Enabling Expert Critique at Scale with Chatbots and Micro Guidance

Carlos Toxtli

West Virginia University
Morgantown, United States
email: carlos.toxtli@mail.wvu.edu

Saiph Savage

West Virginia University
Morgantown, United States
email: saiph.savage@mail.wvu.edu

Abstract—Critique is important to improve creative work and help learners of design to grow. The "gold standard" of critique involves in-person discussion with experts who provide feedback. However, scaling expert critique is difficult as experts are scarce, have limited time and privacy concerns. Online alternatives, such as forums, rarely facilitate specialized critique. To enable at scale access to expert critique, we present Micro Apprenticeship Through Tutorials (MATT), a chatbot that micro-guides experts to critique in short bursts of time. This empowers more experts to critique as the activity becomes more accessible to their busy schedules. MATT's "bot aspect" also provides a mediated form of communication between experts and learners, helping to address experts' privacy concerns. Additionally, MATT helps to delegate critique work to experts in a way that can match experts' and learners' time constraints. We conduct a field experiment comparing MATT to current alternatives. We find that, contrary to other approaches, MATT's conversational micro-guidance facilitates leading a large number of experts to critique learners' creative work. We conclude by providing data-backed design implications to empower and facilitate at scale collaborations between experts and learners.

Keywords–Chatbot; Mediated communication; feedback; experts.

I. INTRODUCTION

Feedback is essential to creative work. Creators can receive many kinds of feedback for their work, from informal reactions/kudos to more detailed, critical analyses. *Critique* is the most prestigious type of feedback a creator can receive because this feedback can truly help the person to improve their work. Critique is characterized by (1) identifying decisions made in the creative piece being analyzed; (2) relating those decisions to best practices; (3) and then describing how and why the decisions made support (or not) the best practices [1]. Critique is especially enhanced when done by experts who can more easily discuss the state of the art and connect the work to impactful societal outcomes [2], [3].

Critique directly enhances creative work, and also helps the creators to learn new techniques and methods [4]. Critique is starting to be considered one of the most effective learning strategies [5].

In Section 2, we present how experts have historically provided critique to creative work within physical studios where experts were directly collocated with creators [6], individuals whom experts had usually never met before. Being physically together in a space with strangers helped experts to provide structured, spontaneous, open feedback, and facilitated an efficient exchange of information [7]. However, getting experts and creators together at the same time in one same physical space is hard [8]. Experts generally have limited time, complex schedules, and are distributed across the globe [9].

To overcome these difficulties, online platforms have emerged to support and act as a companion to physical studios [1]. These platforms aim to facilitate communication between experts and creators (who, in these settings, are considered to be "learners" due to the educational benefits associated with receiving critique). Such systems, however, assume that experts and learners have met previously offline at a design studio [10]. Consequently, these platforms fail at connecting individuals who have never physically attended a design studio, a space relatively foreign to most experts [11]. As a result, such platforms usually have a limited number of experts.

There are, however, many other tools that do facilitate interactions between experts and learners who have never met offline, e.g., online forums like Reddit. Here, learners can post photos/videos of their creative work; and then their peers or experts provide feedback to the creative artifacts [12].

But, experts on online forums generally get stuck in understanding what the creator tried to make. As a result, expert critique is rare [13], [14]. Another problem is that experts usually have concerns about providing feedback on forums [1] and thus prefer not to participate in the activity due to fears of saying something wrong and damaging their reputation [15]. Reputation is a longitudinal social evaluator about a person's actions and can be used as a measure of trustworthiness [16]. Performing in a manner that is unexpected can damage an individual's reputation as well as the organization that the individual represents [17]. Critiquing the work of novices can become a risky activity for experts because they might not have experience interacting with learners and could accidentally do or say things outside the norms, damaging their reputation [18].

Given the difficulties of coordinating experts online, recent research [19] has focused on obtaining critique from nonexperts, e.g., crowd-workers. However, individuals also use critique to learn about best practices, new topics, and even to network [1], activities which crowd workers can rarely complete. Expert critique is, therefore, still needed and should be something that researchers aim to facilitate, especially at scale, to benefit and empower more learners.

To enable learners at scale access to expert critique, we introduce, in Section 3, MATT (Micro Apprenticeship Through Tutorials), a chatbot that guides experts to critique creative work, especially of novices starting to create designs. Figure 1 presents an overview of MATT. In a conversational way, MATT guides experts to critique learners' work. MATT's guidance helps experts to rapidly understand what the learner tried to make. This empowers experts to be able to focus more on critiquing the work instead of interpreting it.

MATT breaks down its guidance into a set of micro-
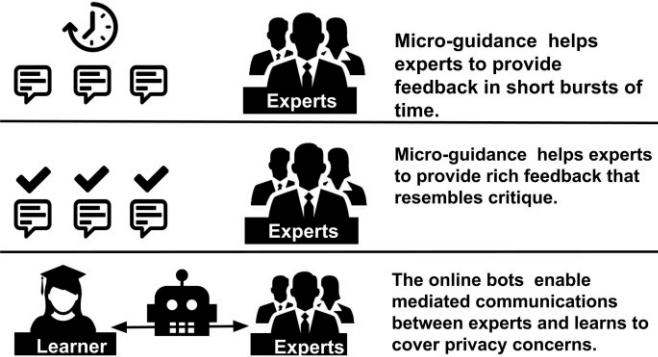
Figure 1. MATT integrates micro-guidance and mediated communication to enable experts to critique online.

tasks embedded in the conversation it has with experts. These micro-tasks facilitate the participation of more experts as they do not have to invest a large portion of their day in the activity. Experts, instead, are empowered to provide feedback throughout their spare time. Each micro-task asks experts to provide feedback on a particular aspect of the design, always tying it back to best practices. By guiding experts to focus on specific design elements, MATT ensures quality feedback resembling a critique.

It is important to note that most related platforms tend to assume that experts will work under prolonged and focused runs [20], [21]. One of our design challenges is thus to create small tasks that experts can do in short bursts of time throughout their day. Micro-tasking also becomes important because in our design we consider that experts are volunteering their time and knowledge (prior work has identified that experts are more likely to participate in such activities if they are intrinsically motivated [22]; therefore we limited providing them with monetary rewards and assumed they would volunteer their time). Given this setting, it becomes important not to burden experts. Through MATT, we broaden the design space of expert/learner systems to include the volunteer participation of specialists without requiring a large commitment. This enables providing specialized critique to a larger and broader number of learners.

Another of our design challenges is that experts can feel "insulted" from receiving guidance [23] (especially as they are allegedly the most knowledgeable in the area and they are volunteering in the activity). As a result, experts could be reluctant to follow directions on how they should critique. It is thus necessary to design guidance mechanisms that do not feel too imposing. We explore how such guidance can be designed via chatbots, which can provide structure without it feeling too commanding [24].

Our chatbot also acts as a proxy between experts and learners: learners first share with MATT their work; MATT then distributes the work to experts who are guided to critique the piece. Next, MATT presents to learners the feedback that experts' produced to help them improve. By creating mediated communication between experts and learners, MATT helps to address experts' privacy concerns. Notice that privacy is a natural concern since any information sent by learners or experts (who many times are public figures) is susceptible to misuse when shared with strangers. MATT addresses this by providing a mediated communication channel via a chatbot.

In section 4, we conducted a field deployment with MATT, where it coordinated a crowd of experts to critique the creative work of a large number of learners, which included posters, logos, and t-shirt designs. Through our study, we find that utilizing chatbots with micro-guidance empowers experts to provide feedback that approximates the gold standards of critique more closely. We finish by discussing in section 5 the design implications of our work.

II. RELATED WORK

The design of MATT is based on two main areas: (1) platforms for generating critique; and (2) platforms for eliciting specialized information from people online.

1. Platforms for Generating Critique. For many disciplines participating in the review of creative work is considered essential to develop skills in that area [1]. Many consider that being able to communicate with experts and use their feedback to improve is just as important as having particular knowledge and skills [25]. While the goal of critique varies across areas, its usefulness as an educational tool is consistent [26].

Related work has explored generating critique within online environments. However, given that even online, it is difficult to coordinate experts [27], [28], [29], most related work uses crowd workers to provide feedback to learners [30], [31]. While learners do appear to value such feedback as it has helped them to make substantial adjustments to their work [32], crowd workers have still not been able to match the range and depth of expert feedback [19]; even when having access to more direction and examples of expert type critique [33]. We should therefore not see feedback from crowd workers as a replacement to expert feedback, but rather a supplement. Focusing on the educational aspect that expert critique provides to learners, this paper explores the potential of orchestrating specialists to critique the creative work of learners at scale.

Similar to a design studio where experts volunteer their time, MATT assumes that experts are working pro-bono. This design facilitates providing access to expert knowledge to a broader range of learners, especially those from marginalized communities. This is not an eccentric idea, given that many experts have an interest in social good [34], especially if it is part of a revolutionary program [22].

However, experts usually lack the time necessary to identify how to best help others [35]; it can be especially timeconsuming to find volunteering opportunities that effectively utilize their specialization. In this sense, MATT facilitates the volunteering process of experts by directly dispatching to them micro-volunteering opportunities that utilize their expertise.

2. Eliciting Specialized Information from People Online. Recently we have seen the emergence of systems that ask people online to share specialized and specific information to benefit strangers [36]. Several human computation workflows have successfully driven strangers to share their

knowledge to help others learn [37]. These studies have found that online strangers can indeed provide quality information [38], even when asked by bots [39].

Researchers have also started to investigate the type of feedback that is possible to manually obtain from different online sites, especially crowd markets, social networks, and forums [40]. The problem, however, is that in these platforms, much time is spent interpreting what the learner produced [14].
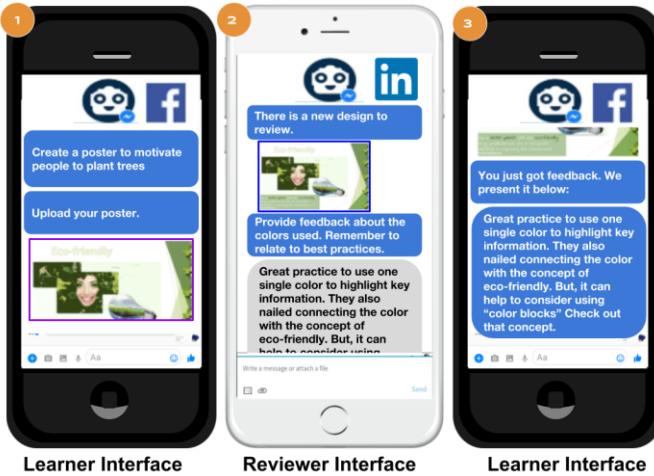
Figure 2. Overview of MATT's workflow: 1.- Learners submits to MATT their creative work. 2.- MATT finds an expert, sends the work to the expert, who is guided to review and provide micro-feedback approximating critique about the work. 3.- MATT then presents the micro-feedback from the expert to the learners who can use it to improve their work.

We motivate the design of MATT on some of the key findings of this previous research: it is possible to drive online strangers to provide useful information [38], [40], even when asked by bots [39]. We hypothesize that if we integrate guidance, we could orchestrate experts to effectively critique the creative work of learners they have never met before at scale.

III. MATT

MATT is a chatbot that: (1) collects creative work from learners; (2) presents the creative work to experts and guides them to critique the work; and (3) then gives the critique back to learners to help them improve. Figure 2. presents an overview of how MATT functions.

To accomplish these three steps, MATT consists of two main components: 1) the "Learner Helper" module that collects learners' creative work, distributes the work to experts and then shares experts' feedback to learners; 2) " Expert MicroGuidance" module that orchestrates experts to volunteer in short bursts of time quality micro-feedback that resembles online critique to help learners at scale.

A. Learner Helper Module

The goal of the Learner Helper component is threefold: 1) allow learners to submit their creative work easily; 2) find experts who can critique their work; 3) present back to the learner the feedback from experts. Figure 2 presents an overview of this workflow. Notice that the Learner-Helper acts as a proxy between learners and experts to address the privacy concerns of experts. Having mediated communication can also make it less awkward for an expert to reject reviewing a piece of work or say that they will review it once they are free. The learner would never know about the incident, but rather only MATT would be informed and would just search for another expert who can volunteer.

While there are many possible interfaces that could act as proxies between learners and experts, we consider a design that bootstraps on social media as it helps both learners and experts to easily share and receive critique from anywhere without needing to download or learn how to use new tools. Working on social media also facilitates finding people with particular specializations who can produce more relevant critiques [41], e.g., MATT can identify and recruit experts in "website design" based on the job title they present on their social media profile.

Our current design, therefore, considers that both learners and experts use social media, and we can utilize chatbots to act as proxies to connect these parties. We especially work within the Facebook messenger. Notice also that the design of MATT's Learner Helper module is based on intelligent conversational tutoring systems [42], which have shown to be effective for assisting learners.

B. Expert Micro-Guidance Module

MATT's Expert Micro-Guidance module focuses on orchestrating experts to produce, in short bursts of time, quality feedback that resembles critique. MATT, a chatbot on Facebook messenger, displays the learner's work to the expert and then asks the expert to complete small micro-tasks related to critiquing the learner's creative piece. The micro-tasks aim to guide experts to provide all the different types of feedback involved in a critique (especially identifying decisions the creators made in their design, and what are the best practices in each of the cases.) An example of these micro-tasks is to ask experts to provide feedback on the type of color used in the design and how it might or might not relate to best practices. Another similar micro-task is to ask an expert to "Provide feedback about the font type and size used, and how it relates (or not) to best practices."

The module has four features to enable this interaction.

1) Critique in Short Bursts of Time: Experts' time is limited, and experts also generally lack knowledge of how to effectively produce online critiques [35], [43]. MATT tackles this problem by guiding experts to provide critique to creative work in short bursts of time by leveraging task decomposition from crowdsourcing. Crowdsourcing has studied how long and complex work can be done via micro-tasks that are quick to finish. A long review and analysis of a piece of work can also be finished in small steps using the same process. MATT changes the nature of online critique by enabling experts to do it in small bursts of time. This design helps experts to take advantage of the time that might otherwise be wasted. To guide experts, MATT asks them a set of questions related to their perspectives and analysis of the creative work. MATT is designed to help experts also recall points they had covered in their feedback previously to aid learners' growth. Each of

MATT's question can be seen as a type of micro-task. These questions are based on prior work [1] that has defined guidelines or best practices to critique a piece of work. MATT empowers experts to critique in short bursts of time and at any point in time.

2) Critique Anywhere: MATT communicates via Facebook Messenger with experts. This design facilitates portability, and on-the-go experiences as experts can provide feedback wherever they use Facebook messenger, which can be on their desktop, their mobile device, or both. Experts can potentially provide feedback from anywhere, e.g., while waiting in line or on a shuttle.

We believe that these two functions enable more experts to participate in online critique, as they no longer have to invest consecutive hours at a physical desk reviewing work [44].

3) Privacy: MATT's mediated form of communication enables experts to remain anonymous to learners, which facilitates bringing experts' privacy. Our design builds upon privacy research that showcases that with anonymity, higher quality feedback is produced [45]. Our goal is that through MATT's mediated form of communication, experts will be more open and critical in their feedback, leading them to more deeply analyze creative pieces and consequently offering better learning opportunities to creators.

4) Conversational: MATT guides experts to produce critique within a conversational setting. We opted to use chatbots to guide experts because previous work had identified that they were viable sources for guiding strangers to provide specialized information [36] or to volunteer for a cause [46]. The conversational aspect of MATT might also help experts not to feel that MATT's guidance is too dictatorial. While previous work had identified that having chats incorporated into MOOCs did not necessarily increase student engagement [47], we adopt chat-based interfaces because they can help to create more "casual" environments that do not feel too "authoritarian" [48], which is important when working with experts who might feel they have the best knowledge and know-how of how to interact and provide feedback to novices.

IV. FIELD DEPLOYMENT

This paper hypothesizes that we can lead real-world experts to critique online by utilizing online mediated communication in the form of chatbots combined with micro-guidance. Our evaluation focuses on this claim: In the real-world, do chatbots micro-guiding experts enable a better approximation of the gold standard of studio design feedback? To respond to this question we conduct a real-world deployment of our tool and compare the feedback experts generated on MATT to two alternative interfaces: 1) chatbot lacking micro-guidance, i.e., a chatbot that simply asked experts to critique a piece of creative work without prompting experts on how to critique the piece; 2) online forum (we study online forums as they are a mediated communication channel that experts typically use to provide feedback [49]. Figure 3 presents an overview of these two interfaces and MATT.

Experts used either MATT or one of these two interfaces to provide feedback to learners. Learners were asked to create designs for real-world non-profits. We worked with non-profits because we were interested in having real-world usages of our tool, and this is one of the most common spaces where novice designers start to operate to build their portfolio [50]. Each learner produced one design, and each expert reviewed two designs from two different learners. Figure 4 presents examples of learners' work.

We recruited real word learners and experts using social media. To recruit learners, we posted on Facebook groups related to learning design, inviting people to our live deployment. Learners were offered the opportunity to potentially use new interfaces and obtain feedback from experts on their designs. To recruit experts, we used LinkedIn's search to find and invite individuals who stated they worked in design-related areas and identified themselves as experts. We recruited 153 learners and 76 experts. Each of our three interfaces was used by a total of 51 learners and 25 experts.
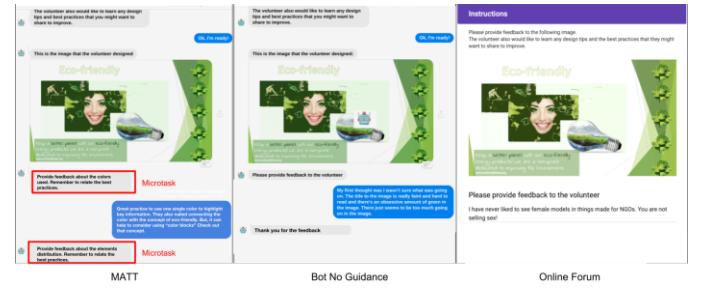

Figure 3. Feedback interfaces: 1) MATT, 2) Bot No-Guidance 3) Online Forum.

A. Categorizing Experts' Feedback

We were interested in understanding the type of feedback that experts in our real-world deployment generated. Our hypothesis was that experts using MATT would produce the most critiques. For this purpose, after experts provided feedback to learners' designs, we categorized their feedback according to the categories in the Feedback Typology of [1]. We recruited three college-educated Upworkers [51] and asked them to categorize experts' feedback into either: *"reactive,"*, *"direction,"*, or *"critique"*, i.e., the feedback categories in the typology that [1] identified. We define each category in detail below.

Reactive Category: emotional or visceral feedback that does not provide information on how to improve the work. Examples: "That's wonderful! Great work!" or "Horrible!"

Direction Category: In this form of feedback, the individuals providing the feedback try to bring the design more in line with their own expectations of what the solution should be. The feedback provides direction but no reasoning behind it. Examples: "I would have..." or "I wish..."

Critique Category: This feedback is considered to be the gold standard of design studios as it helps learners to improve their work and learn new techniques along the way. This type of feedback focuses on identifying decisions made in the creative work, relating that decision to a best practice, and

then describing how and why the decision made supports or does not support the best practices [3].

Two coders classified each of the feedback messages from experts into the category that represented the message the most (either critique, reactive, or direction). The two coders agreed on the classification of 90.1% of all the feedback produced by experts (Cohen's kappa =.89: Strong agreement). We then asked a third coder to act as a tiebreaker in cases of disagreement.

B. Results

Figure 5 presents the amount and type of feedback that experts generated in our real-world deployment with each interface. We observe that when using online forums and the chatbot without guidance, experts produced primarily reactive feedback. This result is in line with previous work that identified that experts online usually get stuck in interpreting the creative work by spending time trying to figure out what the goal of the designer was and consequently provide less critique [14], [52]. We observe from Figure 5 that MATT was the interface that leads experts to critique the most in the realworld.

Figure 4. Examples of the produced creative work.
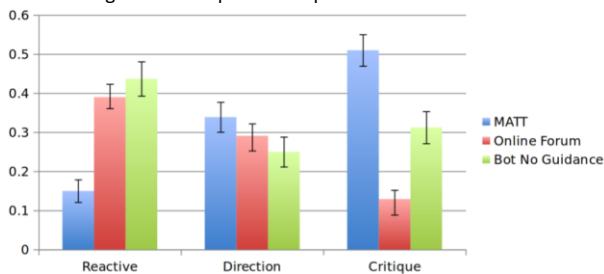
Figure 5. Overview of the type of feedback experts generated. Overall, experts using MATT generated the most critiques.

Given that we are primarily interested in whether MATT increases the amount of critique that a learner receives, we conducted a logistic regression predicting the likelihood that a piece of feedback would be classified as critique given its source (i.e., either from the MATT interface, Bot without Guidance interface, or the Online Forum). The logistic regression model showed that a piece of feedback was significantly more likely to be classified as critique when it came from MATT, compared to feedback from the online forum condition (B=1.12, z=4.96, $p < .01$) or the Bot No Guidance condition (B=0.83, z=3.31, $p < .01$). The overall model was a statistically significant fit to the data, Likelihood Ratio Test χ^2 (2) = 26.33, $p< .01$.

We were also interested in understanding experts' perceptions of the interfaces. It could be that although MATT lead experts to critique more, experts got annoyed with MATT "bossing" them around. We had a post-survey that asked experts about their experiences with MATT and the alternative interfaces. Experts first provided their impressions via fivelevel Likert questions and open responses.

Overall, experts enjoyed moderately the chatbot interfaces (mean=4.85 for MATT and for the chatbot without guidance). The forum interface was also enjoyed, but slightly less (mean=4.77). Experts considered all interfaces to be moderately easy to use (mean=4.8). Open-ended responses reinforced that experts felt that MATT helped them to produce meaningful feedback by directing the communication into what mattered: *"...Chatbots can direct communication efficiently which you don't really get with other technology [...] Suppose you want some information but are accidentally putting off the topic. The chatbot can steer you..."*

None of our experts expressed that MATT was too imposing. On the contrary, they felt that it presented a "sequential and clean" interface. Some experts expressed that the automated aspect of MATT made its guidance not feel too "bossy" because there was nothing personal about it. It was "just" a machine: *"Machines don't have feeling at all, so also nothing to feel on my side."* MATT's automation also helped experts to accept their guidance, as they felt that machines were made to help humans in their daily work. Thus, if a machine was trying to guide them, it must be for something beneficial: *They [machines] are just made to make human work easier [...] I felt the bot was steering towards meaningful communication. Just a good way to communicate..."*

Experts also felt that MATT addressed their privacy concerns (median = 5). Some seemed to especially like the format that MATT had for interacting with learners as they could help others while maintaining their privacy: *"I will get no benefits for not working anonymously. I don't want to be exposed to strangers....I just want to help. That's it...Chances of becoming more famous from doing this are too low to risk exposing my personnel details to strangers [...] I am completely satisfied with the bot [MATT], I am just providing feedback and not mentioning my personal information. So, providing feedback won't affect my privacy..."*

MATT's design also helped experts to not feel restricted in the feedback they shared. As one participant mentioned: *"If other people knew who I was by name, they might ask me later why I answered the way I did, or tell other co-workers what I said or did here. I'd have to then explain myself."* Similar to conversing with "strangers at a bar," this type of mediated communication likely facilitates being more open.

However, some experts noted that there were instances where they would like to possibly meet with learners and

further help them in their career growth (if the learner was willing). Experts' biggest requests for improving MATT involved adding different levels of privacy (e.g., being able to share where they worked with learners while keeping other information confidential). In the future, we will explore having more flexible privacy configurations.

Experts also mentioned that they would have liked to have a "better mental model" of the questions that MATT would ask them. In the future we will explore how we can better convene to experts the questions that MATT plans on covering. Perhaps here it is a matter of designing better conversations for MATT, so the questions feel more natural, and participants are not wondering about what will be asked next.

V. DISCUSSION

In our real-world deployment of MATT, hundreds of learners and experts collaborated to produce creative work and share critique. Here, we reflect on open challenges and opportunities for systems that orchestrate experts to help learners, in particular, to provide useful feedback.

An interesting implication from our study is that interactive and guided mediated communication (i.e., MATT) was the most helpful in leading experts to critique. This result might be arising because the interactive aspect of MATT might have lead experts to feel that they are working in a more conversational environment. Research has shown that "conversations" are an effective method to enhance learning [53]. It might be that this type of medium is also optimal for experts to express themselves and learn how to critique, and hence why MATT was the most optimal.

From our field deployment, we also observed that experts were empowered to provide quality feedback when working within a conversational type channel and when they focused their attention on specific features of the creative work (MATT's questions to experts were aimed at analyzing particular aspects of learners' work). We speculate that these conditions approximate the optimal conditions for critique that experts set for themselves in the design studio. Physical design studios facilitate *focus* (experts generally work on only one task at a time, inspecting one particular feature each time). Guided mediated communication through chatbots was also likely effective because it mitigated experts' time and task distribution concerns. Experts' were able to do the tasks in the time frame that they decided. Experts also had expectations of bots that seemed to facilitate the interactions. Some experts expressed how bots were there to help, and they were, therefore, willing to listen to the automated agent.

In our long-term vision of MATT, experts are given a platform where they can volunteer to share their knowledge in short bursts of time to support the learning process of any large crowd. We believe that it may be possible to lead experts to provide useful micro-feedback beyond our deployment of online critique. Opportunities include obtaining on-demand feedback for emergency response, accessibility, scientific discovery, citizen science, and a variety of other areas.

In MATT's design, the motivation of learners is clear: they gain support to improve their creative work. The incentives from experts are not as clear. Are experts motivated in providing feedback that impacts and helps the growth of other individuals or it simply to help in the creation of interesting creative work? Moving forward, we would like to explore the best way to motivate the continuous micro-participation of experts. This is especially important as having a large network of reliable experts can facilitate learning about any concept or topic. We believe there are important design opportunities in thinking about how to best match experts' intrinsic motivations with micro-volunteering opportunities and covering experts' privacy concerns.

A. Limitations

The insights from this work are limited by the methodology and population we studied. While our deployment allowed us to start to understand how experts engaged with systems like MATT, where a bot asks them to provide feedback to others, we cannot extrapolate to how experts would respond if this approach gained popularity and was widely used. In such a case, it might be relevant for these approaches to consider not pinging experts so frequently to avoid being ignored or labeled as spam. Additionally, while we recruited real-world experts and all creative work produced by learners resembled real-world creative projects, our results might not yet generalize to populations at large. Further analysis is needed to understand how systems that leverage experts and chatbots play out in helping learners to improve their work in different areas. Experiments that compare the type of feedback that experts generate for different areas would help quantify more broadly the effectiveness of using chatbots to guide expert critique. Future experiments that control for the social media platform or online ecosystem could be conducted to further understand what type of platform might facilitate accessing expert knowledge for on-demand feedback. Similar to [38], [39], the goal of this paper was to shed light on how micro-guidance embedded in chatbots facilitated expert critique. Future work could conduct longitudinal studies and engage in in-depth interviews with experts to understand their motivations and perspectives of these types of systems and approaches. Future work could also explore how learners react and benefit from the feedback that experts provide with MATT as well as their overall impressions of such technology. Some interesting questions for future work to explore with learners: what type of skills does MATT help learners to improve? Does MATT help learners to make better design decisions after feedback (in what way)? Are learners improving because they follow experts' advice (in which case they are not really learning a skill, but rather using MATT to get support with their performance)? Do learners' career prospects improve in some measurable way?

VI. CONCLUSION AND FUTURE WORK

In this paper, we introduced MATT, a chatbot that guides experts to critique the creative work of learners at scale. MATT embodies the vision that chatbots facilitate orchestrating experts to critique while addressing experts' privacy concerns and without creating an imposing environment on specialists. A field deployment provided evidence that MATT could guide experts to critique the creative work of hundreds of learners.

Future work lies in three main areas. First, further analysis is needed on best methods to combine chatbots and experts to

improve the engagement of learners long term, as well as workflows that enable crowds of learners and experts to best benefit from systems like MATT. Second, it will be important to devise mechanisms that can motivate experts to continuously micro-volunteer critique to learners. Third, in the long run, it will be important to design how experts and chatbots could help learners for more complex tasks. Experts are generally busy and consequently cannot do community work that is too time-consuming or demanding. This means that specialized social good work is generally not completed as the volunteers who do have the time lack the needed knowledge to complete the work [54]. We envision MATT's potential for combining crowds of experts, chatbots, and learners to complete complex volunteer work and create impactful change. In the future, we also plan to explore the impact of MATT on the type of creative work that learners produce and how they improve their work.

REFERENCES

[1] D. P. Dannels and K. N. Martin, "Critiquing critiques: A genre analysis of feedback across novice to expert design studios," Journal of Business and Technical Communication, vol. 22, no. 2, 2008, pp. 135–159.

[2] G. Fischer, K. Nakakoji, J. Ostwald, G. Stahl, and T. Sumner, "Embedding critics in design environments," The knowledge engineering review, vol. 8, no. 4, 1993, pp. 285–307.

[3] K. Luther, J.-L. Tolentino, W. Wu, A. Pavel, B. P. Bailey, M. Agrawala, B. Hartmann, and S. P. Dow, "Structuring, aggregating, and evaluating crowdsourced design critique," in Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. ACM, 2015, pp. 473–485.

[4] M. Fasli and B. Hassanpour, "Rotational critique system as a method of culture change in an architecture design studio: urban design studio as case study," Innovations in Education and Teaching International, vol. 54, no. 3, 2017, pp. 194–205.

[5] J. Corbin and A. Strauss, "Basics of qualitative research: Techniques and procedures for developing grounded theory," Sage Publications, 2008.

[6] K. H. Anthony, "Design juries on trial: The renaissance of the design studio," Van Nostrand Reinhold, 1991.

[7] S. Kiesler and J. N. Cummings, "What do we know about proximity and distance in work groups? a legacy of research," Distributed work, vol. 1, 2002, pp. 57–80.

[8] M. R. Louis, B. Z. Posner, and G. N. Powell, "The availability and helpfulness of socialization practices," Personnel Psychology, vol. 36, no. 4, 1983, pp. 857–866.

[9] J. Campbell, C. Aragon, K. Davis, S. Evans, A. Evans, and D. Randall, "Thousands of positive reviews: Distributed mentoring in online fan communities," in Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing. ACM, 2016, pp. 691–704.

[10] M. W. Easterday, D. Rees Lewis, C. Fitzpatrick, and E. M. Gerber, "Computer supported novice group critique," in Proceedings of the 2014 conference on Designing interactive systems. ACM, 2014, pp. 405–414.

[11] L. D. Setlock, S. R. Fussell, and C. Neuwirth, "Taking it out of context: collaborating within and across cultures in face-to-face settings and via instant messaging," in Proceedings of the 2004 ACM conference on Computer supported cooperative work. ACM, 2004, pp. 604–613.

[12] J. S. Brown, "The social life of learning: How can continuing education be reconfigured in the future?" Continuing Higher Education Review, vol. 66, 2002, pp. 50–69.

[13] K. Luther and A. Bruckman, "Leadership in online creative collaboration," in Proceedings of the 2008 ACM conference on Computer supported cooperative work. ACM, 2008, pp. 343–352.

[14] Y. Kou and C. Gray, "Supporting distributed critique through interpretation and sense-making in an online creative community," PACMHCI, 2017.

[15] A. Xu and B. Bailey, "What do you think?: a case study of benefit, expectation, and interaction in a large online critique community," in Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. ACM, 2012, pp. 295–304.

[16] M. Anwar and J. Greer, "Reputation management in privacy-enhanced e-learning," in Proceedings of the 3rd Annual Scientific Conference of the LORNET Research Network (I2LOR-06), Montreal, Canada, 2006.

[17] T. Casserley and D. Megginson, Learning from burnout: Developing sustainable leaders and avoiding career derailment. Routledge, 2009.

[18] M. Rhee and M. E. Valdez, "Contextual factors surrounding reputation damage with potential implications for reputation repair," Academy of Management Review, vol. 34, no. 1, 2009, pp. 146–168.

[19] A. Xu, H. Rao, S. P. Dow, and B. P. Bailey, "A classroom study of using crowd feedback in the iterative design process," in Proceedings of the 18th ACM conference on computer supported cooperative work & social computing. ACM, 2015, pp. 1637–1648.

[20] H.-M. Lee, J. Long, and M. R. Mehta, "Designing an e-mentoring application for facebook," in Proceedings of the 49th SIGMIS Annual Conference on Computer Personnel Research, ser. SIGMIS-CPR '11. New York, NY, USA: ACM, 2011, pp. 58–61. [Online]. Available: http://doi.acm.org/10.1145/1982143.1982175

[21] C. Toxtli, A. Monroy-Hernandez, and J. Cranshaw, "Understanding´ chatbot-mediated task management," in Proceedings of the 2018 CHI conference on human factors in computing systems, 2018, pp. 1–6.

[22] D. A. Joyner, "Scaling expert feedback: Two case studies," in Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale. ACM, 2017, pp. 71–80.

[23] P. A. Kohler-Evans, "Co-teaching: How to make this marriage work in front ofthe kids," Education, vol. 127, no. 2, 2006, pp. 260–264.

[24] C. Beaumont, "Beyond e-learning: an intelligent pedagogical agent to guide students in problem-based learning," Ph.D. dissertation, University of Liverpool, 2012.

[25] K. Reily, P. L. Finnerty, and L. Terveen, "Two peers are better than one: aggregating peer reviews for computing assignments is surprisingly accurate," in Proceedings of the ACM 2009 international conference on Supporting group work. ACM, 2009, pp. 115–124.

[26] T. Barrett, "A comparison of the goals of studio professors conducting critiques and art education goals for teaching criticism," Studies in art education, vol. 30, no. 1, 1988, pp. 22–27.

[27] E. Foong, D. Gergle, and E. M. Gerber, "Novice and expert sensemaking of crowdsourced design feedback," Proc. ACM Hum.Comput. Interact., vol. 1, no. CSCW, Dec. 2017, pp. 45:1–45:18. [Online]. Available: http://doi.acm.org/10.1145/3134680

[28] D. Retelny, S. Robaszkiewicz, A. To, W. S. Lasecki, J. Patel, N. Rahmati, T. Doshi, M. Valentine, and M. S. Bernstein, "Expert crowdsourcing with flash teams," in Proceedings of the 27th annual ACM symposium on User interface software and technology. ACM, 2014, pp. 75–85.

[29] R. Vaish, K. Wyngarden, J. Chen, B. Cheung, and M. S. Bernstein, "Twitch crowdsourcing: crowd contributions in short bursts of time," in Proceedings of the 32nd annual ACM conference on Human factors in computing systems. ACM, 2014, pp. 3645–3654.

[30] A. Xu, S.-W. Huang, and B. Bailey, "Voyant: Generating structured feedback on visual designs using a crowd of non-experts," in Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, ser. CSCW '14. New York, NY, USA: ACM, 2014, pp. 1433–1444. [Online]. Available: http://doi.acm.org/10.1145/2531602.2531604

[31] A. Xu and B. Bailey, "What do you think?: A case study of benefit, expectation, and interaction in a large online critique community," in Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, ser. CSCW '12. New York, NY, USA: ACM, 2012, pp.

295–304. [Online]. Available: http://doi.acm.org/10.1145/2145204.2145252

[32] K. Luther, J.-L. Tolentino, W. Wu, A. Pavel, B. P. Bailey, M. Agrawala, B. Hartmann, and S. P. Dow, "Structuring, aggregating, and evaluating crowdsourced design critique," in Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, ser. CSCW '15. New
York, NY, USA: ACM, 2015, pp. 473–485. [Online]. Available: http://doi.acm.org/10.1145/2675133.2675283

[33] M. D. Greenberg, M. W. Easterday, and E. M. Gerber, "Critiki: A scaffolded approach to gathering design feedback from paid crowdworkers," in Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition. ACM, 2015, pp. 235–244.

[34] H. Bussell and D. Forbes, "Understanding the volunteer market: The what, where, who and why of volunteering," International Journal of Nonprofit and Voluntary Sector Marketing, vol. 7, no. 3, 2002, pp. 244–257.

[35] E. Brady, M. R. Morris, and J. P. Bigham, "Social microvolunteering: Donating access to your friends for charitable microwork," in Second AAAI Conference on Human Computation and Crowdsourcing, 2014.

[36] J. Nichols and J.-H. Kang, "Asking questions of targeted strangers on social networks," in Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. ACM, 2012, pp. 999–1002.

[37] S. Weir, J. Kim, K. Z. Gajos, and R. C. Miller, "Learnersourcing subgoal labels for how-to videos," in Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, ser. CSCW '15. New York, NY, USA: ACM, 2015, pp. 405–416. [Online]. Available: http://doi.acm.org/10.1145/2675133.2675219

[38] J. Nichols, M. Zhou, H. Yang, J.-H. Kang, and X. H. Sun, "Analyzing the quality of information solicited from targeted strangers on social media," in Proceedings of the 2013 conference on Computer supported cooperative work. ACM, 2013, pp. 967–976.

[39] S. Savage, A. Monroy-Hernandez, and T. Hollerer, "Botivist: Calling volunteers to action using online bots," arXiv preprint arXiv:1509.06026, 2015.

[40] Y.-C. G. Yen, S. P. Dow, E. Gerber, and B. P. Bailey, "Social network, web forum, or task market?: Comparing different crowd genres for design feedback exchange," in Proceedings of the 2016 ACM Conference on Designing Interactive Systems. ACM, 2016, pp. 773–784.

[41] C. Grevet and E. Gilbert, "Piggyback prototyping: Using existing, largescale social computing systems to prototype new ones," in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM, 2015, pp. 4047–4056.

[42] V. Rus, D. Stefanescu, N. Niraula, and A. C. Graesser, "Deeptutor: Towards macro-and micro-adaptive conversational intelligent tutoring at scale," in Proceedings of the first ACM conference on Learning@ scale conference. ACM, 2014, pp. 209–210.

[43] J. Sch, M. Raab, J. Altmann, E. Kapsammer, A. Kusel, W. Retschitzegger, W. Schwinger et al., "A survey on volunteer management systems,"
in 2016 49th Hawaii International Conference on System Sciences (HICSS). IEEE, 2016, pp. 767–776.

[44] J. R. Kogan, "How to evaluate and give feedback," in The Academic Medicine Handbook. Springer, 2013, pp. 91–101.

[45] R. Lu and L. Bol, "A comparison of anonymous versus identifiable epeer review on college student writing performance and the extent of critical feedback," Journal of Interactive Online Learning, vol. 6, no. 2, 2007, pp. 100–115.

[46] S. Savage, A. Monroy-Hernandez, and T. Hollerer, "Botivist: Calling¨ volunteers to action using online bots," in Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing. ACM, 2016, pp. 813–822.

[47] D. Coetzee, A. Fox, M. A. Hearst, and B. Hartmann, "Chatrooms in moocs: all talk and no action," in Proceedings of the first ACM conference on Learning@ scale conference. ACM, 2014, pp. 127–136.

[48] G. Ball and J. Breese, "Emotion and personality in a conversational agent," Embodied conversational agents, 2000, pp. 189–219.

[49] J. Luca and C. McLoughlin, "Using online forums to support a community of learning," in EdMedia: World Conference on Educational Media and Technology. Association for the Advancement of Computing in Education (AACE), 2004, pp. 1468–1474.

[50] A. Acus, "Volunteering in non-governmental organizations as social policy expression," Tiltai, no. 3, 2018, pp. 149–163.

[51] "In-demand talent on demand. upwork is how." https://www.upwork.com/, (Accessed on 02/10/2020).

[52] R. Vahidov and R. Elrod, "Incorporating critique and argumentation in dss," Decision Support Systems, vol. 26, no. 3, 1999, pp. 249–258.

[53] R. Bellamy and K. Woolsey, "Learning conversations," SIGCHI Bull., vol. 30, no. 2, Apr. 1998, pp. 108–112. [Online]. Available: http://doi.acm.org/10.1145/279044.279170

[54] L. S. Hartenian, "Nonprofit agency dependence on direct service and indirect support volunteers: An empirical investigation," Nonprofit Management and Leadership, vol. 17, no. 3, 2007, pp. 319–334.