CPR) facility, and Peer Help System (PHelpS) [9]. The CPR comprises multi-modal peer communications such as a discussion forum. The PHelpS collects evidence of student knowledge and communicates with student models to match students requesting help to qualified peer helpers. The student models rely on a profile

of various evaluations of the student along with a numeric overlay on a concept-topic structure that represents what each student has been taught as well as what concepts each student has demonstrated expertise.

Both the I-Help system and Intelligent Help-Desk require maintaining individual student models of expertise and preference. However, Quick Helper relies on recommendation algorithms and rule-based scheduling constraints to select three helpers. The recommendation algorithm operates over log data from the MOOC, and due to this approach it is easier to implement Quick Helper in new environments. Additionally, in this paper, the purpose of Quick Helper reaches beyond peer helper matchmaking into serving as a platform for better understanding student decisions to pursue help. We are not only interested in providing peer helpers to students, but also what motivates students to seek help from a particular peer.

2.1 Reputation Systems in MOOCs

A more recent approach to fostering high quality interaction in MOOCs is through reputation systems. The history of such approaches is rooted in large online communities for establishing trust relationships as well as crowd-sourcing the organization of content. Auction websites, review websites, as well as Question & Answer forums all use reputation systems as a way of fostering trust amongst strangers on the Internet [10]. Reputation systems lubricate the process for online commerce and exchange of services, goods, or expertise between strangers online. Judgments of reputation and reliability are involved anytime we work with new, unknown people and these evaluations are based on the information available to us [10]. In this way, reputation systems provide structure for trust-based interaction, but also provide crowd-sourced organization.

As an example, Stackoverflow.com is a Question and Answer forum for both "professional and enthusiast" computer programmers. Users post questions to the website that are answered by other users. Users receive points for all activities, including asking and answering questions. However, other users can vote on answers and questions, and so obtaining more positive votes earns more reputation points. A user that votes down a particular answer will lose 1 reputation point, possibly as a means to stop users from downvoting excessively. As a user gains more points, they are able to access progressively more features on the site, including the ability to vote up, vote down and act as a moderator (i.e., edit other users' content). This functions as a reward system to not only encourage users to produce higher quality or more popular content, but also to engage increasingly more with the website and the community. Posts that are upvoted more than other posts may appear under the "Interesting", "Featured", or "Hot" tags, and so the reputation system is also leveraged as crowd-sourcing the organization of the website content.

Many MOOCs implement reputation systems in their discussion forums for many of the same reasons as www.Stackoverflow.com: to encourage engagement with

the discussion forums and organize content from thousands of students. Coatze et al. [11] determined that using reputation systems improves the response time and number of responses to discussion threads. The authors used a reputation system based on upvoting of user discussion posts and comments. Students earned a reputation score over time that could provide them with forum moderation capabilities. While students using the forum version with the reputation system experienced improved response time to posts, people were less likely to post to that forum. That is, the reputation system with upvoting (and no downvoting) increased the number of posts, but decreased the number of people posting. The authors also found that the basic forum, without a reputation system, actually contained more questions. This suggests that reputation systems might negatively impact help seeking. However, it remains an open question why the forum with the reputation system had this effect on help seeking and why the reputation system discouraged users from posting.

Furthermore, there was no effect of the reputation system on final grades or on forum retention [11] which suggests that a reputation system provides some benefits to increasing engagement in MOOC discussion forums, but perhaps not enough to impact longer term outcomes. Particularly when considering evaluative up/downvoting, course technology designers must consider help seeking and other learning-relevant behaviors. Interaction archetypes such as voting that were designed for purposes outside of learning contexts may have unintended consequences on valuable behaviors in learning contexts.

Often included within reputation systems, and commonly used in online learning environments are badging systems. In some MOOCs, badges are awarded to students for achieving particular goals or completing certain activities [12][13]. These badges can be viewed as an approach to providing feedback of successful progress to students, but can also be used as extrinsic rewards motivating more interaction and participation in course materials. Mozilla's Open Badge Infrastructure and other badging frameworks attempt to standardize badging systems so an individual's learning can be more easily understood across the Internet [13]. However, these badges may also be provided to a reputation system. Badging systems have the potential to not just reward students for good behavior, but also to signal to other students the achievement of milestones.

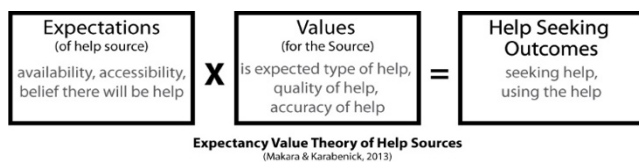


Fig. 1. Expectancy Value Theory for help sources model [3].

3 THEORETICAL FRAMEWORK: EXPECTANCY VALUE THEORY FOR HELP SOURCES

In order to achieve the positive effects of reputation systems without inadvertently introducing negative side-effects, we must understand the space of underlying factors and their interplay in MOOC contexts. The space is best defined and understood within an appropriate theoretical framework. We adopt and expand upon Makara and Karabenick's (2013) Expectancy Value model of Help Seeking for Help Sources (EVT-HS) [3].

Once the student has determined that she requires help, she must select a source from which to seek help and then follow through with the help request [14]. EVT-HS, as illustrated in Fig. 1, can be leveraged to better understand what motivational and design features influence student decisions to seek help from a particular source. At an abstract level, Expectancy Value Theory for Help Sources states that whether or not a person decides to seek help from a particular source is determined by her expectation that help source will provide accessible help and perceptions that the help source will provide high quality help.

Expectations for help from a particular help source are based on beliefs about whether that source will be available to provide help, whether that source is accessible, and a basic belief that there will be obtainable help from that particular source. Values for a help source originate from whether that help source will provide the expected type of help such as the anticipated quality and accuracy. This model functions as an initial theoretical explanation for how students seek help from a particular resource.

Expectancy Value Theory for Help Sources also largely focuses on utility value for the values for the help source. This is likely due to the other, more intrinsically-related value types being less relevant in a help seeking context. Outside of EVT-HS, other motivating student values include attainment, intrinsic, and cost values. Attainment value was measured using surveys [15] including items such as, "Mathematics/English is important to me personally." Intrinsic value was measured by items such as, "I enjoy puzzling over mathematics/English problems." It seems unlikely that students would enjoy seeking help and consider attaining help as a personal value. However, Expectancy Value Theory for Help Sources could certainly have included cost items in their model of Expectancy Value Theory for Help Sources. While cost can be measured as an amount of time-required to achieve the task as in Trautwein et al. (2012) [16], it can also be measured as public and private threats to self-esteem, and inconveniencing a particular help source. While costs were not explicitly included in the original introduction of EVT-HS [3], these social costs are often included in higher-level models of expectancy and value motivational frameworks. Public threats to self-esteem or evaluation anxiety is important when considering the social nature of help seeking.

3.1 Costs of Help Seeking

One contribution of our work is an expansion of focus on costs in help seeking. Costs are typically considered one of many possible values inside an Expectancy Value Theory model. But there are also many different types of costs. Costly outcomes can include private threats to self-esteem (i.e., “If I ask for help, it means I’m not competent”), public threats to self-esteem (i.e., “If I ask for help the teacher will think I’m not competent”), face threatening acts (i.e., “It will inconvenience the teacher to help me”), among others [15]. Costs were not included in the Expectancy Value for Help Sources Theory explicitly, but certainly one help source could induce more costs than another (i.e., “If I ask the teacher for help, she will think I’m not competent, but if I ask for help from this discussion forum, strangers on the Internet might think I’m not competent”).

With regard to our investigation of a peer help exchange system in a discussion forum, social costs are of particular interest. When looking specifically at voting in reputation systems, one of the most apparent costs of help seeking is evaluation anxiety or the fear of being judged.

Evaluation anxiety, or a person’s concern about being evaluated [17], can be impacted by numerous contextual factors and is also similar to perceived public threats to self-esteem [18]. Both of these factors are related to impression management strategies to prevent others from perceiving one as incompetent. In this section, we focus on the effect of evaluation anxiety in learning contexts.

Learning often requires evaluation, either from others such as the teacher, or from within when self-monitoring one’s progress. Hence, the issue of anxiety around evaluation potential is relevant to learners. However, a review of the literature does not appear to reveal evaluation anxiety systematically studied with regards to its effects on help seeking. Evaluation anxiety is referred to in reference to its relationship with threats to public self-esteem, as in Nadler (1997) in which he “suggests that one avoids the seeking of help because of evaluation apprehension concerns” [19].

Additional investigations of evaluation anxiety look at impact on dominant responses depending upon the task difficulty and whether the evaluation being provided is presented as instrumental for future performances [20]. This work also began measuring participants’ perceived levels of evaluation anxiety through general anxiety measurements before and after the experimental task. Geen’s (1983) measure of evaluation anxiety consisted of the state form of the Spielberger State-Trait Anxiety Inventory [20]. Items in this part of the inventory include: “I am tense; I am worried” and “I feel calm; I feel secure.” Current methods to measure evaluation anxiety also look at the cause of the evaluation anxiety [21][2]. To measure experienced evaluation anxiety, a subset of items used to measure negative affect are used as a scale. These items include negative affects specifically related to anxiety: nervous, worried, calm, tense, and relaxed [21].

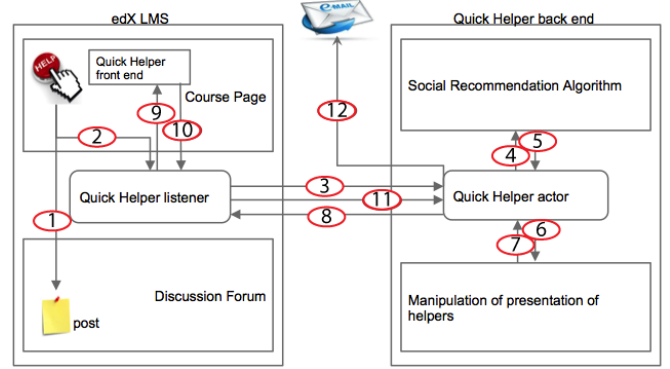


Fig. 2. The steps in the Quick Helper system for students to have potential peer helpers recommended to answer their question.

Howley et al. (2014) explored the impact of robots on evaluation anxiety and help seeking in a one-on-one tutoring setting [22]. Results showed that students learned significantly less from a human teacher as compared to a robot teacher, human helper, and robot helper. This significant reduction in learning was partially due to the fact that participants asked the human teacher marginally fewer questions. While students asked marginally fewer questions from human teachers than the human helper, the authors did not see the same distinction made for robot teachers and helpers. So, human teachers were hypothesized to increase evaluation anxiety more than human helpers and either of the robot conditions. This experiment shows that evaluation anxiety may very well have impact on help seeking, and that the fear of being judged can possibly be reduced through the intentional design and presentation of the technology enhanced learning environment.

4 QUICK HELPER EXCHANGE SYSTEM

As an expansion of the focus on costs in help seeking, this paper addresses the role evaluation anxiety and other constructs from Expectancy Value Theory for Help Sources plays in help seeking in educational contexts, through a MOOC help exchange system. We do this using our Quick Helper tool as an experimental infrastructure. Our experiment explores the design space specifying how to present and use the information returned by a social recommendation algorithm for help seeking. We experimentally explore this space using Quick Helper, implemented using a context-aware Matrix Factorization model to predict students’ preferences for involvement in discussion so that it would be possible to select potential help providers who would be likely to participate in a discussion related to the issue raised in the help request. In prior work, a similar approach was designed to recommend discussion threads to students [6], but only evaluated in corpus analysis experiments, not in a deployment study in real time. A real-time comparable recommendation algorithm had not been implemented in a live MOOC to operate in real time until the experiment reported in this paper.

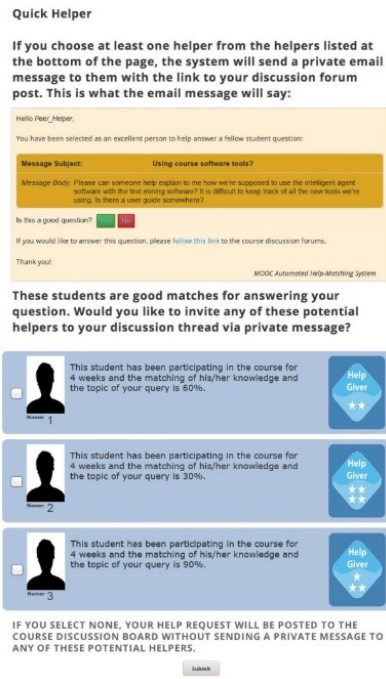


Fig. 3. A screenshot of the Helper Selection Window showing the preview email message on top and the helper selection screen below.

The detailed process of using Quick Helper is as follows, as shown in Fig. 2: *Step 1*. When the Quick Helper button is clicked, the action is logged and the student submits her question. The question is posted to the course discussion forum. *Step 2*. The Quick Helper component listens to the question-posting event and in *Step 3* it forwards this question (the title and body) and other details like help-seeker id and URL of the question posted on the discussion forum to our external Quick Helper server. *Step 4*. The information is passed to an intelligent social recommendation algorithm which in *Step 5* selects and recommends three potential helpers to answer the student's question. *Step 6-7*. A decision is then made about the presentation of these helpers. This is where the experimental manipulation of our study took place. *Step 8*. Information (i.e., the user name / user ID, profile image, user expertise description, and badges for user rating) about these recommended helpers is sent back to the Quick Helper client on the edX course page *Step 9*. A window is presented to the user showing this information. At the top is a sample private message that will be sent to the selected helpers (should the student select any), as shown in Fig. 3. Below the preview message are the helpers from which to select, as in Fig. 3. *Step 10-11*. The student selects 0-3 helpers. The selected helpers are passed on to the Quick Helper backend. *Step 12*. An email is sent to each of the selected helpers including a hyperlink to the forum thread that was posted in Step 1, inviting a response.

Whether or not the student selects a helper, Quick Helper always publishes the help request into the course discussion forum. Our intervention added the option to

explicitly invite helpers to the thread as a means of increasing likelihood of a response.

When the student submits her question through the Quick Helper, the matrix factorization algorithm uses the content of the question and metadata regarding the student's social connections and pattern of discussion involvement to select three appropriate peer helpers. It first maps a student's question to a similar question, and then estimates students' preferences for answering that question by taking into account features from students, questions and student connections similar to the approach in [6]. The algorithm also has the ability to consider load balancing issues to avoid any particularly capable student being overwhelmed by requests for help, although we did not use this feature in the current experiment.

5 EXPERIMENT 1: EMPIRICAL EVIDENCE FOR EVT-HS

Using this Quick Helper tool, when students seek help they are given the option to select up to three potential helpers to assist before posting their question to the course discussion board. We apply our EVT-HS lens when we display the three helpers to the student. Our three experimental dimensions consist of components of EVT-HS as well as evaluation anxiety.

We selected three features that are commonly part of reputation systems to examine how they impact student expectancies, values, and costs for a help source: badges, helper expertise information, and up/downvoting. These features are ambiguous in design as there is not necessarily a direct path between our design features and components of EVT-HS. This path is investigated at length in Experiment 2. While it is possible to hypothesize the valence of the features' effect on help seeking, it should be acknowledged that we might not be able to accurately predict the exact effect on EVT-HS beliefs these features may have.

These three common MOOC features provide us the ability to investigate how emphasizing components of Expectancy Value Theory for Help Sources and evaluation anxiety impact helper selection in a help exchange system, yielding the following hypotheses:

1. **Topic Match Sentences.** Placing an emphasis on the helpers' knowledge should raise the perceived expectancies or values of the help that helper can provide. This should be reflected by an increase in help seeking outcomes, as compared to the control sentences. (Marginally supported by our results).
 - a. An increasing number of weeks enrolled in the course should increase the number of helpers selected. (Not supported).
 - b. An increasing topic match percentage shown should increase the number of helpers selected (Supported, $X^2(1, N=171) = 8.86, p < 0.01 (R^2 = 0.05)$).

2. **Badges.** The presence of badges implying information about a peer's help giving should increase the likelihood students will seek help. In our system design, this increased likelihood will be reflected by a larger number of peers privately invited to view a public thread. (Partially supported as an interaction with voting in our final analysis).
 - a. An increasing number of badge stars should increase the number of helpers selected (Not supported).
3. **Voting.** Being evaluated via up/downvoting increases the cost of seeking help, yielding a reduction in help seeking outcomes. (Supported as an interaction with badges, $F(4, 129) = 4.07, p < 0.05, R^2 = 0.05 (\Delta R^2 = 0.025)$).

To investigate how these expectations, values, and costs influence help seeking in MOOCs, we performed a 2 (badges) X 2 (topic match sentences) X 2 (voting) factorial experiment in the context of MOOC discussion forums. Our experiment manipulates how potential helpers are presented to the help-seeking student. Number of helpers selected is the main help seeking outcome we investigated.

The common reputation system features employed in our experimental manipulation were adapted for our Quick Helper help exchange system for MOOC discussion forums. When students used our Quick Helper system, they were presented with three potential peer helpers to invite to answer their question in the course discussion forum. We emphasized the different components of our model through methods currently employed in MOOCs and other online learning systems. Without an associated survey instrument, we could only hypothesize about whether the manipulations had a positive or negative impact on helper selection in our help exchange system.

A "Help Giver" badge system with one to four stars, one of which is shown in Fig. 4, should have a positive impact on help seeking. If the help-seeking instance was not assigned to the badge condition, potential helpers were displayed without a badge. The number of stars on the help giver badge is determined by rank ordering the three potential helpers, although we provided no explicit explanation of the stars' meaning to students. We based these badges on the visual appearance of the OLDS MOOC badges [12], but our Help Giver badges were displayed within our Quick Helper system and not rewarded to students for display on personal pages or posts. In this way our badges were not applied in their typical way as motivational and extrinsic rewards, merely as an informational tool to help seekers. By providing explicit insight into the potential helpers' knowledge, we should also elicit a positive impact on helper selection in our help exchange system. In this way, the student could evaluate the potential helpers' ability to provide accurate help. The sentence displayed was "This student has been participating in the course for <#> weeks and the matching of his/her knowledge and the topic of your query is <#>%." The

numbers were provided by the system, but no further explanation of their meaning was provided to students. If not assigned to the topic match sentence condition, students were shown control sentences about their potential helpers from the following four:

1. "This colleague has a computer and is ready to go."
2. "This colleague is involved in the course."
3. "This colleague answers email on a regular basis."
4. "This colleague uses Web 2.0 technologies."

Fig. 4 shows two examples where the top image displays the positive manipulations and the bottom image exemplifies the control counterparts.

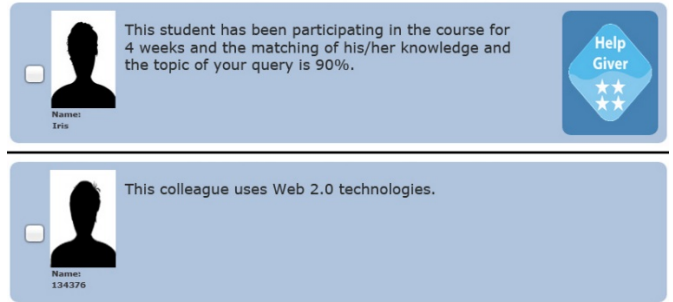


Fig. 4. The top image shows badge, topic match sentences, and username conditions, while the bottom image exemplifies the control version of these manipulations.

We emphasized a potential cost of seeking help by displaying to the help seeker a preview of the email selected helpers would receive. Help seekers could see from this preview email message that their selected helpers will be invited to evaluate whether the student's question was good. We did this through an exaggerated up/downvoting interactional archetype using buttons commonly used in MOOCs [11] and other help request discussion forums such as StackOverflow. Our implementation of up/downvoting is shown in Fig. 5. Knowing that one's post will explicitly be evaluated by any selected helpers should also increase public threats to self-esteem, thereby emphasizing the costs of selecting helpers. In the non-voting condition, the preview email message did not have the "Is this a good question?" with green and red buttons following.

We also manipulated whether or not students saw their potential helper's usernames. Students' selected usernames are most commonly displayed in discussion forums. However, knowing your helpers' names may impact perceived expectations and values about their help-giving abilities. And so, we included helper anonymity as a fourth dimension in our experiment, so that we might explore how EVT-HS lives in both a real world setting, as well as in a slightly more controlled experimental setting. However, our analyses showed no effect of this manipulation, therefore we drop it from further discussion.

5.1 Qualitative Results from Deployment Study

Students in a learning analytics course hosted by edX had the option to post their questions directly to the course discussion forums, or to use our “Quick Helper.” Using the Quick Helper would still post the question to the public discussion forum, but also privately invite selected peers to view the thread’s URL. When the student’s question is posted to the course discussion forum, the student is randomly assigned to one of our intervention conditions.

Throughout the learning analytics MOOC, approximately 20,000 individuals were enrolled, of these, 6,240 students were considered “active” as they clicked a course video at least once. After the initial 3 weeks, no more than 2,493 students were active in a given week. 285 MOOC students posted a total of 671 threads to the discussion forum throughout the course and one third of these students used the Quick Helper at least once (i.e., 96 users). Number of Quick Helper posts increased over time, relative to non-Quick Helper forum posts.

In the initial two weeks of the course when data on students was lacking, we used a TA Version of Quick Helper. Teaching Assistants were volunteers recruited by the MOOC instructors. The TA Version was different from the Student Version in the following ways: 1) badges always showed four stars for the TAs, 2) the topic match sentence was always “This is one of the Teaching Assistants selected for this course. All of our Teaching Assistants are highly qualified to answer student queries”, 3) The TAs’ usernames were always shown. Our analyses controlled for differences in the TA and student version of the MOOC.

We had numerous successful cases in which a student used the Quick Helper, invited three potential helpers to her forum post, and one of those invited peers responded, such as in the example below:

Student151: I don't remember being able to participate in a hangout. In fact all I got was George in a parking lot and then some guy talking about data.

Student 102: Hi Student151 The hangout you are referring to was the TONY HIRST HANGOUT from Week 2. I will usually get an email sent to me informing me of the date and time of the upcoming hangouts so if I want to participate, I will know when they are happening...

However, in doing a more thorough step-by-step analysis,

TABLE 1 NUMBER OF HELPERS SELECTED	
<i>Total Number of Helpers Selected per QH Instance</i>	
0 selected	98
1 selected	23
2 selected	7
3 selected	30
<i>Helpers Selected Overall</i>	
Yes (Selected)	127
No (Not Invited)	347

The top table shows how many helpers were selected per Quick Helper instance (dependent variable for the binary independent variables). The bottom table is the proportion of helpers that were selected overall (sub-level variables).

we realized that the Quick Helper system was often inviting potential peers who may have become inactive, although they had been active in the past. This suggests that going forward our algorithm needs to incorporate students’ last active date and a threshold for recent inactivity as a feature in the social recommendation algorithm.

Our analysis also revealed a few Quick Helper instances in which the student was not seeking help, but was perhaps using the Quick Helper system as more convenient access to the discussion forums. This suggests that as a new MOOC feature, students are still developing an understanding of the purpose and benefits of the Quick Helper. With more widespread use the learning curve for using the system might be reduced.

5.2 Results of Experimental Manipulation in Deployment Study

Our dataset for testing our hypotheses includes 158 of the Quick Helper instances by 66 users, who selected a mean of 0.79 helpers ($\sigma = 1.17$). Participants were randomly assigned to one level in each of the three factors of our experiment: 50% of Quick Helper instances were assigned to the Badges condition (50% to the No Badges condition), 50% to the Topic Match condition (50% to control sentence), and 46% to the Voting condition (54% to No Voting). Prior to our analysis, we removed overly brief Quick Helper posts,

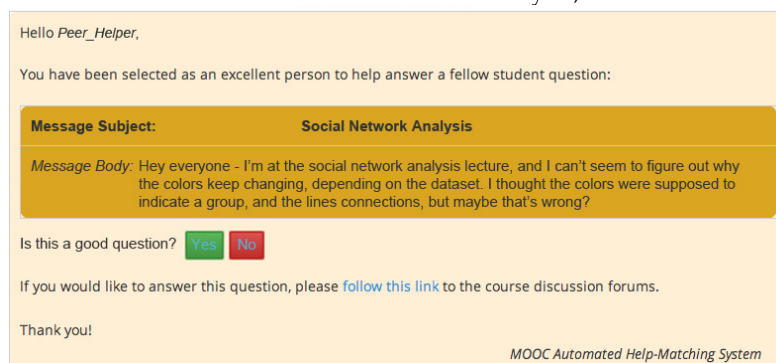


Fig. 5. A screenshot of the Preview email message, showing the voting manipulation (i.e., the green and red boxes).

as well as data points with a timestamp occurring after the course had officially ended.

We have two dependent variables at two levels of analysis. Our main dependent variable, 'Number of Helpers Selected', is at the Quick Helper instance level. We can use our binary badges, voting, and sentence conditions to predict number of helpers selected. Within the badges and sentence conditions, we have sub-level independent variables. These sub-level variables include the number of stars shown on the badge as well as the number of weeks enrolled and topic match percentages. These independent variables are at the helper level and 'Helper Was Selected' is the relevant dependent variable. There were three helpers shown per Quick Helper instance, so it is not possible to investigate individual helper sentence level variables with respect to instance level variables (i.e., three different sentences were displayed at once). The proportion of helpers selected with our Quick Helper system is shown in Table 1. Version (i.e., student or TA) is maintained as a covariate throughout the analyses, and post-hoc analyses are performed as Student's t-tests.

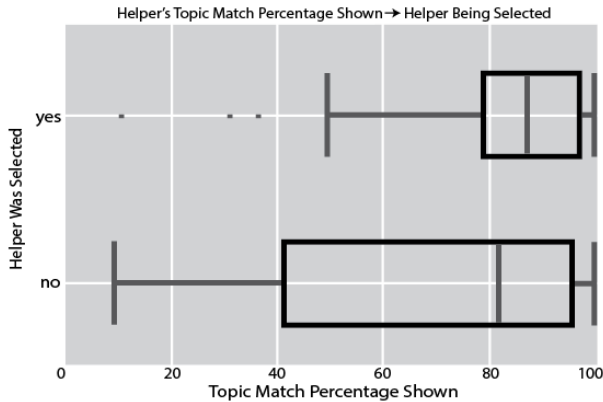


Fig. 6. A comparison of the topic match percentages shown next to helpers, by whether or not the help-seeker selected them.

Hypothesis 1 – Topic Match Sentence Hypothesis.

An ANCOVA analysis, controlling for version, showed that the topic match sentence condition had a marginal effect on number of helpers invited to the question thread, $F(2, 149) = 3.38$, $p < 0.07$ (Cohen's $d=0.21$). A Student's t-test post-hoc analysis revealed that students in the topic match sentence condition selected marginally more helpers to be invited to their help request thread. This marginal result follows the predictions of Hypothesis 1, although is described further by analysis of Hypothesis 1a.

We also investigated how the information displayed in each condition impacted whether a helper was selected (Hypothesis 1a, 1b). Only instances where the value emphasis sentence was shown were included in this analysis. Topic match percentage shown in the value emphasis condition had a significant effect on whether the helper was

selected, $X(1, N=171) = 8.86$, $p < 0.01$ ($R^2 = 0.05$), as shown in Fig. 6. The number of weeks the helper participated in the course did not appear to have an effect on students' choices of helpers.

From Fig. 6 we can see that helpers with topic match percentages displayed under approximately 80% were considerably less likely to be invited to the question thread. This suggests that lower topic match percentages are not emphasizing a potential helper's quality of help, but rather a lack of quality. This 80% threshold is the likely reason why the effect of topic match sentences on number of helpers selected was marginal and not significant. The majority of the helpers shown had a topic match above 70%, which is why we maintain some effect of topic match sentence.

Hypotheses 2 & 3 – Badges and Voting Hypotheses

Further investigations of our usage of badges and voting revealed no statistically significant relationship, until we look at the interactions between our conditions. There was a significant interaction between badges and voting, $F(4, 129) = 4.07$, $p < 0.05$, $R^2 = 0.05$ ($\Delta R^2 = 0.025$), with a post-hoc analysis revealing that voting only appears to have an effect when no badges are present. A Student's t-test (and Fig. 7) shows that in the absence of badges, significantly more helpers are selected in the non-voting condition. This interaction supports Hypothesis 3 which predicts a negative relationship between increasing the cost of help seeking, and the number of helpers selected. Hypothesis 2 is also supported as part of this interaction which also introduces the potential of using help giver badges to alleviate the negative effects brought about by the use of up/down-voting. The most number of helpers were invited on threads when helpers were shown without voting and without badges, although this difference is not significantly better than conditions where badges are present.

In investigating how the number of stars on the Help Giver badges related to whether a helper was selected, we did not find a significant relationship, although the trend was in the expected direction: more badge stars shown increased the likelihood of the helper being selected.

Limitations

Our Quick Helper system was designed to test commonly used MOOC discussion forum features for their positive or negative impact on student help seeking behaviors. Some of our manipulations are particularly exaggerated and may not represent the design of reputation systems in more standard environments. Furthermore, our dependent variables (i.e., helper selection) were obtained immediately after exposure to the experimental manipulations. This provided ideal control over environmental variables, but 'helper selection' may not be a valid dependent variable in all MOOC discussion forum systems.

While our MOOC had 6,240 users who clicked a course video at least once, 4.5% of these users posted to the dis-

cussion forum at least once. In other MOOCs we have previously observed higher discussion forum engagement. Therefore, it is unclear to what extent these exact results generalize to other MOOCs. However, one third of our discussion forum users interacted with Quick Helper, which suggests among users of MOOC discussion forums Quick Helper may be a desirable tool. Furthermore, our Quick Helper results align with the results from Coatsee et al. [10], which is a different MOOC forum context from a different research group, suggesting some generalizability of our results on evaluation anxiety and up/downvoting.

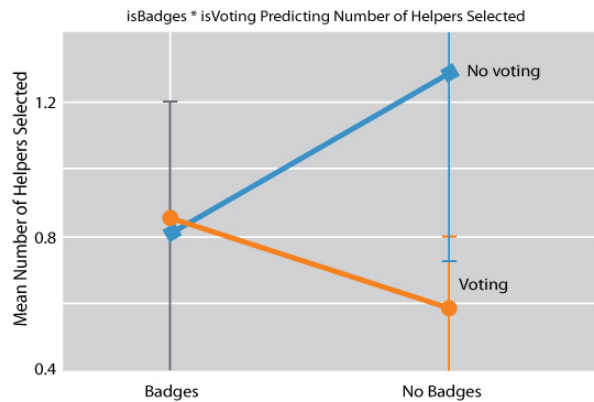


Fig. 7. An interaction plot between Help Giver Badges and Up/Downvoting, showing badges eliminating the negative effect of voting.

Additionally, this experiment was designed for a much larger sample size, but due to Quick Helper's novelty we did not see as much use of our system as anticipated. Due to this, we have limited statistical power to draw reliable conclusions about external validity. It might be informative for improving the help exchange system for a more in-depth analysis to examine why students clicked the Quick Helper, but then did not invite any helpers to their thread. It is possible that the Quick Helper interface was overwhelming or that inviting helpers was too effortful. It is also possible that students viewed the Quick Helper as a shortcut to using the discussion forum, and not as limited to purposes of help seeking.

Discussion

Fig. 8 shows a visualization of the relationships between variables hypothesized previously. Our results suggest that two commonly applied features of reputation systems have a complex relationship with student help seeking. That is, without badges providing information about whether the help source is a help giver, up/downvoting facilities may have negative effects. The presence of the badges alleviates the potential harmful effects to public self-esteem, resulting in students inviting more helpers to their discussion threads. This suggests that our Help Giver badges and up/downvoting mechanism might be influencing the same student beliefs as part of the EVT-HS model. Designers of MOOCs and SPOCs (Small Private

Online Course) need to be mindful of which features they decide to deploy in their course, and how those decisions impact student help seeking.

The marginal effect of value emphasis on number of helpers selected supports our first hypothesis. Knowing that help sought will be high quality increases the number of helpers a student invites to her forum thread. Information about peer expertise may be important to share with help seekers. Designers of online courses may want to consider how they present the expertise of their potential helpers. In our case, a knowledge topic match below 80% had a negative effect on the number of helpers invited.

6 EXPERIMENT 2: SURVEY SUPPORT FOR EVT-HS

While our field experiment showed relationships between help seeking behavior and interface elements, it does not illuminate the connection between the interface elements and the psychological variables referred to in the theory. It is currently unknown if Expectancy Value Theory for Help Sources explains the causal mechanism behind the behavior observed in Experiment 1. Therefore, Experiment 2 is designed to determine if reputation system features can be mapped to Expectancy Value Theory for Help Sources through self-report surveys.

Understanding the impact that common reputation system features have on expectancies, values, and costs for help sources is a novel line of research, but this survey also provides us with the opportunity to explore our irrelevant/control sentences to ensure that they did not have unintended effects on student perceptions. The more general features such as Help Giver Badges, voting, and helper expertise contribute to our understanding of EVT-HS in practice. The specific investigation of the wording of irrelevant/control sentences may function more as a manipulation check in this particular survey experiment.

6.1 Study Design and Methodology

The purpose of this experiment is to serve as initial empirical evidence for relationships between commonly used reputation system features and expectancy for the help source, values for the help source, and help seeking outcomes. This experiment provides support for EVT-HS as a causal mechanism for the interaction between badge and voting conditions in Experiment 1. We measured participant perceptions of expectancies, values, evaluation anxiety, and intention to seek help in response to reputation system features from the Quick Helper field experiment.

Recruited participants were shown screenshots from the Quick Helper system and for each screenshot their perceptions were measured via self-report.

Dependent Measures

To measure participant perceptions of expectancies, values, evaluation anxiety, and help seeking outcomes we employed a series of survey instruments. The dependent measures and survey items employed were: Expectancies

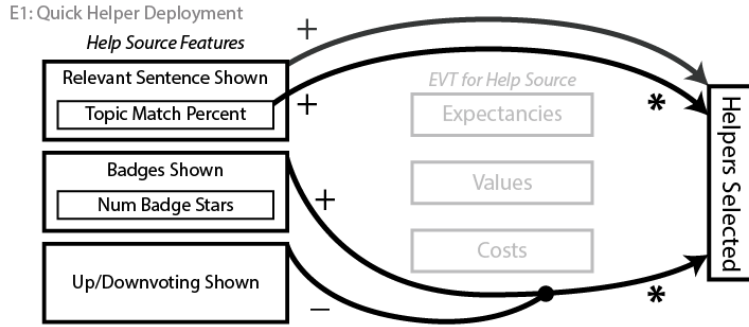


Fig. 8. Statistically significant relationships between variables in the Quick Helper MOOC Experiment. Positive independent variables had a positive effect on helpers being invited to the thread, while voting had the hypothesized negative effect on helper selection. An interaction between badges and voting is represented by connecting lines.

and Values of Help Sources inspired by Makara & Karabenick [3], Intention to Seek and Avoid Help from this help source, adapted from Wolters et al. [15] and Costs of Seeking Help in a Particular Context: Evaluation Anxiety measures from Leary et al. [21]. Measures for Intention to Seek and Avoid Help and Evaluation Anxiety were pulled from their original sources and mildly adapted to fit the nature of the experimental set up. However, measures of EVT-HS were previously nonexistent and as such, we constructed six Likert scale items:

Expectancy

- 1) <This person> is available to give me help.
- 2) If I ask for help from <this person>, they will give me help.
- 3) If I have a question for <this person>, they will answer me.

Value

- 4) The help from <this person> will be what I need to answer my question.
- 5) <This person> will provide me answers of high quality.
- 6) <This person> can give me accurate help.

Independent Measures

Experiment 2 has independent variables derived from the Experiment 1 field experiment. Participants were shown a variety of Helper Screenshots adapted from the Quick Helper Help Exchange system. These screenshots were manipulated along several dimensions:

- No Badges or Badges (with 1, 3, or 4 stars)
- Irrelevant/Control Sentences (4 possible) or Expertise Topic Match Sentences (4 weeks participation, and 30%, 60%, or 90% topic match)

Voting or No Voting (manipulated separately from the above two dimensions) The helper screenshots from this experiment looked similar to the samples shown in Fig. 4. The sentences displayed included:

- Irrelevant/Control Sentences: 1) "This colleague has a computer and is ready to go.", 2) "This colleague is involved in the course.", 3) "This colleague answers email on a regular basis.", and 4)

"This colleague uses Web 2.0 technologies."

- Relevant Topic Match Sentences: "This student has been participating in the course for 4 weeks and the matching of his/her knowledge and the topic of your query is <#>%." where <#>% is either 30%, 60%, or 90%.
- TA Sentence: "This is one of the Teaching Assistants selected for this course. All of our Teaching Assistants are highly qualified to answer student queries."

There were over twenty different versions of the screenshot. 54 participants were recruited from a private university's participant pool, as they share common age and educational levels with students in a MOOC. Each participant saw three different screenshots (except for one participant who completed only one question). An additional ten participants viewed four different sentence screenshots each. As such, this is a within-subjects experimental design. 30% topic match sentences occurred 20 times, 60% topic match had 13 instances, 90% had 27 instances, the TA sentence had 26, the Web 2.0 irrelevant/control sentence had 31, "involved in the course" had 26, "answers email" had 36, and "has a computer" had 41 instances.

6.2 Research Hypotheses

Experiment 1 suggested that badges and increasing expertise have a positive effect on help seeking, which could be achieved through many paths: increasing expectancies and values, or decreasing costs. Based on Expectancy Value Theory of Help Sources and the experimental manipulations from Experiment 1, we generated the hypotheses below. These hypotheses linking manipulation to belief are tentative as badges and sentences were adapted from commonly used reputation system features and not necessarily designed to target EVT-HS specifically:

1. Badges. Help Giver Badges might increase student expectations that there will be help, and so:

- a. The presence of badges will increase perceived expectancies for the help source. The presence of badges may also increase self-reported intentions to seek help from that source. (Unsupported)

- b. An increasing number of stars on the badge will result in an increase in the perceived expectancies for that help source. More badge stars should also result in more self-reported intentions to seek help from that help source. (Unsupported, except for a significant negative effect on evaluation anxiety, $F(1,81)=8.19$, $p=0.005$, $R=0.83$)

2. Topic Match Sentences. The topic match sentences from Quick Helper might increase student's perceived value of help from that help source. As the expertise expressed in the topic match sentences might manipulate values for the help source, prior work suggests the sentences will also manipulate expectancies.

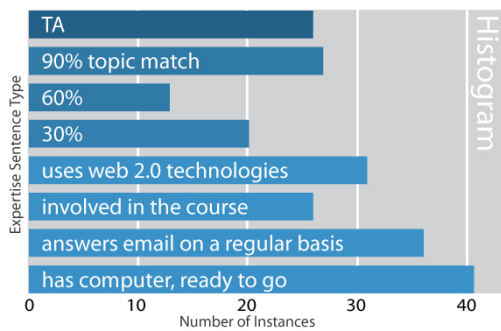


Fig. 9. The number of each sentence type shown to participants.

- Teaching Assistant and Topic Match sentences will predict significantly higher value for the help source than Irrelevant/Control Sentences. (Supported $F(2,172)=16.08$, $p<.0001$, $R=0.68$)
- Teaching Assistant and Topic Match sentences will predict significantly higher expectancies for the help source than the control sentences. (Supported, $F(2,170)=10.91$, $p<.0001$, $R=0.68$)
- Teaching Assistant and Topic Match sentences will predict significantly higher Intention to Seek Help than the control sentences. (Supported, $F(2,176)=11.76$, $p<.0001$, $R=0.64$)
- Increasing Topic Match Percentages in the Topic Match Sentences will result in significantly higher value for the help source (Supported, $F(1,34)=17.26$, $p=0.0002$, $R=0.91$) as well as significantly higher Intention to Seek Help (Supported, $F(1,51)=6.3$,

$p=0.02$, $R=0.66$). Increasing Topic Match Percentage may also result in increased expectancy beliefs for the help source (Supported, $F(1,27)=9.56$, $p=.05$, $R=0.94$)

3. Expectancy Value Theory for Help Sources. Standard hypotheses relating Expectancy Value Theory for Help Sources to self-reported Intention to Seek Help and Intention to Avoid Help:

- Self-reported expectancies will significantly predict intentions to seek help from the shown help source. (Supported, $F(1,195)=297.00$, $p<.0001$, $R=0.75$)
- Self-reported values will significantly predict intentions to seek help from the shown help source. (Supported, $F(1,190)=441.66$, $p<.0001$, $R=0.85$)
- Self-reported costs (i.e., evaluation anxiety) will significantly [negatively] predict intentions to seek help from the shown help source. (Supported, $F(1,222)=28.94$, $p<0.0001$, $R=0.66$)

4. Irrelevant/Control Sentence Hypotheses. Hypotheses related to the four different irrelevant sentences.

- TA and 90% topic match sentences will predict significantly more perceived values for the help source, followed by the irrelevant/control sentences followed by 60% and 30% topic match sentences (due to how they emphasized a lack of value in the Quick Helper experiment). (Supported, $F(7,168)=10.55$, $p<.0001$, $R=0.74$)
- Sentence Type may have a significant effect on Intention to Seek Help, in the ordering described in the previous hypothesis (Supported, $F(7,170)=9.13$, $p<.0001$, $R=0.71$)
- Sentence Type may have a significant effect on expectancies for the help source, in the ordering described by Hypothesis 4a (Supported, although the "email" sentence performs significantly better than the 90% topic match sentence on expectancies, $F(7,166)=7.10$, $p<.0001$, $R=0.73$)

6.3 Statistical Approach

The data's structure reflected its many dimensions. There were three separate columns for the binary-level variables (i.e., *isBadges*, *isTopicMatchSentence*, *isVoting*), and a separate column for the type of sentence that is listed

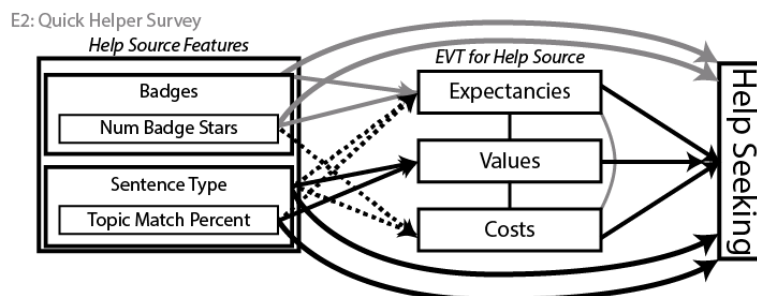


Fig. 10. The hypotheses results model for Experiment 2. Solid black lines indicate supported hypothesized relationships, solid grey lines indicate unsupported hypotheses, and dotted black lines indicate unhypothesized relationships

in Fig. 9. Although, the `isTopicMatchSentence` condition technically had three levels: `isTopicMatchSentence`, `isControlSentence`, and `TAsentence`. 73 badges with 1 star were shown, while 74 badges with 4 stars were shown, and 93 screenshots shown did not have any badge. Topic match percentage and number of badge stars were treated as continuous variables.

All analyses connecting categorical experimental manipulations to numerical beliefs scales were performed as an ANOVA with `RespondentID` as a random effect to account for the within-subjects experimental design. Analyses connecting the theory beliefs scales to intention to seek help were performed as a linear regression with `RespondentID` as a random effect as well. Post-hoc analyses were performed via Student's t-tests.

6.4 Results

Results can be seen in the hypotheses results model in Fig. 10, and are further explained next to the individual hypotheses described in section 6.2. Overall, we see the majority of our hypotheses supported.

Badges did not have the hypothesized effect on expectancies, but number of badge stars did have a significant inverse relationship with our measures of evaluation anxiety. The value manipulation, topic match sentence condition, had a significant relationship with value beliefs and expectancy beliefs, as anticipated. The relationship between expectancies, values, and evaluation anxiety with intention to seek help was once again repeated in the hypothesized directions.

The irrelevant/control sentence hypotheses can be further explored as a manipulation check to ensure that the control sentences were neutral. These hypotheses generally predicted that the TA and 90% topic match sentences would perform better on the positive dependent variables (expectancies and values for the help source, intention to seek help) than the control sentences, with the 60% and 30% topic match sentences following at the end. This relationship generally held through all three of the positive outcomes, except the "This colleague answers email on a regular basis" sentence performs consistently higher than all the other control sentences. For the expectancies for the help source, this email sentence actually performs statistically significantly better than the 90% topic match sentence and statistically indistinguishable from the TA sentence.

6.5 Discussion

The results of this survey experiment are three-fold:

- 1) Help Giver badges do not manipulate expectancy for the help source beliefs, but they might manipulate evaluation anxiety. This result better explains the interaction between badges and voting we saw in Experiment 1. Quite possibly, the negative effect of voting was reduced due to a direct manipulation of evaluation anxiety via the badges, although we did not see any relationship between number of badge stars shown and whether or not the helper was

selected in the Quick Helper Experiment 1. Further exploration of why a Help Giver badge reduces evaluation anxiety is necessary. Perhaps when participants see a helper listed as a "Help Giver" they assume that this peer is more altruistic and less likely to evaluate their help seeking.

- 2) The value manipulation manipulated both value and expectancy beliefs for the help source. Expectancy and value beliefs for the help source may be difficult to manipulate separately, although increasing both expectancies and values has a positive effect on help seeking.

- 3) One of the irrelevant/control sentences did not function as a control sentence for expectancy beliefs for the help source. While the "email" sentence performed better than expected for one outcome (expectancy beliefs for the help source), overall the control sentences performed at the hypothesized levels. Any information provided has the potential to impact student beliefs for the help source, and so from a theoretical standpoint minimizing that effect for control sentences is important.

6.6 Limitations

The nature of introducing the hypothetical context might introduce validity questions which are the case for the up/downvoting manipulation. Additionally, the survey relies on the specifically selected design features of the Quick Helper system which may not be valid outside of our helper recommendation system, although this concern is most relevant for the design of the irrelevant/control sentences which serve a theory-purpose in the experiments in this chapter.

Currently missing from this section is an analysis of the connection of our cost manipulation with evaluation anxiety (i.e., cost beliefs). Hypotheses would have predicted a positive relationship between up/downvoting and evaluation anxiety, but it was not possible to implement the Quick Helper preview email message screenshots in an interpretable format for survey participants. Pilot testing revealed a general failure of participants to 1) read the content of the messages included in the email preview screenshots, 2) understand that the screenshot was a preview for a peer helper that was not the participant, and 3) realize that the peer helpers were requested to evaluate the quality of the question. Much of this confusion was likely due to the question in the email preview screenshot being arbitrary and not specifically written by the survey participant, as was the case in the actual Quick Helper system. However, due to the explicit evaluative nature of up/downvoting and the results from Experiment #1, our confidence in up/downvoting manipulating evaluation anxiety is relatively high. Evidence to empirically link voting to self-reported perceptions of evaluation anxiety is still needed.

7 CONCLUSIONS

Portions of Expectancy Value Theory for Help Sources can be used to explain actual student behavior in a help exchange system. However, as shown in Experiment #2, the

survey experiment, Value Beliefs are difficult to manipulate separately from Expectancy Beliefs for the help source. However, the expected relationship between EVT-HS beliefs and help seeking outcomes was observed in our analyses. From a practical perspective, aiming for a positive effect on expectancies or values, and a negative effect on evaluation anxiety should result in increased help seeking. It may be difficult to design manipulations to specifically impact only one EVT-HS belief independently.

Fig. 11 shows a synthesis of the results from the MOOC field experiment (Experiment #1) and the survey experiment (Experiment #2). We see that our hypotheses were essentially supported: expectancies and values can be increased through the use of topic match sentences, and perceived costs for a help source can be impacted by Help Giver Badges and up/downvoting. The figure also shows that topic match sentences, due to their ambiguous design, influenced expectancies, values, and costs in consistent directions. An interaction between Help Giver Badges and up/downvoting is two separate and opposing forces on perceived costs of seeking help from the help source.

The Quick Helper MOOC experiment and associated survey experiment have generated some design recommendations. Specifically, up/downvoting impacts the helper selection process negatively, but this can be mitigated through the use of Help Giver badges which reduce self-reported evaluation anxiety. If a reputation system's representation for student expertise is not high, then students will be unlikely to seek help from that person.

Design Recommendations

In pursuing an understanding of how Expectancy Value Theory for Help Sources relates to student behavior, reputation system features, and evaluation anxiety, we have discovered two key design recommendations for improving online courses. Design recommendations for course instructors include: 1) reducing student evaluation anxiety to increase help seeking and 2) emphasizing a peer helper's potential value and expectancy to increase help seeking.

These design recommendations originated from results

showing that up/downvoting forum interactional archetypes increase student evaluation anxiety, and likely should not be used in educational contexts if the voting is not being leveraged to organize large quantities of content. Reducing the options of the up/downvoting was one observed way to decrease evaluation anxiety. Implementing Help Giver badges is another path to achieving this goal.

When emphasizing a peer helper's expertise, it should be noted that displaying expertise levels below "above average" will have a negative impact on how often that student is invited to help. In order for a display of expertise to positively impact expectancies and values, that expertise must exceed a threshold of perceived competence.

Future Work

In this paper, we showed that the Quick Helper system is useful as a theory proving ground, and also potentially as a way to connect more students to the help that they need. While this paper focuses primarily on motivations for selecting peer helpers in Quick Helper, it did not address whether help seekers received the desired help. According to the EVT-HS theory, experiences gained using Quick Helper previously could influence perceived expectancy of help being available or useful the next time. If previous experiences with Quick Helper were successful, this might motivate future use of the tool. Investigating help received and how it influences longterm usage of Quick Helper is a clear next step in this line of inquiry.

We investigated in great depth the first few steps of the help seeking process, but the help giving portion of the process remains as yet unexplored. A better understanding of whether the peer helpers from the Quick Helper system actually answer the student in need is necessary, as well as better approaches for encouraging the peer helpers to provide help when requested. Redesigning the email messages used to invite helpers are one possible path to encouraging more help giving. Future work includes examination of the help giving experience in Quick Helper.

Ensuring generalizability of our results on EVT-HS and reputation systems requires a second experiment. This experiment would go beyond Coatzee et al. [11] and include

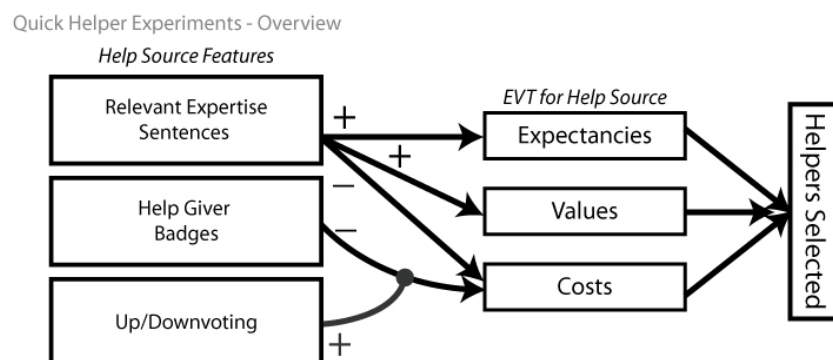


Fig. 11. A synthesis of the [field] Experiment #1 and [survey] Experiment #2 results. Note that the connection between Up/Downvoting and EVT-HS costs was unable to be empirically examined in the survey experiment.

follow-up surveys on student expectatations and values for help seeking in the forum. The experiment would also include a wider user base, and better instrumentation for measuring the success of Quick Helper interactions.

Overall, this article serves as a proof-of-concept for Quick Helper as a potential means for connecting students to helpers. More work is necessary to refine the help exchange process and evaluate its effectiveness in a larger user base. This article also shows Quick Helper's effectiveness as a tool for answering research questions about student behaviors and motivations for help seeking.

Acknowledgment

This work is supported in part by Carnegie Mellon University's PIER Program funded by grant R305B090023 from the US Department of Education.

References

- [1] J. Magnusson and R.P. Perry. "Academic help-seeking in the university setting: The effects of motivational set, attributional style, and help source characteristics." *Research in Higher Education* 33, no. 2 (1992): 227-245.
- [2] P.L. Bagley. "Evaluation Apprehension: An Examination of Affect in the Audit Environment." PhD diss., Doctoral Dissertation, University of Georgia, 2007.
- [3] K.A. Makara and S.A. Karabenick. "Characterizing sources of academic help in the age of expanding educational technology: A new conceptual framework." *Advances in help seeking research and applications: The role of information and communication technologies* (2013): 37-72.
- [4] J. Mackness, S. Mak, and R. Williams. "The ideals and reality of participating in a MOOC." In *Proceedings of the 7th International Conference on Networked Learning* 2010. University of Lancaster, 2010.
- [5] S. Mak, R. Williams, and J. Mackness. "Blogs and forums as communication and learning tools in a MOOC." In *Proceedings of the 7th International Conference on Networked Learning* 2010. University of Lancaster, 2010. D. Yang, M. Wen, I. Howley, R. Kraut, and C. Rose. "Exploring the effect of confusion in discussion forums of massive open online courses." In *Proceedings of the Second (2015) ACM Conference on Learning@ Scale*, pp. 121-130. ACM, 2015.
- [6] D. Yang, M. Pierrgallini, I. Howley, and C. Rose. "Forum thread recommendation for massive open online courses." In *Proc. of the 7th International Conference on Educational Data Mining* 2014.
- [7] S. Bull, J. Greer, G. McCalla, L. Kettel, and J. Bowes. "User modelling in i-help: What, why, when and how." In *International Conference on User Modeling*, pp. 117-126. Springer Berlin Heidelberg, 2001.
- [8] J. Greer, G. McCalla, J. Cooke, J. Collins, V. Kumar, A. Bishop, and J. Vassileva. "The intelligent helpdesk: Supporting peer-help in a university course." In *Intelligent tutoring systems*, pp. 494-503. Springer Berlin/Heidelberg, 1998.
- [9] P. Resnick, K. Kuwabara, R. Zeckhauser, and E. Friedman. "Reputation systems." *Communications of the ACM* 43, no. 12 (2000): 45-48.
- [10] D. Coetzee, A. Fox, M.A. Hearst, and B. Hartmann. "Should your MOOC forum use a reputation system?." In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pp. 1176-1187. ACM, 2014.
- [11] S. Cross. "Evaluation of the OLDS MOOC curriculum design course: participant perspectives, expectations and experiences." OLDS MOOC Project Internal Report, Milton Keynes. 2013.
- [12] J.A. Laso, P. Pernías Peco, and S. Luján-Mora. "Using Open Badges as certification in a MOOC." *ICERI2013 Proceedings* (2013): 1809-1818.
- [13] S. Nelson-Le Gall. "Help-seeking: An understudied problem-solving skill in children." *Developmental Review* 1, no. 3 (1981): 224-246.
- [14] C.A. Wolters, P.R. Pintrich, and S.A. Karabenick. "Assessing academic self-regulated learning." In *What Do Children Need to Flourish?*, pp. 251-270. Springer US, 2005.
- [15] U. Trautwein, H.W. Marsh, B. Nagengast, O. Lüdtke, G. Nagy, and K. Jonkmann. "Probing for the multiplicative term in modern expectancy-value theory: A latent interaction modeling study." *Journal of Educational Psychology* 104, no. 3 (2012): 763.
- [16] B. Guerin. "Mere presence effects in humans: A review." *Journal of experimental social psychology* 22, no. 1 (1986): 38-77.
- [17] S.A. Karabenick. "Seeking help in large college classes: A person-centered approach." *Contemporary educational psychology* 28, no. 1 (2003): 37-58.
- [18] A. Nadler. "Personality and help seeking." In *Sourcebook of social support and personality*, pp. 379-407. Springer US, 1997.
- [19] R.G. Geen. "Evaluation apprehension and the social facilitation/inhibition of learning." *Motivation and Emotion* 7, no. 2 (1983): 203-212.
- [20] M.R. Leary, B.D. Barnes, and C. Griebel. "Cognitive, affective, and attributional effects of potential threats to self-esteem." *Journal of Social and Clinical Psychology* 4, no. 4 (1986): 461.
- [21] I. Howley, T. Kanda, K. Hayashi, and C. Rosé. "Effects of social presence and social role on help-seeking and learning." In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pp. 415-422. ACM, 2014.

Iris Howley is an Assistant Professor of Computer Science at Williams College. Her current research interests lie in the overlap between the learning sciences, human-computer interaction, and social psychology research.

Gaurav Tomar is a research assistant at the Language Technologies Institute of Carnegie Mellon University. His research is deeply interdisciplinary and situated at the intersection of computer science, language technologies, learning science and technologies.

Oliver Ferschke is a postdoctoral researcher at the Language Technologies Institute at Carnegie Mellon University. He studies collaboration at scale and collaboration in online communities through the lens of language and computational linguistics.

Carolyn P. Rosé is a professor with the Language Technologies Institute and the Human-Computer Interaction Institute at Carnegie Mellon University. Her research spans the fields of language technologies, learning sciences, and human-computer interaction.