What Makes Creators Engage with Online Critiques? Understanding the Role of Artifacts' Creation Stage, Characteristics of Community Comments, and their Interactions

Qingyu Guo
qguoag@connect.ust.hk
The Hong Kong University of Science and Technology
Hong Kong, China

Chao Zhang
zhangchaodesign@zju.edu.cn
Zhejiang University
Hangzhou, China

Hanfang Lyu
hanfang.lyu@connect.ust.hk
The Hong Kong University of Science and Technology
Hong Kong, China

Zhenhui Peng
pengzhh29@mail.sysu.edu.cn
Sun Yat-sen University
Zhuhai, China

Xiaojuan Ma*
mxj@cse.ust.hk
The Hong Kong University of Science and Technology
Hong Kong, China

ABSTRACT

Online critique communities (OCCs) provide a convenient space for creators to solicit feedback on their artifacts and improve skills. Creators' behavioral, emotional, and cognitive engagement with comments on their works contribute to their skill development. However, what kinds of critique creators feel engaging may change with the creation stage of their shared artifacts. In this paper, we first model three dimensions of engagement expressed in creators' replies to peer comments. Then we quantitatively examine how their engagement is affected by artifacts' stage and feedback characteristics via regression analysis. Results show that creators sharing works-in-progress tend to exhibit lower behavioral and emotional engagement, but higher cognitive engagement than those sharing complete works. The increase in the valence of the feedback is associated with a stronger increase in behavior engagement for seekers sharing complete works than works-in-progress. Finally, we discuss how our insights could benefit OCCs and other online help-seeking platforms.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in collaborative and social computing**.

KEYWORDS

Critique, online community, behavioral engagement, emotional engagement, cognitive engagement

*Corresponding author

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3581054

ACM Reference Format:
Qingyu Guo, Chao Zhang, Hanfang Lyu, Zhenhui Peng, and Xiaojuan Ma. 2023. What Makes Creators Engage with Online Critiques? Understanding the Role of Artifacts' Creation Stage, Characteristics of Community Comments, and their Interactions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3544548.3581054

1 INTRODUCTION

For creators, soliciting feedback on their artifacts is an essential step toward polishing their works and enhancing creativity skills [30]. A variety of online critique communities (OCCs) thus emerge as an informal learning space for creators, especially those with little opportunity to receive critiques from private feedback exchange groups or professional services [19, 22]. In OCCs, critique seekers typically post their artifacts (such as artworks [48], graphic designs [19], photography [105], and fanfictions [18]) publicly and then learn from the received peer feedback toward their works [48, 105]. For instance, r/learnart is an active OCC on the Reddit platform with 459k followers as of November 2022 [84], designed for "artists who want to improve" and allows creators to "check for feedback". Different from other interest-driven communities, OCCs are dominated by constructive critiques on artifacts, and members pursue skill development rather than entertainment or self-advertisement [22, 64]. Moreover, seekers can solicit support during the creative process, including the pre-creation stage (such as preparation and ideation), and the ***creation stage*** (i.e., work-in-progress and complete) [48, 52]. In this work, we focus on the creation stage where most critique-requests happen, and seekers could benefit from the received technical suggestions and polish their artifacts [52].

In the creation stage, seeker's engagement with feedback contributes to their artifact improvement and skill development [64, 101], and seekers usually demonstrate engagement with feedback in their replies to the critique providers [14, 68]. Seekers' expressed engagement with feedback not only signals their participation in creative skill learning, but also reciprocally makes the providers feel worthy of knowledge sharing [57]. Previous qualitative studies indicated that seeker's engagement could vary between sharing in-progress works and complete ones [48]. For example, seekers sharing in-progress works are likely to experience frustration and even quit learning, indicating relatively low engagement [51, 56]. Moreover, seekers' expectations of the received feedback may differ given the artifact's creation stage [30]. For instance, timely feedback is more likely to engage seekers sharing works-in-progress than complete works [105]. However, discrepancies may exist between the small-scale, self-reported experiences of seekers obtained in previous studies and general OCCs members' actual engagement with peer feedback. There lacks a quantitative understanding of how feedback should be adapted to the creation stage for the sake of engaging seekers [30]. Therefore, a data-driven analysis is needed to comprehensively investigate relationships between artifact's creation stage, feedback characteristics, and their interactions to seekers' engagement, which is helpful for members in OCCs to achieve the goal of skill development. The resulting insights could also offer design implications for technological support (such as adaptive provider matching [100] and intelligent feedback writing assistance [77, 93]) that benefits seekers in OCCs.

In the scope of this paper, seeker's engagement with feedback can be divided into behavioral engagement, expressed emotional engagement, and expressed cognitive engagement [64, 101]. *Behavioral engagement* stands for seekers' *willingness to* participate in the critique discussion [22, 52]. As critique exchange in OCCs often takes place through free-form dialogues [52], seekers need to communicate with providers to obtain deeper insights for refining their artifacts. *Expressed emotional engagement* refers to seekers' affective states with the received feedback, and a positive emotional response could encourage seekers' creative thinking and commitment to their skill development [70, 103]. *Expressed cognitive engagement* can be conceptualized as seekers' *willingness* to invest efforts in the feedback, an indicator of the extent to which the seeker would accept the received input and polish their works accordingly [59]. In this work, we quantitatively examined the effect of factors on seekers' expressed engagement in OCCs. Specifically, we raise the following research questions: how would the creation stage of a shared artifact, feedback characteristics, and their interactions affect seeker's **RQ1)** behavioral engagement, **RQ2)** expressed emotional engagement, and **RQ3)** expressed cognitive engagement with the corresponding peer feedback.

To this end, we first collect large-scale data from four art-related OCCs on the Reddit platform. We measure the behavioral engagement to be whether seekers reply to the peer responses to their works, and quantify their expressed emotional engagement with VADER [43] – a commonly adopted tool for analyzing user sentiment on social media. We develop a coding scheme to characterize the expressed cognitive engagement and train a BERT-based model for cognitive engagement level classification with 820 annotated samples. Next, we develop a deep learning-based model to classify a target artifact's creation stage (i.e., work-in-progress versus complete) based on its associated critique-seeking post. Then, we characterize received critiques in the comment threads with commonly discussed features in feedback exchange, including content-based features (i.e., actionability, justification, specificity, and valence) and timing (i.e., delay) of the feedback [19, 54]. Finally, we apply regression models to answer the research questions.

Our results show that, in general, seekers sharing works-in-progress tend to present lower behavioral engagement and more negative emotional engagement, but are more likely to express a higher-level of cognitive engagement than those sharing complete works. Although the valence of peer comments is typically positively correlated with seekers' behavioral and expressed emotional engagement, it is negatively correlated with their expressed cognitive engagement. We also found that the role of feedback characteristics in creators' critique engagement may vary in different creation stages. For example, the increase in valence of the feedback is associated with a smaller increase in behavior engagement for seekers posting works-in-progress than sharing complete works.

Our work makes several contributions to research on OCCs and online support. **Empirically**, we extend the understanding of how artifacts' creation stage, feedback characteristics, and their interactions may affect seekers' engagement in OCCs by analyzing five-year-longitudinal data. **Practically**, we contribute methods for quantifying and modeling seekers' engagement and predicting artifacts' creation stage, and propose design opportunities to help members achieve their goal of efficient feedback exchange and skill development in OCCs. **Theoretically**, we discuss how our insights and analysis workflow can be generalized to other online help-seeking communities where members may go through different stages with varying expectations on the received support.

2 RELATED WORK

In this section, we first introduce the benefits of OCCs and motivate our work through OCCs seekers' feedback engagement. Then we survey how the artifact's creation stage and feedback characteristics may affect feedback engagement in a general creative context. Finally, we situate this paper in previous OCCs studies and summarize our research gaps.

2.1 Benefits of OCCs

Soliciting feedback on their creation is a crucial step for creators to refine their work and improve creativity skills [30]. Online critique communities provide a convenient space for creators to share their creative artifacts, solicit support, and provide feedback, especially those with little opportunity to receive formal feedback [22, 30]. Creators join OCCs (e.g., r/learnart [85], r/artcrit [82] on Reddit) and upload their original creative works (such as paintings [48], UI designs [19], photographys [105], and fanfictions [18]) to solicit critiques from peers with primary purposes of skill development and artifact improvement [19, 64, 105]. OCCs differ from other online help-seeking communities and interest-driven communities in several ways. On the one hand, compared with other online help-seeking communities (e.g., for health issues [107]), the communication between seekers and providers revolves around the seekers' artifacts [19, 105]. On the other hand, seekers in OCCs aim to improve their creativity skills by learning from feedback toward their artifacts, instead of for leisure [88] or promoting themselves [50] in general interest-driven networks [19, 22, 64]. Moreover, seekers can solicit support during their creative process, rather than only present the final work [48]. Previous studies generally divided

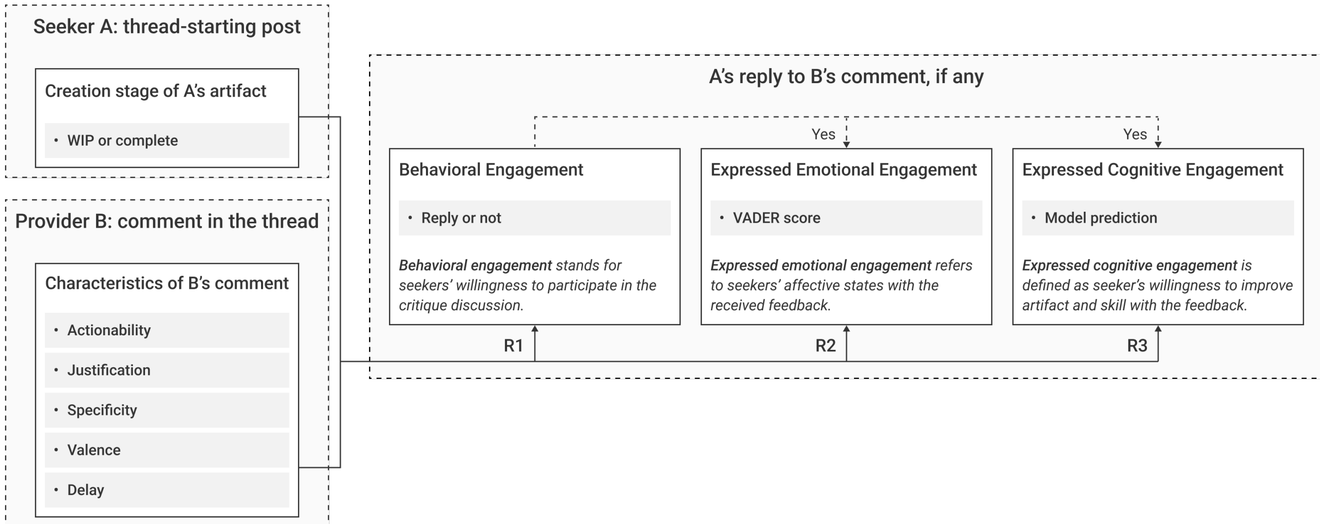


Figure 1: The concept diagram of this study. We analyzed the effect of the creation stage of the shared artifact, characteristics of the received feedback, and their interactions on the RQ1) behavioral engagement, RQ2) expressed emotional engagement, and RQ3) expressed cognitive engagement of the seeker in OCCs. Here, "WIP" stands for work-in-progress.

the creative process into pre-creation stage (such as preparation and ideation) and the creation stage (i.e., work-in-progress and complete), and they found that critique-requests in creation stage dominate the posts in OCCs [48, 52]. In this work, we focus on the creation stage where seekers could learn from the technical suggestions to polish their artifacts [52].

Prior works suggested that short-term cues could be important indicators to ensure the effectiveness of support in online communities [38, 77, 81, 99]. For example, researchers measured the seekers' expressed satisfaction with the received comments and the positive sentiment changes in online health communities to investigate mental health improvements [77, 81, 99]. Likewise, student engagement with feedback in discussion threads serves as an immediate cue to study the learning effects of online learning platforms (e.g., MOOCs) [39, 45, 68, 106]. Winstone et al. [101] also suggested that seekers' engagement with the feedback rather than the feedback itself is crucial for learning. In OCCs, seekers' feedback engagement stands for the extent to which they would like to understand and adopt the feedback, which determines whether they will revise or iterate their artifacts [55, 59, 70, 103]. By engaging with feedback, they recognize how others perceive their work, perform a reflection activity, and further practice for skill development [28, 111]. In text-based online communities, seekers usually express their engagement in their reply to the received comment [14, 68], which could reciprocally make the providers feel worthy of knowledge sharing [57]. We motivate our study with the benefits of OCCs and choose to evaluate seekers' feedback engagement to understand the underlying factors in online feedback exchange.

2.2 Seekers' Feedback Engagement in Creative Context

In the general creative and learning context, previous work approached feedback engagement from three aspects–*behavioral engagement*, *emotional engagement*, and *cognitive engagement* [31, 65]. Firstly, behavioral engagement refers to students participating in a learning activity, such as asking questions [31, 65]. Xu et al. [106] calculated the number of messages each online learner sends as an indicator of behavioral engagement. Likewise, whether original posters (OPs) reply to the feedback or not can also be an important cue of seekers' behavioral engagement [52]. Secondly, emotional engagement refers to the receivers' feelings or sentiments toward feedback [31, 104]. Nguyen and colleagues used the PANAS scale [35] to measure seekers' emotional responses to feedback [31]. In online learning communities, researchers usually conduct sentiment analysis using natural language processing techniques such as VADER [43] to analyze the expressed emotional engagement [34, 61, 68], which inspired us to use the same method in OCCs. Finally, cognitive engagement can be conceptualized as the willingness to invest in learning and improving with the feedback [20, 31, 90]. Prior research gathered tentative actions taken by seekers in response to feedback, such as implementing it, considering it, or ignoring it [22, 55], as well as the perceived helpfulness [54, 113]. In OCCs, OPs usually express their actions towards feedback in their replies, which provides a potential opportunity for us to investigate their expressed cognitive engagement.

Sharing artifacts at different stages in OCCs and engaging with peer reviews provide opportunities for seekers to iterate their works and achieve skill growth [110, 111]. Previous studies have implied that changes in seekers' creation stages may shift their feedback engagement [10, 40, 55, 71, 80]. For example, creators would prefer minor revision suggestions or affirmations for complete works, while expecting suggestions about design alternatives for works-in-progress [56]. Therefore, to offer implications for improving the feedback engagement of OCC seekers, it is important to understand the differences in behavioral engagement, emotional engagement, and cognitive engagement between sharing in-progress works and complete works in current OCCs practice.

In addition, Winstone et al. [101] suggested that feedback attributes may influence seeker's feedback engagement. On the one hand, linguistic characteristics of received feedback content may influence seekers' engagement [6, 12, 27, 47]. For example, Krause et al. [54] extracted the feedback's linguistic features (e.g., actionability, justification, specificity, valence) and analyzed their correlation with the perceived usefulness of design feedback. On the other hand, the timeliness of feedback may also be an important factor affecting seekers' engagement. For instance, Xu et al. [106] found that seekers hope to get quick feedback for their ongoing works. However, previous studies focused more on investigating the connection between feedback characteristics and feedback acceptance (related to cognitive engagement); less knowledge exists on how feedback characteristics may correlate with the seeker's behavioral and emotional engagement, despite the importance for seekers to gain deeper insights from OCCs and continue the creation [64, 101]. Moreover, it is unclear whether and how the effect of these feedback characteristics on seekers' engagement may vary in their different creation stages. Such understanding can provide implications for adapting feedback to the artifact [30] to engage creators in OCCs.

2.3 Studies of Online Critique Communities

Previous studies have widely explored the characteristics and dynamics of OCCs. For example, Kou and Gray conducted a series of empirical studies to understand the content in OCCs, i.e., the distribution of critiques [52] and professional self-disclosure in threads [53]. They found that most support requests occur during the creation stage (i.e., work-in-progress or complete) instead of pre-creation (such as preparation and ideation) [52]. Another group of studies put effort into motivation [64] and socio-psychological factors [30] behind the community members participating in OCCs. For example, Xu et al. [105] explored seekers' expectations and member interactions in an online photo critique community. Notably, some researchers focused on seekers to understand their seeking behaviors [22] and requesting strategies [19] in OCCs. For example, Cheng et al. [19] identified how specific request strategies impact the quantity and quality of feedback in a subreddit design critique community. Kim et al. [48] explored a prototype online critique community where only works-in-progress are allowed to share, aiming to encourage creators to seek support during their creation. In line with these previous studies, we fill in the understanding of how artifact creation stages, community feedback characteristics, and their interactions would affect feedback engagement expressed in the seekers' replies to received comments in OCCs. Based on the findings, we also provide several design opportunities for OCCs to enhance creator engagement with online critiques.

3 METHOD

In this section, we first introduce the research site and the data processing pipeline. We then introduce methods to measure and model the dependent variables (i.e., seeker's behavioral, emotional, and cognitive engagement). Next, we introduce how we extract the independent variables, including the artifact's creation stage and feedback characteristics.

3.1 Research Sites and Dataset

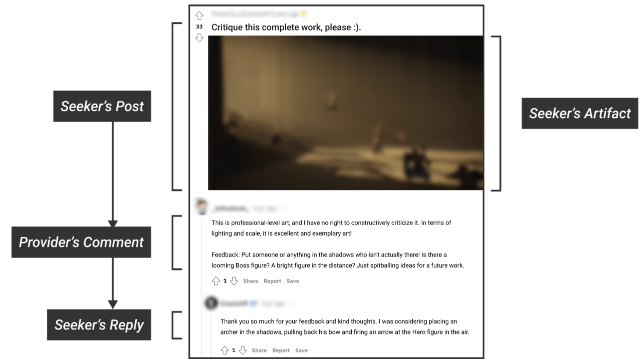


Figure 2: Example feedback exchange thread in r/learnart. A seeker would solicit feedback by initiating a thread with a post, then a provider could criticize it by leaving comments, to which a seeker may reply. We decrease the resolution and obscure sensitive information for copyright and privacy concerns. We slightly paraphrase the content in the post so that the post cannot be searched. These operations apply to all figures in this work. Image © Reddit.

3.1.1 *Reddit Platform for Online Feedback Exchange.*

Our dataset was derived from the art-related communities on the Reddit platform. We used data from the Reddit platform as it offers a wide range of OCCs open to all levels of creators. Creators, especially those with little chance to receive feedback from private feedback exchange groups or professional critique services [22], could discuss artworks and exchange critiques in these communities for free [19, 22]. We further focused on visual art-related communities for the subsequent processing and analysis for two reasons. First, when creating artworks, creators usually seek critiques during their creation process to refine their artifacts and improve their skills [28, 89], which could guarantee the quantity and diversity of the critique-seeking posts. Second, visual artifacts are a representative creation format and have been investigated in several prior studies about OCCs [19, 22, 105].

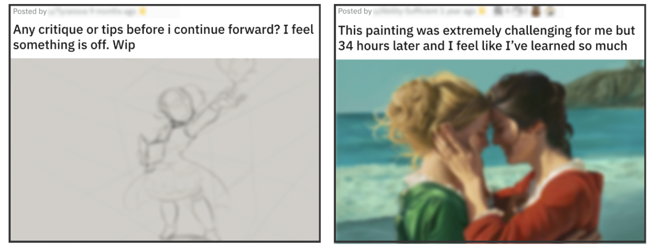
We retrieved appropriate OCCs as our research sites by following the keyword search and snowball sampling methods in [52]. Specifically, we first queried communities with a set of art-related keywords (such as "art", "artwork", "drawing", "painting" and "sketches") via the Reddit community search engine. Then we explored the associated communities of the search results to extend the list of candidate communities for analysis. This led to 53 subreddits (e.g., r/learnart, r/artshub, r/doodles). Since the goal of the study is to examine factors related to seekers' engagement in soliciting feedback in the creation stage, we removed the non-critique oriented communities (such as for entertainment or self-advertisement), and only include those encouraging seeking and providing critiques by checking community rules and by observation. After this step, seven communities remains. We further removed three subreddits with less than 10k followers where seekers rarely receive comments, which could not provide sufficient support-seeking interaction data for analysis. We acknowledge that this sampling procedure may not fully reflect the entire OCC landscape, and we provide suggestions for future research in in Section 6. Finally, we maintained four active subreddits that encourage constructive critiques and forbid creators from promoting themselves when posting – r/learnart, r/ArtCrit, r/DigitalPainting, and r/FurryArtSchool. Figure 2 illustrates a typical critique-seeking thread in our selected community. In the subsequent data preprocessing and analysis, we combined the data from the four communities for generalizability.

3.1.2 Data Collection and Preprocessing. We collected all publicly available data between January 2017 and December 2021 via Pushshift API [5]. The initial data included seekers' posts (thread-starting posts), providers' comments (messages in the discussion threads created by members beyond the seeker), and seekers' replies to the received comments. In this study, we did not restrict the position of the providers' comments on the first level in the thread since critiques in OCCs could appear throughout the thread [52], and we were interested in investigating seekers' engagement with each critique. We took several steps to pre-process the collected data. First, we removed the posts which contained the "NSFW" (a.k.a., Not Safe For Work) tag for ethical concerns. Then, we filtered the posts, comments, and replies without author names to distinguish providers' comments in the discussion thread from seekers. We also removed the posts, comments, and replies whose content is "[removed]" or "[deleted]" [32]. Next, we removed the posts without receiving comments from other community members.

Two researchers who are familiar with OCCs randomly sampled 480 posts from the communities and screened the content. We found that most of the sampled posts ($N = 446$) are critique-seeking posts, with 443 of them containing images of the shared artifacts, which are either in the work-in-progress stage or completed stage. For the other three critique-seeking posts identified based on the textual content in the post, we noticed that the received comments were asking for artifact uploads rather than critiques for the artifact. Meanwhile, we recognized 34 non-critique-seeking posts, mainly about seeking tutorials or emotional support. Only five of them contain images (1.04% of sampled posts), which are seeking learning tutorials with a reference. The exploratory analysis suggests that the exist of images in the post could be an indicator of critique-seeking posts of artifacts in the creation stage (the focus of this study) in art-related OCCs. One possible reason is that the norm exists in these communities. For example, in r/learnart, one community rule explicitly mentioned that "include your own work to get clear feedback". Therefore, we further removed the posts without images and their associated comments based on the scraped metadata. Figure 3 demonstrates two examples of critique-seeking posts in OCCs. Finally, our dataset consists of 81,346 posts, 312,437 provider's comments, and 125,333 seeker's replies.

3.2 Feedback Engagement Measurement

Prior studies in the creativity and learning domains indicated that creators' engagement with the received feedback could be decomposed into three aspects – behavioral engagement, emotional engagement, and cognitive engagement [25, 31, 65]. These three dimensions of engagement contribute to creators' artifact improvement and skill development [64, 101]. In this section, we introduce how we quantified seekers' behavioral, expressed emotional and cognitive engagement toward online critiques.



(a) Example post sharing an in-progress work (b) Example post sharing a complete work

Figure 3: The example posts for critiques: (a) is sharing an in-progress work and (b) is sharing a complete work. We decrease the resolution and obscure sensitive information for copyright and privacy concerns. Images © Reddit.

3.2.1 Behavioral Engagement. Behavioral engagement is defined as whether a seeker participates in the critique discussion thread after receiving a comment [52, 77]. If a comment receives the reply from the seeker, it is defined as 1, otherwise 0. We measured the seekers' behavioral engagement for all comments in our dataset. As shown in Table 3, in our dataset, around 40% of comments received the seekers' replies.

3.2.2 Expressed Emotional Engagement. Seekers' expressed emotional engagement with a comment (i.e., seeker's affective state when receiving a critique [70, 103]) were quantified with VADER [43] – a commonly adopted tool for analyzing user sentiment on social media. The VADER sentiment score ranges from -1 (most negative sentiment) to 1 (most positive sentiment). Table 1 shows some example seekers' replies with the VADER sentiment score. As shown in Table 3, the mean, median, and standard deviation of the measured sentiment score are 0.49, 0.54, and 0.38, respectively.

3.2.3 Expressed Cognitive Engagement. Previous studies suggested that the expressed cognitive engagement stands for the seeker's **willingness** to invest effort in the feedback to improve the artifact and skill [59]. To quantify the expressed cognitive engagement of each reply, we first identified patterns of seeker's reply through open-coding, then grouped them into different levels based on their demonstrated willingness to accept the feedback. Specifically, following the thematic analysis procedure in [11, 66], two researchers first sampled 120 replies and coded them independently. They then merged and synthesized codes and determined the final coding themes over multiple rounds of discussion. Finally, we identified three levels of expressed cognitive engagement (i.e., low, middle and high), and Table 2 summarizes common examples in these three levels. The low level stands for seekers only socializing with providers, such as expressing thankfulness, while without actively digesting the feedback, the prerequisite of being willing to invest effort in the feedback [101]. The middle level stands for the seeker exchanging information with the providers, and the common characteristics are clarifying their initial thoughts, and raising follow-up questions to providers. This reflects that seekers are willing to digest the received feedback, but have not yet decided whether to accept the suggestions. The high level stands for seekers demonstrating their willingness to accept the suggestions. Seekers may mention they will adopt suggestions for current or future works immediately after receiving the feedback (i.e., planned action); alternatively,

Table 1: Expressed emotional engagement examples.

Seeker's Reply	VADER Score
If only I weren't afraid of the microns… Seriously though I need practice not messing up a piece through inking, so maybe I'll give it a shot if I can get over the "oh god I'm gonna ruin it" panic.	-0.77
Thank you! That was probably the hardest part, knowing how light/vibrant to leave it, so the colors didn't get lost in the black background.	0.00
Beautiful advice. Thanks so much for sharing all of that!! I can't wait to try an undercoat next. I feel like that would have made such a huge difference. I'm obsessed with shine and gloss in most work, so I will definitely look into getting that. Thanks again! Love your name btw.	0.97

Table 2: Expressed cognitive engagement coding scheme and examples.

Level	Characteristic	Example Seeker's Reply
Low - Socialization	Socialization	*Hard work pays off - thank you soo much for your comment. Cheers and best wishes.*
Middle - Information Exchange	Provide clarification or explanation	*It doesn't matter, I was going for meaningful instead of accurate.*
	Ask follow-up questions	*I agree. What do you think could make it less empty?*
High - Feedback Acceptance	Will investigate in the future	*Thank you for your feedback! I will work on that next time.*
	Will investigate for the shared artifact	*Thanks for that, I will pay more attention to this valley now.*
	Have investigated for the shared artifact	*I did it and it looked way better! Thank you so much!*

seekers may not express their willingness to follow the suggestion in the first place, but confirm it after they have implemented the feedback (i.e., executed action). We acknowledge that reporting the executed action might suggest a stronger willingness to accept suggestions than the planned action. However, our sample of 120 replies contained only three instances of the executed action. Such imbalanced data makes it difficult to obtain a reliable model, and the classifier trained on four categories achieved an F1-score of 0.4 on the category of executed action in our trial. Since reporting both the planned and executed action could imply the seeker's willingness to accept and act on the feedback, we combine these replies into the high level of cognitive engagement to distinguish them from the low and middle level. Moreover, we acknowledge that the seeker's reply is a proxy of their actual cognitive engagement, as discrepancy may exist between their expression and actual willingness. However, this proxy could shed light on understanding seeker's expressed cognitive engagement in soliciting feedback in OCCs. We elaborate on these points in Section 6.

Based on the coding scheme, these two researchers continued to annotate another 700 randomly sampled seeker's replies from the dataset. Following the method proposed by Zhu et al. [114], they first annotated all levels of cognitive engagement manifested in each sentence of the seeker's reply. They then took the highest level of expressed cognitive engagement in each sentence as the label of the reply. For instance, if one reply contains both the sentence of "*Thank you for your comment.*" (expressing low cognitive engagement) and "*I implemented your suggestions.*" (expressing high cognitive engagement), the final level of the expressed cognitive engagement is aggregated as being high. The initial inter-rater metric Cohen's Kappa is 0.86, and they discussed to resolve conflicts. After these two steps, we got 252/352/216 replies that expressed low/middle/high cognitive engagement.

We then developed computational models to classify a seeker's reply into three cognitive engagement levels. We compared the performance of machine learning models (Support Vector Classifier (SVC), Multi-Layer Perceptron classifier (MLP), and Random Forest (RF)) and the deep learning-based BERT model [24] for this task. Following prior studies in online community content analysis [76, 108], we extracted lexicon-based features (i.e., the word frequency under each category of LIWC dictionary [78], 73 features in total) and statistic features (i.e., number of words, sentences, characteristics, URLs, emojis, and numeric, 6 features in total) in seekers' replies to train machine learning models. Table 5 provides a detailed description of the selected features. We implemented the BERT-based classifier using the pre-trained BERT model from the HuggingFace Transformers library [102].

We split the annotated data into a training set (85%) and a test set (15%). We further split 15% of the training data to validate the BERT-based classifier. We determined the hyper-parameters of machine learning models by 10-fold cross-validation on the training set and fine-tuned the BERT-based classifier with an early stop mechanism [13] on the validation set. Among these models, the BERT-based classifier achieved the highest accuracy (0.87) and macro F1-score (0.86) on the test set. We thus predicted the level of expressed cognitive engagement with the BERT-based classifier and obtained 37,578 replies expressing low cognitive engagement (30.0%), 58,295 for middle cognitive engagement (46.5%), and 29,460 expressing high cognitive engagement (23.5%).

3.3 Creation Stage Classification

Following the conventions of OCCs and findings in previous qualitative research in online feedback exchange [19, 48, 52], we classify the creation stage of an artifact as the work-in-progress stage or completion stage. We introduce the procedures of dataset construction and classifier development in this section.

In the subreddit – r/learnart, creators could add the tag of "In the Works" or "Complete" to their critique-seeking posts, which denote the creation stage of their shared artifacts (work-in-progress or complete). Following prior works [46, 63, 96], we developed creation stage classifiers with the supervision of the user-generated tags. To this end, we randomly sampled 2, 700 posts under each of the two tags, which forms the LEARNART-TAG dataset for training and selecting the creation stage classifiers.

We formulated the creation stage prediction as a binary classification task – whether the post is seeking support for an in-progress artifact. We noticed that the shared artifacts might no longer be downloadable with the retrieved links, probably because seekers were concerned about the theft of their creations [1]. However, whether the seeker would like to remove the shared images may be related to their engagement with the online critiques [48]. To avoid selection bias, we predicted the creation stage of the shared artifact solely based on the textual content in the post, instead of images. We implemented the machine learning models (SVC, MLP, RF) and the deep learning-based model following the approach illustrated in Section 3.2.3. To fully utilize the textual information, we further proposed a linguistic feature-enhanced BERT-based model (LF-BERT model), which integrates both the post's semantic representation and linguistic features. Specifically, we first aligned the semantic representation (obtained with a pre-trained BERT model) and the linguistic features (i.e., 73 LIWC-related features and six statistic features) of the post into one representation space using two independent fully connected layers. Then we concatenated the semantic representation and linguistic features, followed by a fully connected layer for binary classification.

We split the LEARNART-TAG dataset into a training set (80%) and a test set (20%). For the deep learning-based models, we further split 25% of the training data for validation. Following the same training and model selection procedure described in Section 3.2.3, the LF-BERT model achieved the best classification performance (F1-score: 0.77/0.80 for work-in-progress/complete).

To evaluate the generalizability of the classifier, we further prepared an OCC-ANNOTATION dataset by sampling and annotating critique-seeking posts in four communities. We randomly sampled 150 posts from each of the four communities, with no overlap between the LEARNART-TAG dataset. Then we recruited three annotators with art-related backgrounds from a local university in China via word of mouth. Each was compensated with a voucher worth USD 50. They all have an art-related degree or certification and at least eight years of art learning experience. Two annotators individually labeled the creation stages of the sampled posts with the artwork images (if still downloadable) and seeker's text descriptions. They were allowed to label "not sure" as some artifacts' images were not accessible with the retrieved links due to being deleted by the creator after sharing for a period. The Cohen's Kappa is 0.78, indicating a substantial agreement [67]. The disagreements were resolved by discussing with the third annotator and majority voting. Finally, we got 151 works-in-progress and 397 complete posts, which formed the OCC-ANNOTATION dataset. The LF-BERT model also achieved a reasonable performance on the OCC-ANNOTATION dataset (F1-score: 0.72/0.89 for work-in-progress/complete), indicating good generalizability of the model. Using the LF-BERT model, we predicted the creation stage of 81,346 posts with the textual information. Among them, 29,495 were predicted as in-progress works, and 51,851 were predicted as complete works.

3.4 Feedback Characteristic Extraction

We considered four commonly used content-related features in prior studies of creativity critiques [54] – *actionability*, *justification*, *specificity*, and *valence*. The actionability of feedback stands for the number of concrete suggestions toward the seeker's artifact, indicating the revision space of the artifact. We calculated the actionability by counting the number of non-indicative (command or suggestions) sentences in feedback [91]. For example, "*You may check by simply overlaying your sketch over the original photo*" could be recognized as one actionable sentence. The justification of feedback represents the extent to which a provider's suggestions are backed up with evidence and reasoning. It was calculated as the number of sentences containing words indicating reasoning (such as "*because*" and "*since*") [42]. The specificity of feedback stands for the concreteness of the suggestions. It was measured as the maximum depth of the words in the comment based on the Wordnet structure [29]. Words closer to the root are more general (e.g., art), while words deeper in the Wordnet are more specific (e.g., magenta). In addition, the valence of the feedback stands for whether the comment is positive or negative, which may affect the seeker's affective state and the willingness to iterate the artifact [103, 104]. We extracted all the content-related features with Python libraries following the classic pipeline proposed by Krause et al. [54]. More specifically, feedback's actionability, justification, and valence were calculated with the pattern.en package[1], and the specificity was obtained based on the Natural Language Toolkit[2]. Besides these content-related features, we also considered the timing issue of feedback. We measured the feedback *delay* as the time interval between the post creation and the comment creation (in seconds) [19]. Table 3 shows the descriptive statistics of five feedback-related independent variables.

4 RESULT AND ANALYSIS

We conduct three sets of regression models to address our research questions. RQ1 (behavioral engagement) is based on the dataset of 312,437 post-comment pairs (includes all providers' comments, labeled as 1 if a comment receives the seeker's reply, otherwise 0), while RQ2/3 (expressed emotional / cognitive engagement) is based on the dataset of 125,333 post-comment-reply triplets (only includes providers' comments replied by seekers since the label relies on the seeker's reply). The statistics of all independent and dependent variables are depicted in Table 3. Before the regression analysis, all independent variables except for the categorical data (i.e., creation stage) were standardized by centering to zero mean and divided by the standard deviation following the preprocessing methods in previous works [107]. As shown in Figure 4, the Pearson correlations between each pair of variables are less than 0.45, indicating collinearity in RQ1 and RQ2/3 datasets are acceptable following analysis [7]. The variance inflation factor (VIF) among the variables are all less than three, suggesting multicollinearity is not a problem for our regression model.

For RQ1 (behavioral engagement), since the dependent variable is binary (whether the seeker replies to the feedback provider), we adopted logistic regression for analysis. As the dependent variable in RQ2 (expressed emotional engagement) is interval data, we utilized the linear regression model. Furthermore, for RQ3 (expressed cognitive engagement), we adopted the ordinal logistic regression

[1] https://github.com/clips/pattern
[2] https://www.nltk.org/

Table 3: Descriptive statistics of variables for predicting seekers' behavioral (RQ1), expressed emotional (RQ2), and expressed cognitive engagement (RQ3). In the table, "WIP" stands for work-in-progress.

		RQ1 Dataset for behavioral engagement ($N = 312,437$)			**RQ2/3** Dataset for expressed emotional / cognitive engagement ($N = 125,333$)		
	variables	**min/max**	**mean (std)**	**median**	**min/max**	**mean (std)**	**median**
CREATION STAGE	**is WIP**	0/1	0.35 (0.48)	0.00	0/1	0.34 (0.47)	0.00
RECEIVED COMMENT	**actionability**	0/43	0.80 (1.32)	0.00	0/43	0.95 (1.5)	0.00
	justification	0/12	0.11 (0.40)	0.00	0/12	0.15 (0.45)	0.00
	specificity	0/19	10.53 (2.59)	10.00	0/19	10.83 (2.37)	11.00
	valence	0/9	5.59 (1.29)	5.42	0/9	5.6 (1.22)	5.44
	delay (sec)	1/138M	128K (1.59M)	22.1K	3/112M	76K (0.9M)	17.7K
SEEKER REPLY	**behavioral engagement**	0/1	0.4 (0.49)	0	-	-	-
	expressed emotional engagement	-	-	-	-1/1	0.49 (0.38)	0.54
	expressed cognitive engagement	-	-	-	1/3	1.94 (0.73)	2

(a) correlations between independent variables in RQ1 dataset for seeker's behavioral engagement ($N = 312,437$).

(b) correlations between independent variables in RQ2/3 dataset for expressed emotional/cognitive engagement ($N = 125,333$).

Figure 4: Correlations between each pair of independent variables in our RQ1 and RQ2/3 datasets. Here, "WIP" stands for work-in-progress.

model because the cognitive engagement level is ordinal data. The regression coefficients are listed in Table 4.

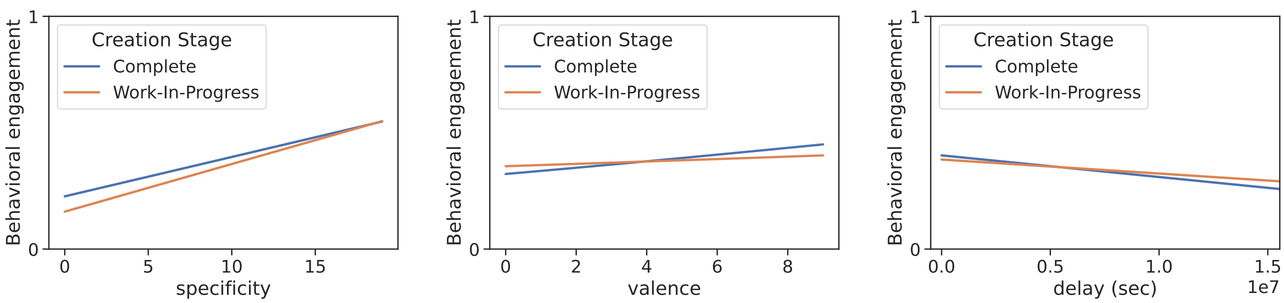
4.1 RQ1. Behavioral Engagement

Model 1 in Table 4 shows the main effects of the artifact's creation stage and feedback characteristics on seekers' reply behavior. As indicated by model 1's coefficient, seekers are less likely to reply to feedback for their in-progress works ($\beta = -0.135$) compared with sharing complete works. Moreover, when receiving feedback with higher actionability ($\beta = 0.12$), higher justification ($\beta = 0.053$), higher specificity ($\beta = 0.153$), and higher valence ($\beta = 0.039$), seekers tend to have higher chance to participate in the discussion threads. On the contrary, seekers are less likely to respond to late feedback ($\beta = -0.124$) compared with earlier ones, which is consistent with the phenomenon in generic community question-answering sites [95].

Model 2 adds the interaction terms between the artifact's creation stage (whether it is work-in-progress) and feedback characteristics. The likelihood-ratio test [8] suggests that model 2 is significantly better ($p < 0.001$) than model 1 in fitting the behavioral engagement. We further depict the interaction between the creation stage and feedback characteristics (specificity, valence, and delay, that show significance in the interaction terms) on seekers' behavioral engagement in Figure 5. As shown in Figure 5a, eliciting the specificity level of feedback is generally associated with a sharper increase ($\beta = 0.011$) in behavioral engagement for seekers sharing works-in-progress in contrast to sharing complete works. Figure 5b demonstrates that receiving feedback with higher valence seems to be associated with an increase in behavioral engagement for seekers sharing both works-in-progress and complete works, but this increase is slower ($\beta = -0.038$) for works-in-progress. In addition, from Figure 5c, we notice that extended delay in feedback is typically associated with a slower decline for those sharing ongoing works ($\beta = 0.036$) than those in the complete stage. The result may contradict a prior study that indicated seekers hoped to receive speedy feedback when their works were uncompleted [105]. However, no significant difference ($p > 0.05$) in the changes of reply behaviors between the seekers sharing in-progress and complete

Table 4: Regression coefficients of models for predicting behavioral engagement (RQ1), expressed emotional engagement (RQ2), and expressed cognitive engagement (RQ3). In the table, ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$. "WIP" stands for work-in-progress.

	Behavioral engagement		Emotional engagement		Cognitive engagement	
predictors	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
creation stage (is WIP)	-0.135 ***	-0.139 ***	-0.017 ***	-0.017 ***	0.264 ***	0.28 ***
actionability	0.120 ***	0.125 ***	0.046 ***	0.051 ***	0.257 ***	0.320 ***
justification	0.053 ***	0.054 ***	0.007 ***	0.007 ***	0.047 ***	0.055 ***
specificity	0.153 ***	0.148 ***	0.034 ***	0.032 ***	0.422 ***	0.436 ***
valence	0.039 ***	0.060 ***	0.047 ***	0.049 ***	-0.236 ***	-0.252 ***
delay	-0.124 ***	-0.145 ***	-0.002	-0.001	-0.017 **	-0.021 *
creation stage (is WIP) * **actionability**		-0.008		-0.007 ***		-0.091 ***
creation stage (is WIP) * **justification**		-0.001		2.99E-05		-0.011
creation stage (is WIP) * **specificity**		0.011 *		0.003 *		-0.032 ***
creation stage (is WIP) * **valence**		-0.038 ***		-0.004 **		0.036 ***
creation stage (is WIP) * **delay**		0.036 **		-0.001		0.004
R^2	0.011	0.012	0.043	0.043	-	-
Log-Likelihood	-2.073E+05	-2.072E+05	-52967	-52951	-124910	-124790



(a) Specificity of the feedback **(b) Valence of the feedback** **(c) Delay of the feedback**

Figure 5: Interaction between creation stage and feedback characteristics on seeker's behavioral engagement.

works is observed when adjusting the actionability or justification of the feedback.

4.2 RQ2. Expressed Emotional Engagement

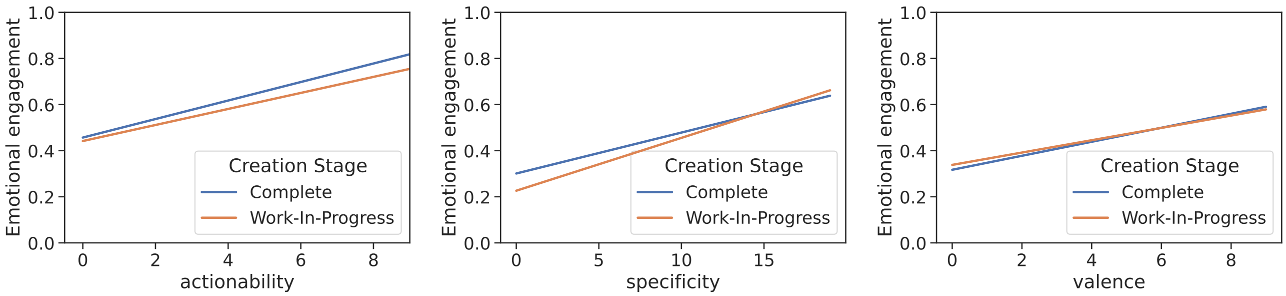
The main effects of the artifact's creation stage and the feedback characteristics on seekers' expressed emotional engagement are shown in Model 3 (Table 4). Seekers sharing in-progress works tend to express more negative emotional engagement ($\beta = -0.017$) than those sharing complete works. The result also suggests that, the actionability ($\beta = 0.046$), justification ($\beta = 0.007$), and specificity ($\beta = 0.034$) of the feedback, are positively correlated with their expressed emotional engagement. Moreover, valence of the feedback is positively correlated ($\beta = 0.047$) with seekers' expressed emotional engagement, which is consistent with the findings about design critiques through lab experiments [70, 72].

Model 4 in Table 4 further considers the interaction effects of the creation stage and feedback characteristics on seekers' expressed emotional engagement. Although the R^2 value is the same as it is in model 3, the log-likelihood value of model 4 is higher than that of model 3, indicating that the goodness of the regression model keeps increasing. The F-test [15] also suggests that model 4 better ($p < 0.001$) fit the expressed emotional engagement than model 3.

We illustrate the interaction between the creation stage and feedback characteristics (actionability, specificity, and valence, which show significant interaction effects) on seekers' expressed emotional engagement in Figure 6. Figure 6a shows that the increase in expressed emotional engagement value for seekers sharing in-progress works is slower ($\beta = -0.007$) than those sharing complete works as the actionability of the feedback rises. Contrarily, as depicted in Figure 6b, when feedback becomes more specific, the increase in expressed emotional engagement value of seekers sharing works-in-progress is more dramatic ($\beta = 0.003$) than those sharing complete works. Figure 6c demonstrates that there exists a lower increase ($\beta = -0.004$) in the expressed emotional engagement value for those sharing in-progress works than sharing complete ones as the valence of feedback increases. In addition, it is suggested that the justification or delay of the feedback makes no significant difference ($p > 0.05$) in the changes of expressed emotional engagement for creators in different creation stages.

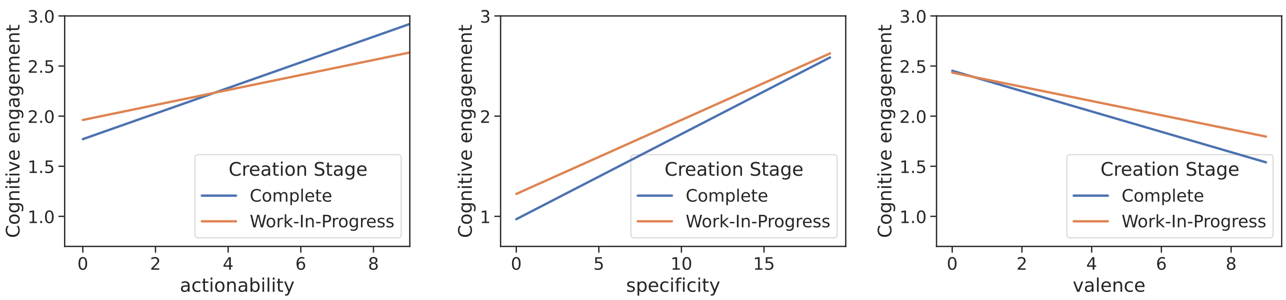
4.3 RQ3. Expressed Cognitive Engagement

Model 5 in Table 4 presents the main effects of the artifact's creation stage and feedback characteristics on seekers' expressed cognitive engagement. Seekers are more likely to express higher cognitive engagement when sharing in-progress works ($\beta = 0.264$) than



(a) Actionability of the feedback **(b) Specificity of the feedback** **(c) Valence of the feedback**

Figure 6: Interaction between creation stage and feedback characteristics on seeker's expressed emotional engagement.



(a) Actionability of the feedback **(b) Specificity of the feedback** **(c) Valence of the feedback**

Figure 7: Interaction between creation stage and feedback characteristics on seeker's expressed cognitive engagement.

sharing complete artifacts, which is in line with a previous finding that creators would balance the cost of revision in the creation process [56]. Moreover, the actionability ($\beta = 0.257$), justification ($\beta = 0.047$), and specificity ($\beta = 0.422$) of the feedback positively predict seekers' expressed cognitive engagement, which confirms findings about design critiques through lab experiments [54]. However, the valence ($\beta = -0.236$) and delay ($\beta = -0.017$) seem to be negative predictors for seekers' expressed cognitive engagement.

Model 6 in Table 4 further combines the interaction effects of the creation stage and feedback characteristics. According to the likelihood-ratio test [79], model 6 provides a significantly better ($p < 0.001$) fit to the expressed cognitive engagement than model 5. We depict the interaction between the creation stage and feedback characteristics (actionability, specificity, and valence, which show significant interaction effects) on seekers' cognitive engagement in Figure 7. Figure 7a shows that as the actionability of feedback goes up, the expressed cognitive engagement grows more slowly ($\beta = -0.091$) for seekers sharing ongoing works than sharing complete works. Similarly, Figure 7b illustrates that increasing the feedback specificity level is usually associated with a lower increase ($\beta = -0.032$) in expressed cognitive engagement for seekers sharing works-in-progress in contrast to sharing complete works. Moreover, Figure 7c demonstrates that when increasing the valence of feedback, the expressed cognitive engagement drops slower ($\beta = 0.036$) for seekers sharing in-progress works than those sharing complete ones. We also observe that seekers sharing works in different creation stages are equally sensitive ($p > 0.05$) to the changes in justification or delay of the feedback.

We also performed the same analysis for each of the community and found two interesting differences. First, we noticed that in r/learnart community (the largest community in this study), the behavioral engagement of seekers sharing in-progress works grows more slowly than those sharing complete works as the actionability level of feedback goes up; however, the opposite trend was observed in r/ArtCrit community (the smallest community in our selected OCCs). Community size could be a possible reason for the difference. Seekers in large communities may be overwhelmed by suggestions from broader providers, and the feedback could be more diverse for in-progress works than completed works [48, 112], making it hard for them to reply to every provider when sharing in-progress works. Another interesting finding is that seekers in the r/FurryArtSchool community (which focuses on furry artworks) tend to express more negative emotional engagement when receiving late feedback, while the effect is not significant in other communities. This finding implies that furry artists are more enthusiastic about timely feedback. Overall, these results suggest that community size and art style might be factors related to seeker's engagement, which could be investigated in the future.

5 DISCUSSION

Our work contributes to the empirical understanding of creators' engagement with online critique in OCCs. We propose methods to quantify and model seekers' engagement levels and predict the artifact's creation stage (code and annotated dataset are available at https://github.com/QingyuGuo/occ_engagement). Through the regression analysis, we add the understanding of how the creation stage of a shared artifact, feedback characteristics, and their interactions may affect seekers' engagement. We find that the main effect of one factor (e.g., artifact's creation stage, valence of feedback) on three dimensions of seeker's engagement may not be in the same direction. The effect of feedback characteristics on seeker's engagement may vary for the artifact in different creation stages. In this section, we discuss the findings and propose design opportunities for OCCs that aim to improve member's creation skills. We also discuss how our analysis workflow may be generalized to other online help-seeking communities, where seekers could go through different stages with varying expectation on the received support.

5.1 Main Effect of Artifact's Creation Stage on Seekers' Engagement

From Model 1 and Model 3 in Table 4, we can see that seekers sharing in-progress works tend to have lower behavioral engagement and express more negative emotional engagement than those sharing complete works. A possible reason is that seekers often feel stressed about sharing their ongoing works due to fear of aggressive providers and a lack of confidence in their skills [48]. Such negative emotions may hinder their motivation to further discuss with critique providers [74, 75]. From Model 5 in Table 4, we also find that seekers who are sharing complete works tend to express lower cognitive engagement than those sharing works-in-progress. This is possible because the cost of revision is typically higher in the complete stage than in-progress works, and creators would be concerned about the cost of revision in accepting the suggestion [56].

5.2 Main Effect of Feedback Characteristics on Seekers' Engagement

Model 5 in Table 4 confirms the previous findings about design critiques through lab experiments [54] where design critiques' actionability, justification, and specificity are positively correlated with the willingness to accept feedback (i.e., expressed cognitive engagement). Model 1 and Model 3 (in Table 4) further show the positive effect of these feedback characteristics on the behavioral and expressed emotional engagement of seekers. A possible explanation is that learners – online creators trying to improve their skills – would feel pleased when they perceive their cognitive engagement is fulfilled [58], and informative comments could encourage participation in the discussion [101]. Moreover, the importance of actionability and specificity of a critique may vary across different dimensions of engagement. In particular, according to model 1 in Table 4, $\beta_{specificity} = 0.153$ is larger than $\beta_{actionability} = 0.12$, indicating that the former might be a stronger factor in predicting the willingness to reply than the latter. Similarly, as shown from model 5 in Table 4, the specificity of the feedback ($\beta_{specificity} = 0.422$) seems to be a more important signal for seeker's expressed cognitive engagement compared with actionability ($\beta_{actionability} = 0.257$). One possible reason is that specificity lowers the barrier for seekers to digest and adapt the feedback [37], and may motivate them to further discuss with the provider [98]; while overly actionable feedback may run the risk of making seekers feel overwhelmed, or even losing the interest of investigating [101]. However, the actionability ($\beta_{actionability} = 0.046$) may be more efficient in increasing the emotional engagement than the specificity of the feedback ($\beta_{specificity} = 0.034$). This may be explained by actionable feedback (i.e., containing more suggestions) often requiring more effort from providers to propose, and seekers may recognize the effort and feel pleased [49]. Overall, our results suggest that different strategies should be adopted to strengthen the associated aspect(s) of online critiques for eliciting a particular dimension of engagement.

In addition, we notice that the feedback valence is positively related to seekers' expressed emotional engagement, which is consistent with the findings about design critiques through lab experiments [70, 72]. Moreover, the valence of the feedback is positively correlated with behavioral engagement, perhaps because encouraging comments would make the learner feel connected and motivate them to acknowledge the support by leaving a response [17]. However, the valence of feedback is negatively correlated with the expressed cognitive engagement. One possible explanation is that negative feedback tend to point out things to improve and thus evokes more reflections [92]. Another potential reason is that negative feedback is considered more trustworthy than positive feedback, thus contributing to more seekers expressing cognitive engagement [21].

Besides the content of the feedback, Model 1 and Model 3 in Table 4 suggest that the delay of the feedback is negatively correlated with the seeker's behavioral and expressed cognitive engagement. On the one hand, the reply behavior in OCCs is consistent with the observation that late answers have a lower reply rate than the timely ones in other generic community question-answering sites [95]. On the other hand, relatively less people expressing cognitive engagement may be explained by the phenomenon of Baby Duck Syndrome [26]. Specifically, if seekers have agreed with some earlier critiques, unless the later feedback is significantly more persuasive than the previous ones, seekers may be relatively reluctant to accept the later ones.

5.3 Interaction Effect between Artifact's Creation Stage and Feedback Characteristics on Seekers' Engagement

According to the results of Model 4 and Model 6 (Table 4), creators sharing works-in-progress have a smaller increase in the expressed emotional and cognitive engagement as the level of actionability of the received critiques goes up than those sharing complete works. A possible cause is that seekers sharing works under development are more sensitive to criticism [48] and may feel anxious when they realize that their works have many deficiencies [3], leading to a more negative shift on expressed emotional engagement than those with finished works. The lower expressed cognitive engagement might be caused by information overload [44]. Specifically, more suggestions (i.e., increased actionability) may cover more diverse topics and have a higher chance of contradicting one another, especially for ongoing works [48]. Therefore, it will be more difficult for seekers sharing in-progress works to merge ideas, resolve conflicts, and prioritize revisions [112], resulting in a lower ratio of suggestions being accepted than those sharing complete works as the actionability of feedback increases.

Model 2 and Model 4 in Table 4 show that elevated specificity in feedback is generally associated with a higher increase in behavioral and expressed emotional engagement for seekers sharing ongoing works than those posting complete ones. This may be because specific comments could decrease seekers' efforts in digesting the feedback compared to the generic ones. Since creators in an ongoing project have a higher tendency to receive ambiguous suggestions [37], they are more likely to feel pleased with clear suggestions and acknowledge them with providers [73]. However, a lower increase in the expressed cognitive engagement is observed as feedback specificity level goes up for creators posting works-in-progress than those publishing complete works. This result confirms a previous finding in a qualitative interview with OCC members that seekers with works under development may also appreciate high-level feedback [105].

Moreover, compared to those in the complete stage (in Table 4), seekers with works-in-progress have a slighter increase in behavioral engagement (Model 2), a smaller positive increase in expressed emotional engagement (Model 4) and a slower decline in expressed cognitive engagement (Model 6) when the valence of the received feedback rises. One possible reason is that seekers sharing works-in-progress may be more suspicious of the sincerity and expertise of the provider who leaves overly positive feedback due to a lack of confidence in their own work [48], and are more likely to reflect on the feedback in contrast to seekers sharing complete works [56].

Model 2 in Table 4 shows that extended delay in feedback is generally associated with a slower drop in the reply probability for seekers posting in-progress works than those sharing complete works. This result contradict a prior research that suggests that seekers hope to get quick feedback for their ongoing works [105]. A possible explanation is that for works-in-progress, creators would like to wait to collect more comprehensive feedback in OCCs [52].

Overall, to fully understand why seekers in different creation stages have different reply behaviors, affective reactions, and willingness to accept feedback as revealed by our quantitative results, future in-depth qualitative interviews with creators in OCCs will be necessary.

5.4 Design Opportunities for Online Critique Communities

Our findings and the engagement modeling methods offer several design opportunities for OCCs to enhance creator engagement with online critiques.

5.4.1 Eliciting Engagement Needs for Seekers based on their Creation Stage. Our results show that seekers sharing works-in-progress are less likely to reply to the critiques; if they reply, their expressed emotional engagement tends to be more negative than those publishing complete works. However, it is beneficial for seekers of ongoing works to participate in discussion threads to gain a deeper understanding of creation alternatives [52]. Meanwhile, a positive affective state could encourage them to complete the artifact creation [104]. Therefore, the OCC services may explore means to motivate members to respond to peer feedback on their works. For example, a chatbot can be developed to help seekers reflect on the received comments. It can first ask seekers whether they have any confusion about the obtained critiques or whether there are suggestions that contradict their initial plans. Then, the chatbot can convert its conversation with seekers into a response template and encourage seekers to build their reply to the provider on top of it. Moreover, if the community detects low emotional engagement expressed in creators' responses to the peer critiques, it may send them comforting messages composed using positive psychology techniques [94] through private channels. One such technique is positive reframing [115], which automatically converts the seeker's original negative text into a positive perspective (e.g., from "*I am sad that there are many problems with the work*" to *"It is good that I can see the problems in my work and I will fix them to make my work even better!"*).

5.4.2 Facilitating Feedback Writing for Providers. In OCCs, although providers are encouraged to offer constructive feedback rather than simple praise, it is often hard for them to know whether their input could meet the community expectations or requirements, especially when they are unfamiliar with the community norm or new to creation [64]. Moreover, our analysis shows that the characteristics of critique that could engage seekers vary when the target artifacts are in different creation stages, which may make it more difficult for providers to understand seekers' underlying needs and tailor their feedback accordingly. Therefore, OCCs can develop technologies that could guide providers in feedback writing. Similar to CritiqueKit [69], OCCs can offer providers an ambient awareness of their feedback quality and the extent to which they satisfy seekers' needs at different creation stages. To this end, a computational system could be proposed to analyze and visualize the characteristics of providers' comments (such as the actionability and valence) on-the-fly with a comparison to the community norms and adaptive expert suggestions, allowing providers to adjust their input wherever they see fit. Such a system can also take the form of a chatbot similar to the MepsBot system [76], which can predict seekers' engagement levels with providers' proposed critiques and suggest improvement opportunities of the feedback. Compared with the original MepsBot system developed for the online health community, extra information (e.g., content of the critique-seeking post, the creation stage or the visual information of the artifact, etc.) should be considered to precisely infer seekers' engagement with peer critiques in OCCs. Alternatively, an automatic rewriting tool could be developed to adapt providers' initial draft of feedback to the targeting artifact, such as by tailoring the content to adjust its actionability, specificity, and valence. Such techniques might be implemented with a similar architecture to PARTNER proposed by Sharma et al. [93], which utilizes a reinforcement-based model that could replace low-quality sentences with high-quality ones while maintaining the sentence coherence and context specificity.

5.4.3 Allocating Community Resources for Moderator. Our results show that seekers posting ongoing works are more likely to reply to late feedback than those in the complete stage. The result indicates that these creators are inclined to continue monitoring peer feedback as their works are still underway. Therefore, OCC moderators can ask creators whose works have been completed whether they have gotten enough critiques. If the answer is yes, moderators can decrease the post recommendation priority or even close the discussion threads with the consent of seekers and increase the priority for those critique-seeking posts containing in-progress works. Moreover, the relative importance of content features (e.g., specificity, actionability) varies for different creation stages. In OCCs, providers could have different commenting habits and abilities. For instance, some may be good at delivering specific examples, while others may prefer to offer high-level suggestions. Moderators can thus invite particular providers to comment on an artifact according to their historical responses to critique-seeking posts in different creation stages. The provider allocation process may also be automated by developing a recommender system based on the community interaction (i.e., posting and commenting) data [4, 100]. However, the workload balance for each provider should be considered to avoid attrition to community members.

5.5 Generalizing the Concept of Stage in Online Help-seeking Communities

One salient contribution of the work is that we explicitly investigate how the creation stage of the shared artifact may affect seeker's engagement in online feedback exchange. Since many seekers may not explicitly articulate their expectations when soliciting feedback, our insights would be beneficial for providers to tailor their feedback according to the artifact's creation stage. In fact, this issue is not unique to OCCs. Research in other online help-seeking communities (such as online communities for patients[23, 60], parents [9, 33], and graduate students [36]) find that, seekers may go through different stages, and they would have distinct expectations of support in various stages. However, seekers may also not clearly demonstrate their need for support, often due to social stigma and a lack of experience [2, 38, 109]. Specifically, in online health communities, cancer patients could be in different treatment stages (i.e., pretreatment, receiving treatment, no evidence of disease, and chronic care), which may vary their expectation on the received support [41, 60]. In online parenting communities, children of parents could also be in different life stages, ranging from infancy to preschool, and the received support should be customized to the stage of their children [9, 33]. In online communities designed for graduate students, students could face different challenges and expect distinct support before, during, and after their graduations [36].

Therefore, it would be promising for researchers in other online help-seeking communities to consider the seeker's stage as one factor that may affect seeker's engagement. It should be noted that the definition of "stage", factors of provider's comment, and measurement of seeker's engagement need to be tailored to the specific communities. Take the cancer-oriented online health community as an example. One possible solution is to first identify the treatment stage of a seeker using the model proposed by Levonian et al. [60], then quantitatively examine the effect of the life stage, received social support (such as informational/emotional support) [107, 108], and their interactions to seeker's expressed satisfaction [77, 99]. Using a similar analysis approach in this work, design opportunities could be derived to support members (such as seekers, providers, and moderators) in other help-seeking scenarios.

6 LIMITATIONS AND FUTURE WORK

Several limitations exist in the presented work. First, the findings are correlational but not causal. Although we quantitatively present the significance of the correlation between independent variables (i.e., artifact's creation stage, feedback characteristics, and their interactions) to seekers' engagement, the causality between these independent variables and seekers' engagement cannot be ensured without conducting random-assignment experiments. Moreover, the reason behind such differences remains unclear, which could be answered in future studies by qualitatively interviewing creators. Second, as mentioned in Section 3.2.3, since seekers rarely explicitly indicated that they had followed the suggestions (∼3% of the annotated replies), we combined both the planned and executed action as the high cognitive engagement to obtain a stable model for analysis. We encourage future work to develop computational models to recognize a fine-grained feedback acceptance level (e.g., by adding more training data of the executed action), and quantitatively explore what factors are related to such behaviors on large-scale data. Meanwhile, in this work we quantify seeker's expressed engagement by following methods in previous studies [77, 99], and the proxy of engagement could shed light on adapting feedback to the seeker's creation stage. However, it is possible that the actual engagement may not be exactly the same as what they expressed. Future studies can thus explore other direct measurements of seekers' engagement, such as seeker-accepted answers in community question-answering sites [95]. Third, we classify the shared artifact's creation stage using only the critique-seeking posts' textual information. Since we conducted the analysis based on large-scale archived data, the retrieved links to the artifacts may have been broken due to the artifacts having been deleted by the seekers, while selection bias could be induced by only considering posts with downloadable images. Although the model achieves reasonable performance, future studies can further consider the visual information by collecting posts in real time [16] to improve the classification performance. Moreover, according to our analysis on each selected community in Section 4.3, the community norm (such as the community size and artifact's style) may also be related to seekers' engagement. While we observed that seekers rarely receive suggestions in relatively small communities, it would be interesting to investigate how seekers engage with critiques in niche communities and compare findings in this work. In addition, previous qualitative work suggested that seeker's skill levels might also affect their engagement with the received feedback [105]. We suggest that future work can complement our research with more diverse factors related to OCCs. Finally, we analyzed seeker's engagement with the received comments within a single critique-seeking post. While in OCCs, seekers may initiate multiple posts for one artifact during the creation [48]. Future studies could investigate how feedback in historic posts may affect seeker's engagement with feedback for the iteration of an artifact.

In this work, our research sites are mainly about art-related visual artifacts, similar to the visual artifacts discussed in previous studies on OCCs [48, 105, 111]. Meanwhile, diverse categories of creativity exist (such as writing [87], dance [83], and music [86]) and may affect the community norms and the creator's engagement [97]. To extend our research and examine the generalizability of our results in other forms of OCCs, the following issues in the analysis workflow need to be considered. First, in art-related OCCs, the presence of an image is an indicator of critique-seeking posts. When investigating other types of OCCs, a classifier is needed to identify critique-seeking posts. Furthermore, although the binary classification of the creation stage is a general framework that can be applied to multiple types of OCCs, we encourage future studies to tailor the creation stage for specific communities [62].

7 CONCLUSION

This paper presents a comprehensive analysis and understanding of the role of the artifact's creation stage, feedback characteristics, and their interactions with creators' engagement (i.e., behavioral, expressed emotional and cognitive engagement) using data from four art-related OCCs. We developed methods and deep learning-based models to quantify creators' engagement and predict the creation stage of their shared artifacts. We also extracted content and timing-related features to characterize online critiques. Through regression analysis, we found that creators sharing works-in-progress generally present lower behavioral and emotional engagement, but higher cognitive engagement. Although the increase in feedback valence is associated with an increase in behavior engagement for seekers posting both works-in-progress and complete works, this increase is stronger for complete works. This study advances the empirical understanding of how creators' engagement may vary with the creation stage and feedback characteristics, and provides practical implications for enhancing creators' engagement in OCCs. It also offers a theoretical contribution for other online help-seeking platforms by highlighting the importance of considering the seeker's stage in support exchange.

ACKNOWLEDGMENTS

Many thanks to the anonymous reviewers for their insightful suggestions. We thank Jianben He, Liwenhan Xie, Ruijia Cheng, and Linping Yuan for their valuable input and discussion. This research was supported in part by HKUST-WeBank Joint Laboratory Project Grant No.: WEB19EG01-d.

REFERENCES

[1] Jay S Albanese. 2009. *Combating Piracy: Intellectual Property Theft and Fraud.* Routledge, London, UK.
[2] Nazanin Andalibi and Andrea Forte. 2018. Announcing pregnancy loss on Facebook: A decision-making framework for stigmatized disclosures on identified social network sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA, Article 158, 14 pages.
[3] GD Atlas, B Fassett, and C Peterson. 1994. Sensitivity to criticism and depressive symptoms. *Journal of Social Behavior and Personality* 9, 2 (1994), 301–316.
[4] Vipin Balachandran. 2013. Reducing human effort and improving quality in peer code reviews using automatic static analysis and reviewer recommendation. In *Proceedings of the International Conference on Software Engineering.* IEEE, New York, NY, USA, 931–940.
[5] Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The pushshift reddit dataset. *Proceedings of the International AAAI Conference on Web and Social Media* 14 (2020), 830–839.
[6] Chris Beaumont, Ella O'Doherty, and Lee Shannon. 2011. Reconceptualising Assessment Feedback: A Key to Improving Student Learning? *Studies in Higher Education* 36 (2011), 671–687.
[7] William D Berry, Stanley Feldman, and Dr Stanley Feldman. 1985. *Multiple regression in practice.* SAGE, Thousand Oaks, CA, USA.
[8] Viv Bewick, Liz Cheek, and Jonathan Ball. 2005. Statistics review 14: Logistic regression. *Critical Care* 9, 1, Article 112 (2005), 7 pages.
[9] Neelma Bhatti, Amarachi Blessing Mbakwe, Sandra Nnadi, Geetha Saarunya Clarke, Aakash Gautam, D Scott McCrickard, and Aisling Kelliher. 2022. Intimate Narratives: An Assets-Based Approach to Develop Holistic Perspectives of Student Mothers' Lives and Their Use of Technology in Parenting. *Proceedings of the ACM on Human-Computer Interaction* 6, Article 522 (2022), 28 pages. Issue CSCW2.
[10] Richard Bounds, Colleen Bush, Amish Aghera, Nestor Rodriguez, R. Brent Stansfield, Sally A. Santen, and MERC at CORD Feedback Study Group. 2013. Emergency medicine residents' self-assessments play a critical role when receiving feedback. *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine* 20, 10 (2013), 1055–1061.
[11] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. In *APA Handbook of Research Methods in Psychology, Vol 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological.* APA, Washington, DC, USA, 57–71.
[12] Deirdre Burke. 2009. Strategies for using feedback students bring to higher education. *Assessment & Evaluation in Higher Education* 34, 1 (2009), 41–50.
[13] Rich Caruana, Steve Lawrence, and C Giles. 2000. Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. *Advances in Neural Information Processing Systems* 13 (2000), 381–387.
[14] Daniela Castellanos-Reyes. 2021. The dynamics of a MOOC's learner-learner interaction over time: A longitudinal network analysis. *Computers in Human Behavior* 123, Article 106880 (2021), 14 pages.
[15] John M Chambers. 2017. Linear models. In *Statistical models in S.* Routledge, London, UK, 95–144.
[16] Jackie Chan, Aditi Atreyasa, Stevie Chancellor, and Eshwar Chandrasekharan. 2022. Community Resilience: Quantifying the Disruptive Effects of Sudden Spikes in Activity within Online Communities. In *Proceedings of the ACM Conference Extended Abstracts on Human Factors in Computing Systems.* ACM, New York, NY, USA, Article 339, 8 pages.
[17] Diane D Chapman, Julia Storberg-Walker, and Sophia J Stone. 2008. Hitting reply: A qualitative study to understand student decisions to respond to online discussion postings. *E-Learning and Digital Media* 5, 1 (2008), 29–39.
[18] Ruijia Cheng and Jenna Frens. 2022. Feedback Exchange and Online Affinity: A Case Study of Online Fanfiction Writers. *Proceedings of the ACM on Human-Computer Interaction* 6, Article 402 (2022), 29 pages. Issue CSCW2.
[19] Ruijia Cheng, Ziwen Zeng, Maysnow Liu, and Steven Dow. 2020. Critique Me: Exploring How Creators Publicly Request Feedback in an Online Critique Community. *Proceedings of the ACM on Human-Computer Interaction* 4, Article 161 (2020), 24 pages. Issue CSCW2.
[20] Michelene T. H. Chi and Ruth Wylie. 2014. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist* 49, 4 (2014), 219–243.
[21] Lerita M Coleman, Lee Jussim, and Jack Abraham. 1987. Students' Reactions to Teachers' Evaluations: The Unique Impact of Negative Feedback 1. *Journal of Applied Social Psychology* 17, 12 (1987), 1051–1070.
[22] Patrick A Crain and Brian P Bailey. 2017. Share Once or Share Often? Exploring How Designers Approach Iteration in a Large Online Community. In *Proceedings of the SIGCHI Conference on Creativity and Cognition.* ACM, New York, NY, USA, 80–92.
[23] Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. 2016. Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA, 2098–2110.
[24] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers).* ACL, Minneapolis, MN, USA, 4171–4186.
[25] Kevin Doherty and Gavin Doherty. 2018. Engagement in HCI: conception, theory and measurement. *Comput. Surveys* 51, 5, Article 99 (2018), 39 pages.
[26] Gokhan Egri and Coskun Bayrak. 2014. The role of search engine optimization on keeping the user on the site. *Procedia Computer Science* 36 (2014), 335–342.
[27] Kevin W. Eva, Heather Armson, Eric Holmboe, Jocelyn Lockyer, Elaine Loney, Karen Mann, and Joan Sargeant. 2012. Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. *Advances in Health Sciences Education: Theory and Practice* 17, 1 (2012), 15–26.
[28] Edmund Burke Feldman. 1994. *Practical art criticism.* Prentice Hall, Englewood Cliffs, NJ, USA.
[29] Christiane Fellbaum and George A. Miller. 1998. *WordNet: An Electronic Lexical Database.* A Bradford Book, Cambridge, MA, USA.
[30] Eureka Foong, Steven P Dow, Brian P Bailey, and Elizabeth M Gerber. 2017. Online feedback exchange: A framework for understanding the socio-psychological factors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA, 4454–4467.
[31] Jennifer A Fredricks, Phyllis C Blumenfeld, and Alison H Paris. 2004. School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research* 74, 1 (2004), 59–109.
[32] Devin Gaffney and J Nathan Matias. 2018. Caveat emptor, computational social science: Large-scale missing data in a widely-published Reddit corpus. *PloS one* 13, 7, Article e0200162 (2018), 13 pages.
[33] Yujia Gao, Jinu Jang, and Diyi Yang. 2021. Understanding the usage of online media for parenting from infancy to preschool at scale. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA, Article 183, 12 pages.
[34] Shuang Geng, Ben Niu, Yuanyue Feng, and Miaojia Huang. 2020. Understanding the focal points and sentiment of learners in MOOC reviews: A machine learning and SC-LIWC-based approach. *British Journal of Educational Technology* 51, 5 (2020), 1785–1803.
[35] Pollyanna Goncalves, Fabrıcio Benevenuto, and Meeyoung Cha. 2013. PANAS-t: A Pychometric Scale for Measuring Sentiments on Twitter. (2013), 11. arXiv:1308.1857

[36] Crystal Gong, Koustuv Saha, and Stevie Chancellor. 2021. "The Smartest Decision for My Future": Social Media Reveals Challenges and Stress During Post-College Life Transition. *Proceedings of the ACM on Human-Computer Interaction* 5, Article 298 (2021), 29 pages. Issue CSCW2.

[37] Jodi S Goodman and Robert E Wood. 2004. Feedback specificity, learning opportunities, and learning. *Journal of Applied Psychology* 89, 5 (2004), 809.

[38] Qingyu Guo, Siyuan Zhou, Yifeng Wu, Zhenhui Peng, and Xiaojuan Ma. 2022. Understanding and Modeling Viewers' First Impressions with Images in Online Medical Crowdfunding Campaigns. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 361, 20 pages.

[39] Wenge Guo, Ye Chen, Jing Lei, and Yan Wen. 2014. The Effects of Facilitating Feedback on Online Learners' Cognitive Engagement: Evidence from the Asynchronous Online Discussion. *Education Sciences* 4, 2 (2014), 193–208.

[40] Karen Handley, Margaret Price, and Jill Millar. 2011. Beyond 'doing time': investigating the concept of student engagement with feedback. *Oxford Review of Education* 37, 4 (2011), 543–560.

[41] Gillian R Hayes, Gregory D Abowd, John S Davis, Marion L Blount, Maria Ebling, and Elizabeth D Mynatt. 2008. Opportunities for pervasive computing in chronic cancer care. In *International Conference on Pervasive Computing*. Springer, Berlin, Germany, 262–279.

[42] Catherine M Hicks, Vineet Pandey, C Ailie Fraser, and Scott Klemmer. 2016. Framing feedback: Choosing review environment features that support high quality peer assessment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 458–469.

[43] Clayton Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media* 8, 1 (2014), 216–225.

[44] Jacob Jacoby. 1984. Perspectives on information overload. *Journal of Coonsumer Research* 10, 4 (1984), 432–435.

[45] Shazia K. Jan, Panos Vlachopoulos, and Mitch Parsell. 2019. Social Network Analysis and Online Learning Communities in Higher Education: A Systematic Literature Review. *Online Learning* 23, 1 (2019), 249–265.

[46] Menglin Jia, Zuxuan Wu, Austin Reiter, Claire Cardie, Serge Belongie, and Ser-Nam Lim. 2021. Exploring Visual Engagement Signals for Representation Learning. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, Montreal, QC, Canada, 4186–4197.

[47] Anders Jonsson. 2013. Facilitating productive use of feedback in higher education. *Active Learning in Higher Education* 14, 1 (2013), 63–76. SAGE.

[48] Joy Kim, Maneesh Agrawala, and Michael S Bernstein. 2017. Mosaic: designing online creative communities for sharing works-in-progress. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, New York, NY, USA, 246–258.

[49] Ran Kivetz and Itamar Simonson. 2002. Earning the right to indulge: Effort as a determinant of customer preferences toward frequency program rewards. *Journal of Marketing Research* 39, 2 (2002), 155–170.

[50] Yasmine Kotturi, Allie Blaising, Sarah E Fox, and Chinmay Kulkarni. 2021. The Unique Challenges for Creative Small Businesses Seeking Feedback on Social Media. *Proceedings of the ACM on Human-Computer Interaction* 5, Article 15 (2021), 27 pages. Issue CSCW1.

[51] Yasmine Kotturi and McKayla Kingston. 2019. Why do Designers in the "Wild" Wait to Seek Feedback until Later in their Design Process?. In *Proceedings of the SIGCHI Conference on Creativity and Cognition*. ACM, New York, NY, USA, 541–546.

[52] Yubo Kou and Colin M Gray. 2017. Supporting distributed critique through interpretation and sense-making in an online creative community. *Proceedings of the ACM on Human-Computer Interaction* 1, Article 60 (2017), 18 pages. Issue CSCW.

[53] Yubo Kou and Colin M Gray. 2018. "What do you recommend a complete beginner like me to practice" Professional Self-Disclosure in an Online Community. *Proceedings of the ACM on Human-Computer Interaction* 2, Article 94 (2018), 24 pages. Issue CSCW.

[54] Markus Krause, Tom Garncarz, JiaoJiao Song, Elizabeth M. Gerber, Brian P. Bailey, and Steven P. Dow. 2017. Critique Style Guide: Improving Crowdsourced Design Feedback with a Natural Language Model. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 4627–4639.

[55] Sneha R. Krishna Kumaran, Deana C. McDonagh, and Brian P. Bailey. 2017. Increasing Quality and Involvement in Online Peer Feedback Exchange. *Proceedings of the ACM on Human-Computer Interaction* 1, Article 63 (2017), 18 pages. Issue CSCW.

[56] Sneha R. Krishna Kumaran, Wenxuan Wendy Shi, and Brian P Bailey. 2021. Am I Ready to Get Feedback? A Taxonomy of Factors Creators Consider Before Seeking Feedback on In-Progress Creative Work. In *Proceedings of the SIGCHI Conference on Creativity and Cognition*. ACM, New York, NY, USA, Article 31, 10 pages.

[57] Hui-Min Lai and Tsung Teng Chen. 2014. Knowledge sharing in interest online communities: A comparison of posters and lurkers. *Computers in Human Behavior* 35 (2014), 295–306.

[58] Eun-Jung Lee. 2016. How perceived cognitive needs fulfillment affect consumer attitudes toward the customized product: The moderating role of consumer knowledge. *Computers in Human Behavior* 64 (2016), 152–162.

[59] Fritz Lekschas, Spyridon Ampanavos, Pao Siangliulue, Hanspeter Pfister, and Krzysztof Z. Gajos. 2021. Ask Me or Tell Me? Enhancing the Effectiveness of Crowdsourced Design Feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 564, 12 pages.

[60] Zachary Levonian, Drew Richard Erikson, Wenqi Luo, Saumik Narayanan, Sabirat Rubya, Prateek Vachher, Loren Terveen, and Svetlana Yarosh. 2020. Bridging qualitative and quantitative methods for user modeling: Tracing cancer patient behavior in an online health community. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 14. ACM, New York, NY, USA, 405–416.

[61] Sannyuya Liu, Shiqi Liu, Zhi Liu, Xian Peng, and Zongkai Yang. 2022. Automated detection of emotional and cognitive engagement in MOOC discussions to predict learning achievement. *Computers & Education* 181 (2022), 104461.

[62] Xiaojuan Ma, Li Yu, Jodi L Forlizzi, and Steven P Dow. 2015. Exiting the design studio: Leveraging online participants for early-stage design feedback. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, New York, NY, USA, 676–685.

[63] Dhruv Mahajan, Ross Girshick, Vignesh Ramanathan, Kaiming He, Manohar Paluri, Yixuan Li, Ashwin Bharambe, and Laurens van der Maaten. 2018. Exploring the Limits of Weakly Supervised Pretraining. In *Computer Vision – ECCV 2018*. Vol. 11206. Springer, Cham, 185–201.

[64] Jennifer Marlow and Laura Dabbish. 2014. From rookie to all-star: professional development in a graphic design social networking site. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, New York, NY, USA, 922–933.

[65] Rosa M. Mayordomo, Anna Espasa, Teresa Guasch, and Montserrat Martínez-Melo. 2022. Perception of online feedback and its impact on cognitive and emotional engagement with feedback. *Education and Information Technologies* 27 (2022), 7947–7971.

[66] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction* 3, Article 72 (2019), 23 pages. Issue CSCW.

[67] Mary L McHugh. 2012. Interrater reliability: the kappa statistic. *Biochemia Medica* 22, 3 (2012), 276–282.

[68] Pedro Manuel Moreno-Marcos, Carlos Alario-Hoyos, Pedro J. Muñoz-Merino, Iria Estévez-Ayres, and Carlos Delgado Kloos. 2018. Sentiment analysis in MOOCs: A case study. In *2018 IEEE Global Engineering Education Conference*. IEEE, New York, NY, USA, 1489–1496.

[69] Tricia J. Ngoon, C. Ailie Fraser, Ariel S. Weingarten, Mira Dontcheva, and Scott Klemmer. 2018. Interactive Guidance Techniques for Improving Creative Feedback. In *CHI*. ACM, New York, NY, USA, Article 155, 11 pages.

[70] Thi Thao Duyen T. Nguyen, Thomas Garncarz, Felicia Ng, Laura A. Dabbish, and Steven P. Dow. 2017. Fruitful Feedback: Positive Affective Language and Source Anonymity Improve Critique Reception and Work Outcomes. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, New York, NY, USA, 1024–1034.

[71] Yeonjoo Oh, Suguru Ishizaki, Mark D. Gross, and Ellen Yi-Luen Do. 2013. A theoretical framework of design critiquing in architecture studios. *Design Studies* 34, 3 (2013), 302–325.

[72] Jonas Oppenlaender, Elina Kuosmanen, Andrés Lucero, and Simo Hosio. 2021. Hardhats and bungaloos: Comparing crowdsourced design feedback with peer design feedback in the classroom. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 570, 14 pages.

[73] Reinhard Pekrun, Thomas Goetz, Wolfram Titz, and Raymond P Perry. 2002. Positive emotions in education. In *Beyond coping: Meeting goals, visions, and challenges*. Oxford University Press, Oxford, UK, 149–173.

[74] Reinhard Pekrun, Stephanie Lichtenfeld, Herbert W Marsh, Kou Murayama, and Thomas Goetz. 2017. Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development* 88, 5 (2017), 1653–1670.

[75] Reinhard Pekrun and Lisa Linnenbrink-Garcia. 2012. Academic emotions and student engagement. In *Handbook of research on student engagement*. Springer, Berlin, Germany, 259–282.

[76] Zhenhui Peng, Qingyu Guo, Ka Wing Tsang, and Xiaojuan Ma. 2020. Exploring the effects of technological writing assistance for support providers in online mental health community. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 566, 15 pages.

[77] Zhenhui Peng, Xiaojuan Ma, Diyi Yang, Ka Wing Tsang, and Qingyu Guo. 2021. Effects of support-seekers' community knowledge on their expressed satisfaction with the received comments in mental health communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

ACM, New York, NY, USA, Article 536, 12 pages.
[78] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The development and psychometric properties of LIWC2015*. Technical Report. The University of Texas at Austin.
[79] Bercedis Peterson and Frank E Harrell Jr. 1990. Partial proportional odds models for ordinal response variables. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 39, 2 (1990), 205–217.
[80] Margaret Price, Karen Handley, Jill Millar, and Berry O'Donovan. 2010. Feedback: all that effort, but what is the effect? *Assessment & Evaluation in Higher Education* 35, 3 (2010), 277–289.
[81] Yada Pruksachatkun, Sachin R Pendse, and Amit Sharma. 2019. Moments of change: Analyzing peer-based cognitive support in online mental health forums. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 64, 13 pages.
[82] Reddit. 2022. r/ArtCrit. https://www.reddit.com/r/ArtCrit/. Retrieved on Sep 1, 2022.
[83] Reddit. 2022. r/Dance. https://www.reddit.com/r/Dance/. Retrieved on Sep 1, 2022.
[84] Reddit. 2022. Reddit Platform. https://www.reddit.com/. Retrieved on Sep 1, 2022.
[85] Reddit. 2022. r/learnart. https://www.reddit.com/r/learnart/. Retrieved on Sep 1, 2022.
[86] Reddit. 2022. r/MusicCritique. https://www.reddit.com/r/MusicCritique/. Retrieved on Sep 1, 2022.
[87] Reddit. 2022. r/writing. https://www.reddit.com/r/writing/. Retrieved on Sep 1, 2022.
[88] Kathryn E Ringland, Arpita Bhattacharya, Kevin Weatherwax, Tessa Eagle, and Christine T Wolf. 2022. ARMY's Magic Shop: Understanding the Collaborative Construction of Playful Places in Online Communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 126, 19 pages.
[89] Howard Risatti. 1987. Art Criticism in Discipline-Based Art Education. *Journal of Aesthetic Education* 21, 2 (1987), 217–25.
[90] Jerome I. Rotgans and Henk G. Schmidt. 2011. Cognitive engagement in the problem-based learning classroom. *Advances in Health Sciences Education* 16, 4 (2011), 465–479.
[91] D. Royce Sadler. 1989. Formative assessment and the design of instructional systems. *Instructional Science* 18, 2 (1989), 119–144.
[92] Joan M Sargeant, Karen V Mann, Cees P Van der Vleuten, and Job F Metsemakers. 2009. Reflection: a link between receiving and using assessment feedback. *Advances in Health Sciences Education* 14, 3 (2009), 399–410.
[93] Ashish Sharma, Inna W Lin, Adam S Miner, David C Atkins, and Tim Althoff. 2021. Towards facilitating empathic conversations in online mental health support: A reinforcement learning approach. In *Proceedings of the Web Conference*. ACM, New York, NY, USA, 194–205.
[94] C Estelle Smith, William Lane, Hannah Miller Hillberg, Daniel Kluver, Loren Terveen, and Svetlana Yarosh. 2021. Effective Strategies for Crowd-Powered Cognitive Reappraisal Systems: A Field Deployment of the Flip* Doubt Web Application for Mental Health. *Proceedings of the ACM on Human-Computer Interaction* 5, Article 417 (2021), 37 pages. Issue CSCW2.
[95] Ivan Srba and Maria Bielikova. 2016. A comprehensive survey and classification of approaches for community question answering. *ACM Transactions on the Web (TWEB)* 10, 3, Article 18 (2016), 63 pages.
[96] Bart Thomee, David A Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. 2016. YFCC100M: The new data in multimedia research. *Commun. ACM* 59, 2 (2016), 64–73.
[97] Kerrie Unsworth. 2001. Unpacking creativity. *Academy of management review* 26, 2 (2001), 289–297.
[98] Don VandeWalle and Larry L Cummings. 1997. A test of the influence of goal orientation on the feedback-seeking process. *Journal of Applied Psychology* 82, 3 (1997), 390–400.
[99] Tatiana A Vlahovic, Yi-Chia Wang, Robert E Kraut, and John M Levine. 2014. Support matching and satisfaction in an online breast cancer support community. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1625–1634.
[100] Xianzhi Wang, Chaoran Huang, Lina Yao, Boualem Benatallah, and Manqing Dong. 2018. A survey on expert recommendation in community question answering. *Journal of Computer Science and Technology* 33, 4 (2018), 625–653.
[101] Naomi E. Winstone, Robert A. Nash, Michael Parker, and James Rowntree. 2017. Supporting Learners's Agentic Engagement With Feedback: A Systematic Review and a Taxonomy of Recipience Processes. *Educational Psychologist* 52, 1 (2017), 17–37.
[102] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. ACL, Dublin, Ireland, 38–45.
[103] Y. Wayne Wu and Brian P. Bailey. 2017. Bitter Sweet or Sweet Bitter? How Valence Order and Source Identity Influence Feedback Acceptance. In *Proceedings of the SIGCHI Conference on Creativity and Cognition*. ACM, New York, NY, USA, 137–147.
[104] Y. Wayne Wu and Brian P. Bailey. 2018. Soften the Pain, Increase the Gain: Enhancing Users' Resilience to Negative Valence Feedback. *Proceedings of the ACM on Human-Computer Interaction* 2, Article 186 (2018), 20 pages. Issue CSCW.
[105] Anbang Xu and Brian Bailey. 2012. What do you think? A case study of benefit, expectation, and interaction in a large online critique community. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, New York, NY, USA, 295–304.
[106] Bing Xu, Nian-Shing Chen, and Guang Chen. 2020. Effects of teacher role on student engagement in WeChat-Based online discussion learning. *Computers & Education* 157, Article 103956 (2020), 11 pages.
[107] Diyi Yang, Robert Kraut, and John M Levine. 2017. Commitment of newcomers and old-timers to online health support communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 6363–6375.
[108] Diyi Yang, Zheng Yao, and Robert Kraut. 2017. Self-disclosure and channel difference in online health support groups. *Proceedings of the International AAAI Conference on Web and Social Media* 11, 1 (2017), 704–707.
[109] Wenjie Yang, Zhiyang Wu, Nga Yiu Mok, and Xiaojuan Ma. 2022. How to Save Lives with Microblogs? Lessons From the Usage of Weibo for Requests for Medical Assistance During COVID-19. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 70, 18 pages.
[110] Yu-Chun (Grace) Yen. 2017. Enhancing the Usage of Crowd Feedback for Iterative Design. In *Proceedings of the SIGCHI Conference on Creativity and Cognition*. ACM, New York, NY, USA, 513–517.
[111] Yu-Chun Grace Yen, Steven P. Dow, Elizabeth Gerber, and Brian P. Bailey. 2017. Listen to Others, Listen to Yourself: Combining Feedback Review and Reflection to Improve Iterative Design. In *Proceedings of the SIGCHI Conference on Creativity and Cognition*. ACM, New York, NY, USA, 158–170.
[112] Yu-Chun Grace Yen, Joy O Kim, and Brian P Bailey. 2020. Decipher: an interactive visualization tool for interpreting unstructured design feedback from multiple providers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 253, 13 pages.
[113] Alvin Yuan, Kurt Luther, Markus Krause, Sophie Isabel Vennix, Steven P Dow, and Bjorn Hartmann. 2016. Almost an Expert: The Effects of Rubrics and Expertise on Perceived Value of Crowdsourced Design Critiques. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, New York, NY, USA, 1005–1017.
[114] Haiyi Zhu, Amy Zhang, Jiping He, Robert E Kraut, and Aniket Kittur. 2013. Effects of peer feedback on contribution: a field experiment in Wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 2253–2262.
[115] Caleb Ziems, Minzhi Li, Anthony Zhang, and Diyi Yang. 2022. Inducing Positive Perspectives with Text Reframing. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. ACL, Dublin, Ireland, 3682–3700.

A APPENDIX

Table 5 introduces the features utilized in the machine learning models for predicting the level of expressed cognitive engagement.

Table 5: List of features used in the machine learning models for predicting the level of expressed cognitive engagement.

Feature Name	Processing Method	Note / Example
73 LIWC features	Counting word frequency that falls under each category of LIWC dictionary words. We select all the provided categories of the LIWC 2015 dictionary. These categories include 21 linguistic dimensions (e.g., pronouns, verbs), 41 psychological processes dimensions (e.g., words demonstrating positive emotion, cognitive processes), 6 personal concerns dimensions (e.g., home), and 5 informal language dimensions (e.g., swear). Please check Table 1 in [78] for more details.	For example, in the reply "I love the suggestion!", the word frequency of the "positive emotion" category is 1, as "love" is one of the keywords in that category. Similarly, we can calculate the word frequency in all the other categories.
Number of words	use word_tokenize() method from NLTK library	number of words in the reply after removing URLs
Number of sentences	use sent_tokenize() method from NLTK library	number of sentences in the reply after removing URLs
Number of characters	len(content_without_url.replace(" ", ""))	number of characters without spaces after removing URLs
Number of URLs	re.findall(r"http\S+", content)	number of the URLs in the reply
Number of emojis	use emoji_count() method from emoji library	number of emojis in the reply
Number of numerics	use isdigit() method	number of numerics in the reply after removing URLs