

CrowdFolio: Understanding How Holistic and Decomposed Workflows Influence Feedback on Online Portfolios

EUREKA FOONG*, Tokyo College, University of Tokyo, Japan

JOY O. KIM, Adobe Research, USA

MIRA DONTCHEVA, Adobe Research, USA

ELIZABETH M. GERBER, Northwestern University, USA

Freelancers increasingly earn their livelihood through online marketplaces. To attract new clients, freelancers continuously curate their online portfolios to convey their unique skills and style. However, many lack access to rapid, regular, and inexpensive feedback needed to improve their portfolios. Existing crowd feedback systems, which collect feedback on *individual* creative projects (i.e., decomposed approach), could fill this need, but it is unclear how they might support feedback on *multiple* projects (i.e., holistic approach). In a between-subjects study with 30 freelancers, we compared decomposed and holistic feedback collection approaches using CrowdFolio, a crowd feedback system for portfolios. The holistic approach helped freelancers discover new ways to describe their work, while the decomposed approach provided detailed insight about the visual attractiveness of projects. This study contributes evidence that portfolio feedback systems, regardless of collection approach, can positively support professional development by impacting how freelancers portray themselves online and reflect on their identity.

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

Additional Key Words and Phrases: crowdsourcing, feedback, freelancing, online professional identity, careers, professional development

ACM Reference Format:

Eureka Foong, Joy O. Kim, Mira Dontcheva, and Elizabeth M. Gerber. 2021. CrowdFolio: Understanding How Holistic and Decomposed Workflows Influence Feedback on Online Portfolios. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 22 (April 2021), 31 pages. <https://doi.org/10.1145/3449096>

1 INTRODUCTION

Millions of freelancers earn their livelihood through online platforms that connect them to paying clients [48]. To attract clients, freelancers must convey their unique skills and style through online portfolios [29, 36]. For example, a furniture designer may attract clients by posting images of the living room furniture she built to show off her wood-working skills and rustic style. As freelancers create new work, they must curate relevant and appealing projects on online portfolio sites to attract clients and job opportunities (e.g., Behance [4], Dribbble [5, 59]).

*Previously affiliated with Northwestern University

Authors' addresses: Eureka Foong, foong.eureka@tc.u-tokyo.ac.jp, Tokyo College, University of Tokyo, Tokyo, Japan; Joy O. Kim, joykim@adobe.com, Adobe Research, USA; Mira Dontcheva, mirad@adobe.com, Adobe Research, USA; Elizabeth M. Gerber, egerber@northwestern.edu, Northwestern University, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Association for Computing Machinery.

2573-0142/2021/4-ART22 \$15.00

<https://doi.org/10.1145/3449096>

To curate relevant projects for their portfolios, freelancers need rapid, regular, and inexpensive feedback from new, diverse, and readily available feedback providers [10, 67], but many lack this opportunity when working outside of support networks, such as schools or companies, that can provide feedback [43, 44]. Freelancers could ask peer freelancers for feedback, but peers may not always be readily available to offer rapid, regular feedback [45]. Peers may also not always reflect the perspectives of their diverse and sometimes non-expert clients. An alternative approach is to hire dedicated, professional reviewers, who are new to the freelancer’s work and could offer diverse perspectives, but this approach may be expensive. Instead, freelancers are left with vague cues, such as the number of “likes” they receive on social platforms, to determine whether their portfolio represents them in a relevant and appealing way.

Not only do freelancers need rapid, regular, and inexpensive feedback, it is unclear how freelancers should collect feedback on their portfolios. On one hand, taking a holistic approach (i.e., asking reviewers to evaluate multiple projects as a group) could lead to more critical feedback due to the ability to compare works [28, 70]. On the other hand, taking a decomposed approach (i.e., asking reviewers to evaluate individual projects independently) may better ensure projects receive detailed and actionable feedback [14, 58, 65]. Decomposed and holistic approaches to collecting feedback on portfolios could differently influence *what* feedback freelancers receive and *how* they use that feedback [33].

Crowd feedback systems could provide freelancers with rapid, regular, and inexpensive portfolio feedback from new, diverse, and readily available feedback providers (e.g., [33, 57]), but existing systems only take a *decomposed* approach to collecting feedback. In current systems, crowds are guided to produce detailed feedback regarding a single project [19, 37, 51, 57, 72]. By using decomposed tasks on individual projects, these systems solicit high quality feedback to prompt project improvements. Some researchers have begun designing systems that solicit feedback on two projects at a time, recognizing that a holistic approach can lead to more specific and actionable feedback (e.g., [16, 49]), but have not yet studied ways to provide feedback on *multiple* projects in a portfolio. In this study, we evaluate holistic and decomposed approaches to collecting crowd feedback in the context of an *entire* portfolio of work (Table 1).

		Focus	
		Single Project	Portfolio
Feedback Approach	Holistic	<i>Paragon</i> (Kang et al., 2018), <i>Juxtapoeer</i> (Cambre et al., 2018), <i>Mechanical Novel</i> (Kim et al., 2017), <i>Mobi</i> (Zhang et al., 2012)	
	Decomposed	<i>CrowdCrit</i> (Luther et al., 2015), <i>Voyant</i> (Xu et al., 2015), <i>Critiki</i> (Greenberg et al., 2017), <i>CrowdForge</i> (Kittur et al., 2011)	

Table 1. In prior work, crowdsourcing researchers have explored using holistic and decomposed approaches to collect feedback for improving single projects. In the current study, we explore the open question of how holistic and decomposed approaches to collecting feedback and data from online crowds can impact the improvement of multiple projects in a portfolio.

To address this gap in the crowdsourcing research, we applied a design research process [30, 78] to design and evaluate CrowdFolio, a crowd feedback system which supports holistic and decomposed approaches to solicit feedback on graphic design portfolios. In this study, we asked: 1) can online

crowds provide helpful portfolio feedback that influences how freelancers reflect on and improve their portfolio, and 2) how do holistic and decomposed approaches to collecting crowd feedback influence a) the characteristics and perceptions of the feedback, and b) how freelancers revise their descriptions of their portfolio and plans for improvement.

We addressed these research questions by conducting a between-subjects experiment with 30 freelance graphic designers in the United States (US). Regardless of how it was collected, feedback from CrowdFolio helped freelancers recognize opportunities to improve their portfolio by highlighting gaps between their and the crowd's perceptions. The system also filled a need for portfolio feedback from a non-expert crowd that provided perspectives unique from freelancers' peers. Study transcripts showed that the holistic approach allowed the crowd to highlight stylistic similarities between projects (e.g., "community-focused", "youthful" projects) that the freelancers had not considered. At the same time, a qualitative analysis of these transcripts suggested that the decomposed approach provided helpful details about the visual attractiveness of individual projects and led to significantly longer comments.

We contribute to crowdsourcing research a novel comparison of holistic and decomposed approaches to collecting crowd feedback in a professional setting [33]. For the computer-supported cooperative work (CSCW) community at large, we show that crowd feedback, regardless of the feedback collection approach, can help users better understand the professional image they portray online. Thus, we contribute to a growing body of research on the benefits of crowdsourcing to support complex, creative, and personal goals (e.g., [16, 50, 77]).

2 BACKGROUND

2.1 Importance of Online Portfolios to Freelance Careers

To seek work online, freelancers must constantly attract clients (e.g., [66]) through their online portfolios, online profiles (e.g., Twitter [29, 36]), and reviews [9, 20, 22]. Unlike client reviews over which the freelancer limited control, most freelancers take the opportunity to regularly update their online portfolios to attract new clients. A portfolio is a regularly curated collection of prior work projects often used in creative fields to quickly communicate one's skills and style [56, 59]. Potential clients come from diverse backgrounds and have a variety of target audiences for their projects. For example, a client might include a writer with no design experience needing a children's book cover, or a hiring manager with 10 years of design experience needing a logo targeted at young women. Therefore, freelancers must be adept at regularly updating and describing their skills and style through their portfolio to attract work opportunities.

2.2 Challenges in Seeking Feedback on Online Portfolios

Despite the importance of online portfolios to freelance work, freelancers currently face challenges obtaining rapid, regular, and inexpensive feedback from new, diverse, and readily-available feedback providers to improve their portfolios [43, 44]. Receiving feedback is a vital part of improving one's creative practice [23]. Freelancers should ideally seek feedback on their portfolios as they complete new work projects [56], which can span every few weeks to every few months [44]. However, freelancers often work outside of support networks, such as schools and companies [43, 44], and may not have regular opportunities to seek feedback from peers. Peers may also not always reflect the perspectives of their diverse and sometimes non-expert clients. While online communities, such as Reddit [7], could provide feedback on freelancers' work [54, 59], it is not clear how to best present portfolios on these platforms, which typically accommodate feedback on single pieces of work. An alternative approach is to hire dedicated, professional reviewers who are new to the freelancer's work and could offer diverse perspectives, but this may be expensive. In addition, interpreting

social media engagement metrics (e.g., number of likes) may lead freelancers to make inaccurate conclusions about the appeal of their work [24, 31, 42]. Hence, there is a need for freelancers to obtain regular, rapid, and inexpensive feedback on their online portfolio from new, diverse, and readily-available feedback providers, regardless of the size and availability of their social network.

2.3 Crowd Feedback Systems as a Way to Collect Feedback on Creative Projects

Crowd feedback systems can provide regular, rapid, and inexpensive feedback from new, diverse and readily-available paid online crowdworkers on individual creative projects [33, 37, 57, 72, 73]. These online crowds complement other sources of feedback, such as peers, who may be more homogenous, familiar, and not always available to provide feedback [57, 73]. These systems accomplish high-quality critique by providing crowdworkers with heavily structured tasks [19, 51]. Collecting crowd feedback on individual projects, without showing them in the context of other projects, enables crowdworkers to write detailed, specific feedback [27, 51, 57]. For example, Voyant [72, 73] asks online crowds to answer structured questions about elements of visual design, such as which areas of a poster are initially most visually noticeable. CrowdCrit [57, 76] asks crowdworkers to critique an individual visual design project using a rubric of design principles, such as layout and hierarchy.

In contrast, to provide portfolio-based crowd feedback, a reviewer may need additional guidance and context about a person's prior work, but it is less clear how we might design effective crowd tasks to support this. Crowdsourcing researchers have begun exploring the potential of collecting feedback on individual projects in the context of other projects. For instance, Paragon [49] asks online crowds to refer to similar examples of projects by other people when providing their critique, and Juxtapeer [16] encourages reviewers to compare classmates' course projects side-by-side as they write their reviews. In comparison, our research focuses on crowd feedback for a *portfolio* of work instead of individual projects, as freelancers use portfolios to present themselves to and attract new clients.

2.4 Lack of Understanding of Holistic and Decomposed Approaches to Collecting Portfolio Feedback

How one collects feedback could influence *what* feedback one receives and *how* one uses the feedback [28, 33, 70]. Taking a decomposed approach (i.e., asking crowds to evaluate projects independently and aggregating the results) may lead to different outcomes from taking a holistic approach (i.e., asking crowds to evaluate multiple projects at the same time). For example, collecting feedback on more than one version of a prototype can lead to more critical feedback [70] and encourages creators to compare feedback to make larger improvements [28].

Holistic and decomposed approaches to collecting feedback may also influence visual perception and judgment. Studies suggest that people are more effective at processing visual scenes holistically (e.g., [60]) and perceive objects differently when placed in the context of other objects (e.g., [38, 39, 75]). However, taking a decomposed approach to collecting feedback could provide more detailed and thoughtful feedback by guiding people to think about specific problems one at a time. In studies of writing, reviewers who were given explicit criteria to critique (e.g., flow, coherence) gave more reliable ratings [58, 65]. In our case, taking a decomposed approach to collecting feedback could prompt more detailed feedback to help freelancers understand how their portfolio is perceived.

Crowdsourcing researchers have also begun to explore the benefits of holistic and decomposed approaches of collecting feedback to coordinate creative tasks. For example, Mechanical Novel [50] takes a holistic approach by providing online crowds with context about the group's progress as they collaboratively critique and compose a fiction story, enabling crowds to react flexibly to changes in the story and make holistic improvements. Similarly, Mobi [77] helps crowds collaboratively create travel itineraries by providing an overview of the existing itinerary and goals to ensure

that end itineraries are coherent. In contrast, systems such as CrowdForge [52] use decomposed tasks (i.e., writing and critiquing parts of an article) and aggregating individual contributions to minimize crowdworkers' time and effort to complete the tasks. Although researchers have explored decomposed and holistic approaches to crowdsourcing tasks, we have yet to directly compare these approaches within a single system.

3 FORMATIVE DESIGN RESEARCH PROCESS

Our research began with the following question: How do creatives develop their portfolios? To answer this question, we observed three portfolio design class sessions at our university and conducted a literature review to understand best practices for portfolio development. Through these observations and literature review, we recognized a need for decomposed and holistic approaches to collecting feedback on how people perceive a freelancer's skills and unique style. Through a design research process [30, 78], we then developed and evaluated prototypes of a portfolio feedback system for graphic design freelancers (described in detail in Appendix A at the end of this paper). We focused on graphic design portfolios in this study because of the ease of sharing graphic design projects in a digital online format and the popularity of graphic design projects in creative online communities [47]. In addition to supporting designers' career goals and collecting feedback rapidly, regularly, and inexpensively, we identified three aspects of a portfolio that would be most appropriate for crowd feedback:

- Visual Attractiveness (which projects capture the most attention from potential clients?),
- Styles (how would other people describe the portfolio?),
- and Target Audience (what type of audience are the creative projects designed for and what primary design skills does the portfolio convey?).

While these early findings were promising, we still needed to understand the practical challenges to collecting crowd feedback on online portfolios and the effects of decomposed (i.e., reviewing one project at a time) and holistic approaches (i.e., reviewing multiple projects) to collecting feedback.

4 CROWDFOLIO SYSTEM DESCRIPTION

While our formative design research uncovered the types of feedback that graphic design freelancers would find helpful to improving their portfolios, we still lacked an understanding of the practical challenges of crowdsourcing this type of feedback online, especially since existing systems are designed to ask crowdworkers to assess individual projects rather than multiple projects at once. Therefore, we designed CrowdFolio, a crowd feedback system built on top of the Amazon Mechanical Turk (MTurk) crowdsourcing platform [1] to study the effects of decomposed and holistic approaches to collecting feedback. Building on top of the MTurk platform is a standard practice of existing crowd feedback systems because of its ease and low cost [57, 72].

The CrowdFolio system consists of two parts:

- (1) **Feedback collection interface.** Given a link to a portfolio hosted on Behance [4], this interface deploys a series of crowdsourcing tasks that collect feedback on the perceived styles or aesthetic in the portfolio, the perceived visual attractiveness of project thumbnails in the portfolio, and the perceived target audience of the portfolio.
- (2) **Feedback presentation interface.** This interface aggregates and displays the collected crowd feedback to the creator of the portfolio.

Crowdfolio is an interactive NodeJS/HTML/CSS web application that uses the *aws-sdk* package to launch and manage crowdworker tasks. To improve the user experience of the feedback presentation interface, we conducted usability tests with three freelance graphic designers (two female, one

Dimension	Holistic Task
Visual Attractiveness	Rank up to 20 projects in a portfolio from a single designer (displayed in random order) in order of most to least visually attractive by clicking on project thumbnails in order of attractiveness.
Styles	Using a drag-and-drop interface, create at least three groups of related project thumbnails in a portfolio from a single designer and title each group of projects with a style phrase. Not all projects in the designer's portfolio must be in a group, but all groups must include at least three projects. In a second series of crowd tasks, we collect relevance ratings of the style phrase with respect to the project group; other crowdworkers are shown a random (style phrase, project group) pair and indicate whether they agree that the style phrase describe the project thumbnails in the group using a seven-point Likert scale.
Target Audience	View a portfolio from a single designer and answer survey questions regarding the designer's perceived target audience (e.g., "What kind of audience do you think would most like the projects in the designer's portfolio?").

Table 2. A description of the crowdsourcing tasks we ran on MTurk to collect feedback on participants' portfolios in the holistic feedback condition.

Dimension	Decomposed Task
Visual Attractiveness	Independently rate 20 randomly selected projects, from different designers, on a five-point Likert scale.
Styles	Given a randomly selected project thumbnail, generate at least three style phrases that describe the project. In a second series of crowd tasks, we collect relevance ratings of the style phrases with respect to the project thumbnail; other crowdworkers are shown a random (three style phrases, project thumbnail) pair and indicate whether they agree the style phrases describe the project thumbnail using a seven-point Likert scale.
Target Audience	Given a randomly selected project thumbnail, answer survey questions about project's perceived target audience.

Table 3. A description of the crowdsourcing tasks we ran on MTurk to collect feedback on participants' portfolios in the decomposed feedback condition.

Dimension	Feedback Presentation
Visual Attractiveness	A ranked list of project thumbnails in descending order of attractiveness, either using mean rankings (holistic) or mean ratings (decomposed; Figure 1).
Styles	Style phrases and projects (decomposed) or project groups (holistic) displayed in descending order of their mean relevance rating (Figure 2). Comments explaining these phrases are displayed underneath each style group.
Target Audience	A text summary of survey answers, comments from crowdworkers explaining their survey answers, and bar charts visualizing the distribution of survey answers (holistic and decomposed; Figure 3).

Table 4. A description of how feedback from the holistic and decomposed study conditions was presented to freelancer participants.

male; ages 22-27) with portfolios on Behance [4]. Below, we describe the crowdsourcing tasks to collect feedback on portfolios and how the system displays this feedback to users.

4.1 Collecting Crowd Feedback

Based on our formative design research and prior crowdsourcing research (e.g., [51]), the CrowdFolio feedback collection interface consists of crowdworker tasks to collect feedback on the three dimensions determined from the formative study. Each task had two variations. One variation collected feedback using a holistic approach (i.e., assess multiple projects at once) while the other used a decomposed approach (i.e., assess individual projects; Tables 2 and 3). To better simulate the experience of a potential client with limited time to form an impression, we only showed project thumbnails to crowdworkers, rather than the full projects (see the Supplemental Materials). Below we discuss how these tasks were informed by prior work on crowd feedback [57, 73].

4.1.1 Visual Attractiveness. One of our goals was to help freelancers understand which projects in their portfolio capture the most attention from potential clients. In prior work [72, 73], systems ask crowdworkers to click on elements of a single project that first capture their attention. Similarly, we asked crowdworkers in the holistic condition to click on project thumbnails in the order of which thumbnails first captured their attention. We then created a comparable task for the decomposed condition by asking crowdworkers to rate the visual attractiveness of different project thumbnails.

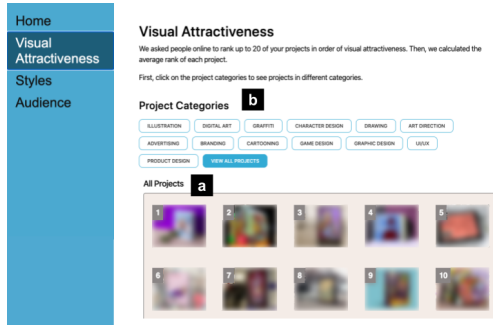


Fig. 1. CrowdFolio Feedback Presentation Interface: Visual Attractiveness. Feedback is displayed as (a) a ranked list of the projects in a freelancer’s portfolio. Projects were shown in descending rank (holistic) or rating (decomposed) order of visual attractiveness. Participants were able to (b) filter projects they wanted to see using the Project Categories options. Project images have been blurred to maintain the confidentiality of the participant.

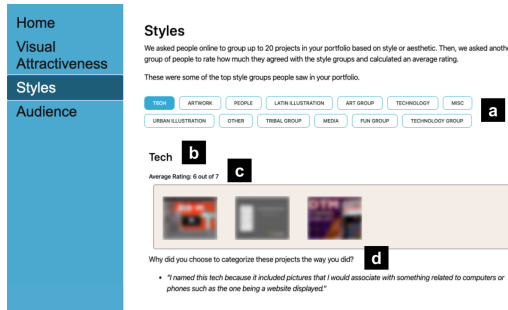


Fig. 2. CrowdFolio Feedback Presentation Interface: Styles. This view displays (a) the style phrases and (b) the projects that match each style phrase in (c) descending order of their mean relevance in both the holistic and decomposed feedback conditions. The relevance of a style phrase is determined by asking crowdworkers whether they agree or disagree that a style phrase (suggested by another crowdworker) appropriately describes a project or group of projects (1 - strongly disagree; 7 - strongly agree). Participants were also shown (d) comments from crowdworkers explaining their choices for style phrases and project groups. Project images have been blurred to maintain the confidentiality of the participant.

4.1.2 *Styles*. Our second goal was to collect feedback so that freelancers could understand how others stylistically perceive their work. In prior crowd feedback systems, [73], crowdworkers are asked to generate phrases to describe their first impressions of project. Beyond the design domain, crowd clustering systems (e.g., [13, 17, 18]) ask crowds to create and label groups of text to generate valuable information (e.g., types of work valued on Wikipedia). Therefore in CrowdFolio, we asked crowdworkers in the holistic condition to group projects in a portfolio and generate style phrases for these style groups. In the decomposed study condition, we asked crowdworkers to generate style phrases for individual projects. Following prior work in crowdsourcing [13, 18, 52], we took steps to improve the quality of these phrases by asking additional crowdworkers to rate how well each phrase described the given project(s).

4.1.3 *Target Audience*. Finally, we collected feedback to help freelancers understand the type of audience they were attracting. While prior crowd feedback systems have not addressed this aspect



Fig. 3. CrowdFolio Feedback Presentation Interface: Target Audience. This view displays (a) a paragraph summarizing the top answers to survey questions related to perceived target audience and (b) comments explaining crowdworker responses. It also displays (c) a breakdown of the survey answers to each survey question using multiple bar charts.

of feedback, they have asked crowdworkers to describe their impressions of a project on several dimensions (e.g., layout, visual hierarchy) [57, 72]. Drawing inspiration from these approaches, CrowdFolio showed crowdworkers either a portfolio (holistic) or individual project (decomposed) and asked them to answer survey questions about the freelancer’s target audience (e.g. “What kind of businesses do you think this designer would work with the most?”) and primary design skill.

4.2 Presenting Feedback to Freelancers

The CrowdFolio feedback presentation interface consists of three web pages, each corresponding to the three crowd feedback tasks (Table 4). On the Visual Attractiveness page, we showed a ranked list of projects in a freelancers’ portfolio in descending order of attractiveness using either mean rankings (holistic) or ratings (decomposed) (Figure 1). On the Styles page, we displayed the style phrases and project groups in descending order of their mean rating, as well as comments to justify these style phrases (Figure 2). In our formative research, we considered computationally aggregating the feedback in the decomposed condition; however, when we used Empath [32] to aggregate style phrases, up to 37% of phrases were unable to be grouped. We describe opportunities to improve on automated grouping of style phrases in the Discussion. On the Target Audience page, we displayed a paragraph and bar charts summarizing the answers to the survey questions (Figure 3).

5 RESEARCH QUESTIONS

To understand whether crowd feedback collected using decomposed and holistic approaches could provide a feasible way for freelancers to improve their portfolios, we asked the following questions:

- RQ 1: Can online crowds provide helpful portfolio feedback that influences how freelancers reflect on and improve their portfolio?
- RQ 2: How do holistic and decomposed approaches to collecting crowd feedback on online portfolios influence a) the characteristics and perceptions of the feedback and b) the way freelancers describe their work to others and plan to improve their portfolio?

6 METHOD

6.1 Participants

We recruited 30 freelance graphic design professionals (56% female, 41% male, 3% non-binary) in the US and 1,386 unique crowdworkers from MTurk [1] in North America. To ensure high quality responses from MTurk workers, we recruited workers with at least 110 approved tasks on the platform and a task approval rate of 90%. We recruited freelancers who had a portfolio on Behance [4] for at least one year with at least seven published projects to ensure that participants had similar portfolios and were at least somewhat active in using their online portfolio. We did not control for portfolio size because this can vary between designers and would have significantly reduced our potential sample size. We also collected feedback on only up to the most recent 20 projects in each portfolio (described later in this section), based on best practices in portfolio design [15, 47, 47]. We recruited freelancers via Behance [4], design forums on websites such as Reddit [7] and Facebook [6], and online freelance labor marketplaces, such as Upwork [8]. All freelancers received \$50 for the one-hour study and crowdworkers received between \$0.40 and \$0.90 per crowd feedback task (depending on the length of the task they completed) to achieve an hourly wage of approximately \$10.80 [41].

6.2 Procedure

To evaluate CrowdFolio, we conducted a between-subjects experiment with two conditions. First, after freelancers indicated their interest in participating and scheduled a study session, we began collecting feedback on their Behance portfolios from MTurk workers using the CrowdFolio feedback collection interface. Each freelancer and MTurk worker in our study was randomly assigned to only one condition (i.e., holistic or decomposed feedback). We launched different feedback tasks to MTurk to collect feedback on 1) the relative visual attractiveness of different projects, 2) the perceived styles, and 3) the perceived target audience of the portfolio. When crowdworkers chose to preview one of our tasks on MTurk, the system randomly assigned their MTurk ID to one of the two study conditions and displayed the appropriate version of the task; this also ensured that they were assigned to the same condition for any future tasks they previewed.

At a high level, crowdworkers in the holistic condition made assessments based on up to 20 projects in a freelancer's portfolio, while crowdworkers in the decomposed condition made assessments based on single projects from various freelancers' portfolios (Tables 2 and 3). We randomized the presentation order of all feedback items shown to crowdworkers. Though freelancer participants were only shown the feedback relevant to the study condition to which they were assigned, we collected both holistic and decomposed feedback for all freelancer participants in our study to make it easier to randomize crowdworkers into the study conditions. To help freelancers understand the background of the crowdworkers who provided feedback, all crowdworkers were asked to answer demographic questions about their age, gender, education, region in the US, and income level.

Once we collected feedback on a freelancer's portfolio, which took between two and 24 hours, we conducted a one-hour remote study session with each freelancer. We asked freelancers to complete a pre-task survey (see Supplemental Materials). To understand how frequently participants were already seeking feedback and whether they already had a clear sense of their professional identity prior to our study, we assessed participants' frequency of seeking feedback on one's portfolio (adapted from a self-monitoring scale by [55]) and the clarity of their professional identity [26]. To understand how the CrowdFolio might influence the way freelancers describe their work, we asked freelancers to type 1) one to two sentences to describe themselves and their work for use on their personal website, and 2) three phrases to tag their work online. To prepare participants, we

asked them to practice conducting a web search task while thinking aloud and view a webpage summarizing the demographic details of crowdworkers who provided feedback [34].

Then, we introduced freelancers to the version of the CrowdFolio feedback presentation interface they were assigned (i.e., holistic or decomposed) and asked them to explore the interface while thinking aloud. To understand if participants' self-perceptions had changed, we asked them to write another self-description, three phrases to tag their work, and a three-point action plan for improving their portfolio. Then, participants completed a post-task survey, in which they rated several items, such as the perceived helpfulness and trustworthiness of CrowdFolio's feedback (see Supplemental Materials). Participants also described their experience in the study in a semi-structured interview. All study sessions were conducted via an online video conferencing application and recorded.

6.3 Analyses

We quantitatively and qualitatively analyzed feedback collected from the crowd, participants' survey and study task responses, and participants' interview and think-aloud transcripts. Following a grounded theory approach [68], we wrote memos during and after each freelancer study session. After the first 15 participants, several themes emerged around participants' use of the CrowdFolio feedback presentation interface, the benefits and limitations of CrowdFolio compared to other sources of feedback, and the benefits and limitations of the holistic and decomposed conditions. We proceeded to group quotes from subsequent memos within these broader categories.

In addition, we quantitatively analyzed participants' survey responses, the feedback participants received, and participants' pre- and post-task self-descriptions and tags. We used independent t-tests to understand differences between participants' survey answers in both study conditions. Then, we computed the length and similarity of the pre- and post-task self-descriptions, as well as the length and quantity of the feedback participants received (i.e., style words, comments justifying style words, and perceived target audience comments).

To provide additional insight into participants' action plans and three tags, we developed coding schemes based on common themes in participants' responses (Tables 5, 6, and 7). To analyze participants' tags, we noted 1) the focus of the tag (i.e., referring either to a person's work or a personal characteristic) and 2) the nature of the changes made to the tags. To refine the coding schemes, two members of the research team independently coded 10% of the action items ($n=9$), individual tags ($n=19$), and sets of tags before and after the study task ($n=3$), following prior work (e.g., [34, 37]). The raters achieved substantial inter-rater reliability on the action items ($\kappa = 0.85$), tag focus ($\kappa=0.72$), and the types of changes made to tags ($\kappa=1.00$, [21]). Then, one rater independently coded the remaining dataset. These analyses helped us better understand the impact of the CrowdFolio feedback presentation interface on participants' reflection processes.

In addition, we viewed all participants' pre- and post-task self-descriptions to understand the types of changes participants made. At a high level, these included adding, modifying, or removing elements from the descriptions (e.g., skills, style words, soft skills, work experience, location). As we were unable to ask all participants to explain these changes, we do not include a detailed analysis of this. Because we were primarily interested in CrowdFolio's impact on participants, we ran an additional analysis to understand whether participants adopted phrases from CrowdFolio in their self-descriptions. We developed an algorithm in R using the stringr package [3] to note when a self-description included a) a phrase from the participant's Styles feedback page or b) one of the skills on the Target Audience page (e.g., illustration, UI). We then manually checked whether the matched phrases had also been used in their pre-task self-description; if not, we considered that the participant had adopted a new phrase from the CrowdFolio interface.

Type of Change	Hol.	Dec.	Description	Example Action Item
Add/Remove	11	5	Suggests adding or removing projects from a portfolio to show a specific type of design skill or varieties of design skills	"I want to include more branding design projects as opposed to personal art projects."
Update	5	8	Suggests updating portfolio by adding new projects or removing older projects	"Add new projects. And also remove old projects."
Projects	0	3	Suggests other changes to individual projects	"Write descriptions/design process."
Project Thumbnails	12	16	Suggests a change to some or all project thumbnails	"Update my Behance thumbnails to better show what TYPE of work each project was (i.e. website, digital ad campaign, branding suite...)"
Tags	5	0	Suggests a change to some or all of the tags for their projects	"adjust my tags for accuracy"
Profile	8	6	Suggests a change to some other aspect of their portfolio	"Adjust bio/description to match with my intended audience/potential clients"
Process	3	4	Suggests changes to other activities that could influence employability, such as learning new skills	"I do want to get back into learning how to use lighting within After Effects since I haven't done anything involving lighting from the program in a long time."
None	1	3	Makes no suggestions for any changes to their portfolio	"NA"

Table 5. Coding scheme for participants' action plan items based on the type of change that was suggested. The number of action items with each code are listed for the holistic (hol) and decomposed (dec) conditions respectively.

Focus of Tag	Description	Holistic		Decomposed	
		<i>M(SD)</i> (before)	<i>M(SD)</i> (after)	<i>M(SD)</i> (before)	<i>M(SD)</i> (after)
Work	Describes the type of work one does (e.g., colorful, contemporary)	2.20 (1.21)	2.33 (1.05)	2.07 (1.28)	2.27 (1.10)
Person	Describes a personal characteristic or the freelancer (e.g., graphic artist, passion for learning)	0.80 (1.21)	0.47 (0.83)	0.93 (1.28)	0.73 (1.10)

Table 6. Coding scheme for the focus of participants' three phrases they would use to tag their work online. We note the average number of tags per participant (out of three) with each focus.

Type of Changes Made	Description	Holistic	Decomposed
		<i>M(SD)</i>	<i>M(SD)</i>
Same Word or Phrase	Uses the same word or phrase (exact match or off by one letter) (e.g., artist → artists)	0.60 (0.51)	1.07 (1.10)
Modified Word or Phrase	Modifies the word or phrase by adding or removing parts of the phrase (e.g., web → web design)	0.80 (0.94)	0.73 (0.88)
Replaced Word or Phrase	Used a word or phrase completely different (i.e., had no words or phrase in common with the initial 3 tags) (e.g., trustworthy → fashion)	1.53 (0.99)	1.20 (0.94)

Table 7. Coding scheme for the types of changes participants made to their three tags. We note the average number of tags per participant (out of three) with each type of change.

7 RESULTS

7.1 Description of the Freelancer Participants

Participants were representative of graphic designers in the US [12, 64]. The gender distribution of the sample (56% female, 41% male, 3% non-binary) was similar to the population of graphic designers in the US in 2019 [12]. Most participants were between 18 and 34 years of age (70%,

$n=21$), had a bachelor's degree (50%, $n=15$), had or were completing a degree in visual or graphic design (60%, $n=18$), had more than five years of work experience (63%, $n=19$), had between one and three years of freelance work experience (47%, $n=14$), and freelanced part-time (63%, $n=19$). T-tests further confirmed that participants in both conditions were similar in terms of the importance of getting freelance work ($M_{\text{decomposed}} = 5.67$, $SD_{\text{decomposed}}=1.63$; $M_{\text{holistic}} = 6.13$, $SD_{\text{holistic}} = 1.36$, $t=-0.85$, $df=27.08$, n.s.) and sense of professional identity ($M_{\text{decomposed}} = 4.40$, $SD_{\text{decomposed}}=1.33$; $M_{\text{holistic}} = 4.50$, $SD_{\text{holistic}} = 0.62$, $t=-0.20$, $df=25.53$, n.s.).

In general, participants browsed creative work on Behance sporadically; about half of participants reported logging onto Behance at least once a month to once a year (46%, $n=14$) and most participants had logged onto Behance less than a month ago (83%, $n=28$). However, almost a third of participants had not updated their Behance portfolio in more than a year (30%, $n=9$). The majority updated their Behance portfolio at least once a month (37%, $n=11$) or once a year (37%, $n=11$), and also had a portfolio that was not on Behance (80%, $n=24$). Participants generally wanted to receive feedback on their portfolios more frequently than they currently did; the majority received feedback at least once a year (47%, $n=14$), but would have liked feedback at least once a month (47%, $n=14$).

7.2 RQ 1: Effects of Receiving Portfolio Feedback on Reflection and Portfolio Improvement Across Study Conditions

Online crowds provided a feasible way to collect rapid, regular, and inexpensive portfolio feedback that participants found helpful, trustworthy, and reflective of new perspectives. Feedback on a portfolio's visual attractiveness, styles, and target audience prompted participants to reflect on how aligned their portfolios were with their career goals; to add or remove projects from their portfolio; and to generate action plans to improve their portfolios.

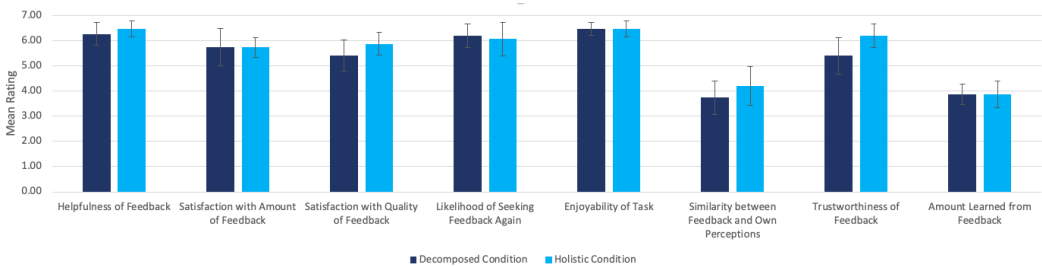


Fig. 4. We did not find any significant differences between participants in either condition on any of the survey measures asking about perceptions of CrowdFolio feedback. These survey items were measured on a scale of 1 (low) to 7 (high) with the exception of "Amount learned from feedback" which was measured on a scale of 1 (nothing at all) to 5 (a lot).

7.2.1 Participants Reported Learning from and Enjoying Portfolio Feedback. We did not find significant differences between participants in either condition on any of the survey measures (Figure 4). Participants in both conditions found the feedback to be helpful for noticing new things about their portfolio ($M_{\text{decomposed}}=6.27$, $M_{\text{holistic}}=6.47$, $t=-0.71$, $df=25.52$, n.s.) and considered feedback to be of moderate quality ($M_{\text{decomposed}}=5.40$, $M_{\text{holistic}}=5.87$, $t=-1.17$, $df=25.74$, n.s.; Figure 4). Participants were also highly likely to seek feedback like this in the future ($M_{\text{decomposed}}=6.20$, $M_{\text{holistic}}=6.07$, $t=0.32$, $df=25.16$, n.s.). Similarly, participants highly enjoyed exploring the feedback from CrowdFolio ($M_{\text{decomposed}}=6.47$, $M_{\text{holistic}}=6.47$, $t=0$, $df=26.80$, n.s.). They felt the feedback was only somewhat similar to their own perceptions ($M_{\text{decomposed}}=3.73$, $M_{\text{holistic}}=4.20$, $t=-0.89$, $df=27.53$, n.s.), found feedback moderately trustworthy ($M_{\text{decomposed}}=5.40$, $M_{\text{holistic}}=6.20$, $t=-1.79$, $df=23.98$, n.s.), and learned a

lot from using the interface ($M_{\text{decomposed}}=3.87$, $M_{\text{holistic}}=3.87$, $t=-0$, $df=26.53$, n.s.; survey item on a scale of 1: learned nothing at all to 5: learned a lot). In short, regardless of approach, participants found the feedback helpful for gaining insights about their portfolio.

7.2.2 Perspectives From the Crowd are Useful, When They Align with Potential Clients. Through CrowdFolio, we found that crowdsourced portfolio feedback afforded several benefits and limitations to freelancers. As we anticipated, freelancers perceived feedback from the crowd to be uniquely valuable and critical. Moreover, because feedback was collected from multiple crowdworkers, freelancers felt confident that the feedback they received was not too heavily influenced by a single feedback provider's opinions. However, CrowdFolio was not as helpful to freelancers who felt that their potential clients had substantial domain expertise, compared to freelancers who felt their potential clients had less domain expertise. We describe these findings below.

Compared to Other Feedback Sources, Crowds Provide Valuable Alternative Perspectives. While it was not a focus of the current study, we additionally asked participants to compare CrowdFolio to other feedback practices to understand how the system might complement these practices. For several participants ($n=7$), the CrowdFolio feedback presentation interface filled a need for portfolio feedback from non-professional audiences. One freelance graphic designer with less than three years of experience said that the crowd could provide a valuable alternative perspective from someone with more graphic design experience:

“... the majority [of them] didn't have experience with designers... Someone who is a graphic designer would more so focus... only on the composition, whereas a regular person wouldn't do that.” (P21)

Another participant said that they valued having their portfolio reviewed by the crowd, who provided access to critical feedback that they had not received since attending college:

“I miss that the most from [school]—having someone rip [my portfolio] apart...” (P24)

Another part-time freelancer who asks their friends and family for feedback said that crowd feedback allowed them to see patterns in opinions of their work, rather than being overly influenced by any single person's opinion:

“With friends and family, you are [only] talking to one person... it could be overly swayed by one person's view...if [only one person has an opinion on a project]...that could be their [own opinion].” (P10)

In short, compared to other sources of feedback, such as design professionals, friends, and family, freelancers appreciated CrowdFolio for the non-expert feedback, critical feedback, and feedback from multiple individuals.

Crowds May Not Reflect Potential Clients. Despite these benefits, some participants ($n=4$) were concerned that crowdworkers did not reflect their potential clients and were hesitant to accept the feedback they received. One participant with more than 10 years of experience felt that she could only trust designers to critique her work because designers would be more familiar with the needs of creative directors, who make up the majority of her client base:

“Part of the difference [between feedback from my friends and crowd feedback] is that my friends are designers [and they] know what creative directors would be looking for...” (P13)

However, freelancers whose clients had less design knowledge may have been more satisfied with the crowd feedback. One participant who was trying to reach clients with less design knowledge was satisfied with the feedback, especially the layman's terms (e.g., “*psychedellica*”) used to describe her work on the Styles page:

“A lot of people don’t know the difference between ‘illustration’ and ‘graphic design’. I love that word ‘psychedelica’ and that sums [my work] up a whole lot better than ‘whimsical’, [which I used in my first self-description].” (P15)

7.2.3 Participants Used Feedback to Plan Changes to Projects in their Portfolio. Although the CrowdFolio feedback presentation interface did not explicitly suggest ways freelancers could improve their portfolio, it nevertheless helped freelancers recognize these opportunities through three dimensions of feedback.

Visual Attractiveness Rankings Prompted Participants to Reflect on the Appeal of their Work. All participants ($n=30$) used rankings on the Visual Attractiveness page to infer their audience’s preferences and inform changes to their portfolio. For example, one participant specializing in logo design and branding guessed that participants were attracted to project thumbnails that were more abstract compared to those that showed an end product, such as a mobile ad:

“I’m definitely seeing...a correlation between things that would garner more interest... some of the things here are a little bit more abstract; they’re more of a teaser image versus things that are a little more clear cut...an email blast or a mobile ad, they rank a little bit lower.” (P18)

In line with prior work (e.g., [34]), some participants made these inferences when they were surprised by the crowd’s feedback. For example, one participant who had been freelancing part-time for less than three years was surprised that one of their project thumbnails, which they thought was simplistic, had been well received by the crowd:

“I’m actually surprised this one got the highest. It’s the most minimalistic. Maybe that’s what people are into. It’s less overwhelming, [it] leaves more to the imagination.” (P22)

In short, participants used the Visual Attractiveness page to speculate what elements of their work might be well received, leading naturally to action items to showcase these elements and downplay others.

Target Audience and Styles Feedback Validated Participants’ Career Direction. Many participants ($n=12$) also used feedback from the Target Audience and Styles pages to seek validation on their career direction. For example, one freelancer with more than 10 years of experience said she felt more comfortable describing herself as a digital or multimedia designer after learning that crowdworkers saw her as a specialist in digital advertising design on the Target Audience page:

“I never think of myself as a multimedia or digitally focused [designer]. But I think it’s OK [to be] pushing that [multimedia focus] a bit more [if other people are thinking that].” (P9)

Another designer who had been freelancing for less than a year felt positive after seeing that the crowd had categorized him as a motion graphics designer on the Target Audience page:

“So a lot of people do seem to realize that I specialize in motion graphics, which I am trying to be.” (P30)

When the feedback from the Styles and Target Audience pages did not validate participants’ perceptions of their own work, participants became motivated to improve their portfolio to clarify their career direction. For example, one freelancer who was transitioning toward three-dimensional design learned that their portfolio did not portray this career direction well after seeing disparate style phrases being used to describe his portfolio (e.g., “text only”, “typographic”, and “3D”). This prompted him to clarify his focus in his portfolio:

“From their adjectives [on the Styles page], it seemed like I could solidify my brand, my specialties and focus... How do I make that make sense for people, other than [saying] I’ve done a lot of stuff in different areas over the years, but I am trying to focus on 3D right now, moving forward?” (P25)

Another participant who had not updated his portfolio in more than a year was motivated to remove his portfolio entirely after seeing the misalignment between the crowd’s perceptions on the Target Audience and Style pages and his current career direction. For example, on the Target Audience page the crowd perceived his primary industry as retail, healthcare, publishing, broadcasting, or computer manufacturing, but not his current intended industry (i.e., software). He then described his intent to remove his portfolio entirely:

“I should remove my portfolio entirely if it’s not up to date. It does not represent me accurately.” (P19)

Portfolio feedback from the CrowdFolio feedback presentation interface encouraged participants ($n=10$) to add or remove projects by prompting them to reflect on aspects of their professional identity that were less apparent in the portfolio. For example, one freelancer specializing in branding was surprised that crowdworkers saw her as a digital advertising designer on the Target Audience page. She imagined that adding branding projects to their portfolio would lead more crowdworkers to categorize her as a ‘branding’ designer. In the following quote, she points to different columns in a graph summarizing her portfolio’s perceived target industry:

“Digital ads’... that’s really interesting, because I do more ‘branding.’ It’s probably because I don’t put a lot of my branding stuff. Yeah, I think if [the crowd] knew more about my stuff, they would put more here, here, here, and here (pointing to bars in a graph on the Target Audience page)...” (P22)

The Target Audience feedback also motivated some participants to make changes to their portfolio after realizing that they were limiting the types of audiences they were targeting. For example, one participant noticed that they shared the gender, age range, and income level of the perceived target audience. This motivated them to make changes to their future projects to reach a more diverse audience:

“[The feedback said] that they thought my target audience was essentially me... [For] me to be more successful, I need to break free from designing for myself, and... reach a broader scope or audience.” (P17)

In conclusion, the Target Audience and Styles feedback helped participants seek validation on their career direction and prompted them to add or remove projects from their portfolio when the feedback conflicted with their goals.

7.2.4 Some Participants Made Changes to their Tags and Self-Descriptions After Using CrowdFolio.

At least some participants made changes to their three tags and self-descriptions after viewing the CrowdFolio feedback interface. In general, participants were more likely to use tags that described characteristics of their work (e.g., modern), rather than themselves (e.g., hard-working) (see Table 6 for a breakdown by study condition). In addition, participants modified 0.77 of their tags ($SD=0.90$) and replaced 1.37 tags on average ($SD=0.97$) after using CrowdFolio (see Table 7 for a breakdown by study condition). For example, one participant in the holistic condition replaced all three of her tags, based on feedback that suggested her target audience was female:

“Presentation Designer, Branding Design, Experienced Creative Designer” (P9, pre-task tags)

“luxury, female-focused, modern” (P9, post-task tags)

The majority of participants ($n=26$, 86.7%) made at least some changes to their self-description, which included adding, modifying, or removing elements from the description (e.g., skill or specialty, aesthetic or style word, soft skills, work experience, location). Our textual analysis revealed that 36.7% of participants ($n_{\text{holistic}}=7$, $n_{\text{decomposed}}=4$) had used a phrase from the Styles or Target Audience page in their post-task self-description. For example, one participant in the holistic condition changed how she referred to herself by adding several words from her Styles and Target Audience pages (i.e., “abstract,” “psychedelica,” “digital,” “print”):

“I am a traveling artist whose work is influenced by nature and the outdoors. I have a whimsical, contemporary illustration style, and I do graphic design work as well.” (P15, pre-task description)

“I create vibrant illustrations and text designs based on travel and nature. Some of my work crosses over into abstract psychedelica. Finished pieces are good for print and digital products and marketing efforts.” (P15, post-task description)

Another participant in the holistic condition also added the phrase “motion graphics” from the Target Audience page in his post-task self-description:¹

“My name is Pat Stevens. I am a self-taught designer in Northern California with over 20 years experience. My skillsets include design, illustration, art direction, UI/UX, strategy and much more.” (P11; pre-task self-description)

“I am Pat Stevens. I am a self-taught designer in Northern California with over 20 years experience. My skillsets include graphic design, illustration, art direction, UI/UX, strategy, *motion graphics* and more” (P11; post-task self-description; emphasis added)

Besides adopting phrases from CrowdFolio, participants added other information about themselves to their self-descriptions. One participant in the decomposed condition added information about his location, his role as a trainer in Adobe software, and his skill in brand identity, after considering these additional facets about himself during the study:

“I’m a design craftsman and typographer with special interest in book design and large format.” (P7; pre-task self-description)

“An independent Los Angeles-based designer and Adobe trainer specializing in book design and typography, brand identity and large format design.” (P7; post-task self-description)

In contrast, a participant in the decomposed condition shortened her self-description by removing a call-to-action and describing her skills and role as a creative director:²

“Claire Sandoval is a Santa Monica based creative who would like it to be summer all of the time. Reach out to her for work in branding, graphic design, and typography.” (pre-task self-description)

“Claire Sandoval is a Santa Monica based creative director who crafts premium brands for a variety of audiences.” (post-task self-description)

While there were some changes to participants’ self-descriptions, we still found that participants wrote self-descriptions of similar length before ($M_{\text{pre-task}}=30.7$, $Mdn_{\text{pre-task}}=31.0$, $SD_{\text{pre-task}}=13.9$) and after the study task ($M_{\text{post-task}}=28.9$, $Mdn_{\text{post-task}}=27.5$, $SD_{\text{post-task}}=13.9$, $t=-0.51$, $df=58.00$, n.s.). Furthermore, we learned that many participants ($n=16$) still struggled to write a self-description because they lacked a clear idea of how to describe themselves succinctly for an online audience. Some even chose to avoid the initial self-description writing task and copied pre-written descriptions

¹We have used a pseudonym in place of the participant’s real name.

²We have used a pseudonym in place of the participant’s real name.

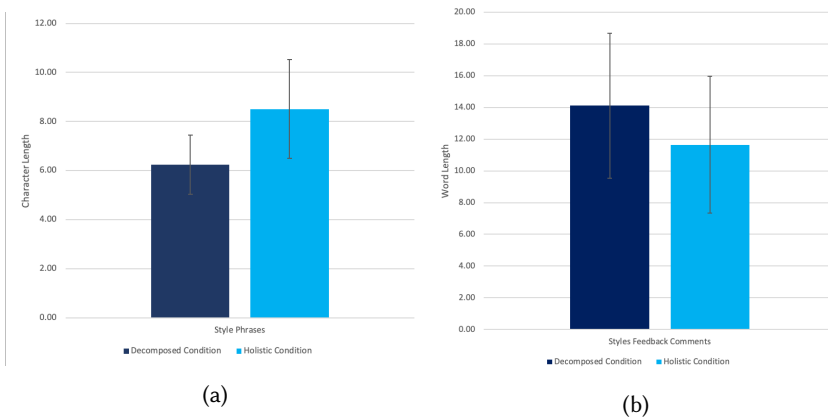


Fig. 5. (a) Style phrases collected in the holistic condition were significantly longer than style phrases in the decomposed condition. (b) Comments on the style phrases collected in the decomposed condition were significantly longer than comments in the holistic condition.

from their personal websites ($n=4$). One participant with more than three years of work experience struggled to describe himself because it was challenging to generate unique phrases that succinctly communicated the variety of work he did:

“What was difficult... is setting yourself apart and recognizing the things that would be different about you that would be appealing to a potential client [and] trying to think of a way to quickly and succinctly do this... I think [this is challenging] because my work tends to be not super focused. If I was a graphic designer who specialized in logos, it would be much easier to say.” (P14)

Additionally, some participants said they did not use the style phrases generated by the crowd in their self-descriptions because they were unsure of how unique and acceptable the phrases were relative to the broader professional design community. Other participants did not use the suggested style phrases because they felt some style phrases were too general to help clients to find their work. For instance, one participant with more than 10 years of experience felt that the style phrases were not as helpful as the search terms and tags he would have chosen for himself:

“Some [of the words on the Styles page] were really subjective adjectives... none of them felt like SEO (Search Engine Optimized) keywords. One guy just said the word ‘good’, and that’s good. But it’s not really [helpful]... [I think] ‘gig posters’ is a good [tag]. It is very specific and people who are searching for this stuff are people going to search for that.” (P11)

These results suggest that, while CrowdFolio feedback may serve as a useful *internal* reflection tool, composing self-descriptions to represent oneself online requires an additional understanding how one’s work compares to others and the technical details of how clients find and hire freelancers.

7.3 RQ 2: Effects of Holistic and Decomposed Approaches to Collecting Feedback

In this section, we describe how the holistic and decomposed approaches for collecting CrowdFolio feedback influenced the characteristics and perceptions of the Styles, Visual Attractiveness, and Target Audience feedback, as well as participants’ self-descriptions, tags, and action plans.

7.3.1 Styles Feedback.

Holistic Approach Led to Longer Style Phrases but Decomposed Approach Led to Longer Comments. On average, the style phrases in the holistic condition were significantly longer than the style phrases in the decomposed condition ($M_{\text{decomposed}}=6.24$ characters, $M_{\text{holistic}}=8.51$ characters, $t=-10.58$, $df=486.05$, $p<.001$; Figure 5). On the other hand, comments explaining style phrases contained more words in the decomposed condition ($M_{\text{decomposed}}=14.11$) compared to the holistic condition ($M_{\text{holistic}}=11.65$, $t=3.70$, $df=682.81$, $p<0.001$). The following is an example of two style phrases and comments on one participant's portfolio:

'Latin Illustration': 'A lot of these illustrations had Mexican, central American, and south american themes. They also had elements of aztec and mayan cultures. Learning about those styles could also give some insight to the cultures there and how and why they identify with the indigenous peoples of the past.' (holistic condition)

'Colorful': 'This is a cartoon picture of the Aztec people. It is colorful and beautiful.' (decomposed condition)

While it is promising that the CrowdFolio feedback collection interface was able to collect specific comments, about 10% of the comments in both conditions were irrelevant or seemed to be copied from other sources. For example, one crowdworker provided the following comment to justify their agreement with a style phrase (i.e., "winter") describing one participant's projects: "*The length of the astronomical seasons varies between 89 and 93 days, while the length of the meteorological seasons is less variable and is fixed at 90 days for winter in a non-leap year (91 days in a leap year), 92 days for spr.*" It seems that the crowdworker had copied the definition of "winter," rather than explaining why they agreed that "winter" appropriately described the freelancer's project. This reflects a larger challenge in crowdsourcing of managing the variability in the quality of responses from crowdworkers (e.g., [51]).

Holistic Approach Helped Participants Understand Project Themes through Styles Feedback. Participants in the holistic condition used the Styles feedback to discover new ways of describing their style ($n=8$). Prior to using the CrowdFolio feedback presentation interface, one digital media designer shared that she did not feel her work had many positive or unique qualities, despite having more than 10 years of work experience. She described herself in terms of general qualities: "*Focused on clients' needs and bringing their visions to fruition. Adept at synthesizing multiple modes of media.*" (P5) However, after viewing the feedback on the Styles page, she identified a unique focus of her work, which was working on multicultural community projects:

"Oh that's cool, I really like the 'cultural and activities' [style phrase]. These three [projects] are all from the same job that I had... a long time ago. I guess that's kind of cool that I have a certain amount of multicultural type stuff on there. I like that. I didn't think about it that way before so that's cool." (P5)

This feedback prompted her to create a point in her action plan to change the way she grouped her projects in their portfolio: "*Change the groupings to highlight some of the positive qualities that the people in the study talked about such as colorful, retail, multicultural, etc*" (P5). After viewing the feedback, she was also motivated to use more specific words to describe herself: "*Eye catching, easy to understand targeted retail advertisement. Community focused, appealing to a younger demographic with bold and colorful design*" (P5). Feedback collected through the holistic approach helped her notice connections between her projects and discover new ways of describing her work.

Participants in the holistic condition were also motivated to add projects to their portfolio after viewing the Styles feedback. One full-time freelancer who specialized in creating various kinds of reports was surprised to see that crowdworkers had grouped many of her projects under the single

style phrase “*business docs*.” This freelancer assumed that the crowd did not see her as a versatile designer. Hence, this freelancer decided to add more variety to her portfolio:

“Seeing how people were interpreting how my documents were makes me want to give more examples [of the types of documents I can design].” (P31)

Other participants in the holistic condition used Styles feedback to understand which projects might need to be clarified. One full-time freelancer who used live-streaming to build his online audience on Behance used the feedback to recognize when some of his projects were missing from the style groups:

“I wish these first two [projects] were in [the] ‘*art direction*’ [group], because that’s kind of what I’m after – creating content that would be used in branding.” (P29)

In sum, the holistic approach to collecting feedback helped participants easily notice stylistic themes between their projects and take action based on those themes.

Decomposed Approach Provided Specific Feedback on a Smaller Range of Projects. In contrast, some participants in the decomposed condition found the Styles feedback to be too focused on describing individual projects. One participant in the decomposed condition with less than a year of freelance experience expressed the following:

“...the [style] words that they chose were a little bit too narrow in scope. Like [the phrase] ‘*valentines*’ and only having one [project] in that section.” (P17)

The decomposed feedback condition also led freelancers to receive feedback on a smaller range of projects. In this condition, crowdworkers provided style phrases to describe a random project from the portfolio of a randomly selected freelancer participant. As a result, not all projects in freelancers’ portfolios received Styles feedback. Several participants ($n=5$) in the decomposed condition were unsatisfied when they noticed that some of their projects had not received feedback from the crowd:

“I was slightly dissatisfied. I wish I had feedback on that illustrated map [project]...[and] on those recent projects. The older ones were trash.” (P16)

Despite these drawbacks, participants in the decomposed feedback condition received more style phrases on average ($M=39$ words) than participants in the holistic feedback condition ($M=12$ words). However, this was not surprising given that crowdworkers in the decomposed condition were asked to provide three different style phrases for each project they viewed, while crowdworkers in the holistic condition provided three style phrases for an entire portfolio.

7.3.2 Visual Attractiveness Feedback. The decomposed approach provided more detailed feedback regarding the visual attractiveness of project thumbnails in participants’ portfolios, but may have also led participants to make inaccurate judgments about the crowd’s preferences. In this study condition, crowdworkers rated a project thumbnail on a scale of 1 (not interested to learn about the project at all) to 5 (very interested to learn about the project). As a result, freelancers got a better sense of how visually attractive each project was on its own. When ratings were low, freelancers still gained a sense of their performance and skill level. For example, a freelance web designer in the decomposed condition used these detailed ratings to assess whether their own perceptions were aligned with the crowd’s:

“Threes and twos, which is not that great... The [project] I have been the most proud of only has a one, so that’s kind of interesting to see. I like how another [project] I was pretty proud of didn’t get a one. I think those two are the ones I was most interested in seeing the feedback for.” (P28)

In contrast, participants in the holistic condition could only learn which projects were more visually attractive than others, but not how attractive projects were overall. We did not observe

any participants in the holistic condition making statements about how visually attractive their portfolio was overall.

At the same time, the decomposed approach may have led participants to make inaccurate judgments about the crowd's preferences. Some participants in the decomposed condition assumed that crowdworkers had compared multiple projects in their portfolio when this was not the case. This could have led participants to make inaccurate inferences. For example, one full-time freelancer in the decomposed condition inferred that one of her projects received the highest rating because it "stood out" as a black-and-white project against her other colorful projects:

"[This] black and white Instagram concept project is the highest rated. I don't know if that's because it has the title in it... or maybe with everything else being color, it just stands out more?" (P12)

In short, the decomposed approach provided detailed Visual Attractiveness feedback on standalone projects, but participants seemed to assume that crowdworkers could compare their projects when providing the feedback.

7.3.3 Target Audience Feedback. We did not observe substantial differences in the way participants used Target Audience feedback collected through different approaches. Independent t-tests revealed that comments about the target audience in both conditions were about the same length in characters ($M_{\text{decomposed}}=86.33$, $M_{\text{holistic}}=94.79$, $t=-1.71$, $df=526.56$, n.s.) and words ($M_{\text{decomposed}}=15.74$, $M_{\text{holistic}}=16.95$, $t=-1.36$, $df=529.66$, n.s.). A high percentage of these target audience comments overall (84.41%, $n=536$) were specific (i.e., more than five words long) [62]. The following is an example of two comments about one participant's perceived target audience in both conditions:

"These images look like they would appeal to inner city kids who enjoy comic books because of the graffiti like style." (holistic condition)

"The main audience would probably be people looking for somewhere to eat when they across this type of image." (decomposed condition)

As described earlier, the Target Audience feedback prompted participants in both conditions add or remove projects to better convey their identity. We did not observe any substantial differences in the way participants perceived this feedback across conditions.

7.3.4 Feedback Collection Approach Did Not Influence Self-Descriptions, Tags, and Action Plans. To check if one approach to collecting feedback resulted in more change in participants' self-perceptions than the other, we analyzed the length and similarity of participants' self-descriptions before and after viewing CrowdFolio feedback. Using a t-test, we found that participants in both conditions changed approximately the same number of words in their post-task self-descriptions ($M_{\text{decomposed}}=-3.47$, $SD_{\text{decomposed}}=8.54$, $Mdn_{\text{decomposed}}=0$, $M_{\text{holistic}}=-0.2$, $SD_{\text{holistic}}=9.83$, $Mdn_{\text{holistic}}=2$, $t=-0.97$, $df=27.47$, n.s.). Using the SequenceMatcher class in Python [35], which provides a character similarity score of 1 for phrases that are exactly the same and 0 for phrases that are completely different, we also found no significant differences between the degree of character similarity of the self-descriptions in either the decomposed ($M_{\text{decomposed}}=0.48$, $Mdn_{\text{decomposed}}=0.36$, $SD_{\text{decomposed}}=0.26$) or holistic conditions ($M_{\text{holistic}}=0.44$, $Mdn_{\text{holistic}}=0.40$, $SD_{\text{holistic}}=0.24$, $t=0.36$, $df=27.77$, n.s.).

Our analysis of the tags suggested similar results. Participants in both conditions were equally likely to focus on describing their work (versus personal characteristics) in their tags; a Wilcoxon Rank-Sum test did not find significant differences in the focus of participants' post-task tags between conditions ($W_{\text{work-focus}}=110.5$, n.s.; see Table 6). They were also equally likely to modify or replace words in their tags after using CrowdFolio ($W_{\text{tags modified}}=109$, n.s.; $W_{\text{tags replaced}}=90$, n.s.; Table 7). Hence, our results suggest that both the holistic and decomposed approaches to collecting crowd feedback led to few differences in the way freelancers described themselves.

In addition, there were no significant differences in participants' action plans across study conditions. Participants in both conditions frequently suggested action steps that were related to changing their project thumbnails (31.1% of action items), adding or removing projects to emphasize certain skills (17.8%), changing other aspects of their Behance profile (15.6%), and updating their Behance portfolio by either adding more recent projects or removing older projects (14.4%). Although slightly more participants in the holistic condition suggested action items related to adding or removing projects (24.4%) compared to the decomposed condition (11.1%), a chi-square analysis found that these differences were not statistically significant ($\chi^2=12.94$, $df=7$, n.s.). While participants in the decomposed condition wrote action items that were longer than those in the holistic condition, these differences did not reach statistical significance ($M_{\text{decomposed}}=28.8$ words, $Mdn_{\text{decomposed}}=20.3$, $SD_{\text{decomposed}}=32.4$; $M_{\text{holistic}}=13.7$ words, $Mdn_{\text{holistic}}=13$, $SD_{\text{holistic}}=5.58$, $t=1.78$, $df=14.83$, n.s.). We calculated this by taking the average length of each participant's action items. Hence, the approach to collecting feedback did not have a significant effect on the action plans participants made.

8 DISCUSSION

In this study, we built on prior work showing how the crowd can provide rapid, regular, and inexpensive feedback by exploring the promise of systems for collecting crowd feedback on *portfolios* of design work. We studied two approaches to collecting portfolio feedback: holistic, in which crowdworkers assessed an entire portfolio at once, and decomposed, in which crowdworkers assessed individual projects. Each approach provided its own benefits; the holistic approach provided insight into stylistic themes present across projects, while the decomposed approach provided detailed feedback on the visual attractiveness of individual work. Nevertheless, both approaches provided freelancers opportunities to discover gaps between their own and the audience's perceptions of their portfolio, complementing existing crowd feedback systems targeted at improving individual projects (e.g., Voyant [73], CrowdCrit [57]).

8.1 Holistic and Decomposed Approaches Helped Users Understand Gaps in Perceptions of Their Work

We expected that the holistic and decomposed approaches to collecting feedback would lead to strong differences in the characteristics and perceptions of the feedback. However, regardless of the way feedback was collected, CrowdFolio helped users *understand gaps* between their perceptions and the crowds' perceptions. For example, the Target Audience feedback in both conditions helped freelancers understand the crowd's perceptions of their professional identity. Because the portfolio feedback was anonymous and aggregated from multiple opinions further reassured freelancers that these reflected newer perspectives than if they had come from friends or family. We speculate the aggregation of feedback on these three dimensions and the anonymity of the crowd contributed to the similarities in effects of the CrowdFolio feedback presentation interface across both conditions.

8.2 Holistic and Decomposed Approaches Helped Freelancers Gain Different Insights about their Portfolio

Nevertheless, our study results suggest that holistic and decomposed approaches to collecting feedback can still present some unique benefits and challenges to freelancers. The holistic approach led to longer style phrases, which we speculate may have been due to crowdworkers describing multiple projects at once. Crowdworkers in the decomposed condition may have also found it more difficult to generate three unique phrases to describe a single project. In addition, participants' think-aloud transcripts suggested that the holistic approach provided insight into the stylistic themes present *across* projects, as crowdworkers were able to compare projects as they made

their impressions. Although participants in the decomposed study condition *could* have manually aggregated the style phrases themselves, we found little evidence from our transcripts that they did so. In fact, this decomposed approach reflects existing practices on platforms, such as Behance, where users manually aggregate comments on their work. In future research, we will explore the potential for using other crowd clustering and machine learning approaches (e.g., [17]) to automatically aggregate style phrases and compare these to phrases from the holistic study condition.

Our study transcripts also suggested that the decomposed approach provided freelancers with helpful, detailed information on the visual attractiveness of their projects. We attribute this to the design of the feedback task—in the decomposed condition, crowds rated the visual attractiveness of each project separately. However, the *holistic* approach may have been more valid in approximating how freelancers expect portfolio feedback to be given. A future version of CrowdFolio could ensure that all projects in a portfolio receive feedback at least once before launching tasks showing repeat projects, but this approach would be roughly four times more expensive and time-consuming than the holistic condition (i.e., crowdworkers in the decomposed condition provided feedback on only five projects, but in the holistic condition provided feedback on up to 20 projects at a time).

8.3 Design Implications for Crowd Feedback Systems

Our findings have implications for the design of future crowd feedback systems for *holistic* sets of creative work. First, system designers must consider the goals of their system (e.g., [74]). To encourage reflection on individual items in the set, consider using a decomposed approach; in contrast, to enable the discovery of underlying categories, use a holistic approach. Echoing prior research in crowdsourcing (e.g., [51]) these systems must also ensure equal coverage across a set of projects in a portfolio, either by asking crowdworkers to evaluate items holistically or distributing tasks equally among the set. System designers should consider the relative benefits and challenges of these approaches when designing crowd feedback systems.

Our work demonstrates that portfolio-based feedback systems can prompt freelancers to reflect on their professional identities in a way that is not possible with project-based crowd feedback systems (e.g., [57, 72]). CrowdFolio prompted freelancers to plan holistic changes to their portfolios, such as adding or removing projects to better showcase one’s primary specialty. In contrast, project-based systems excel at prompting visual changes to individual projects, such as changing the color or layout of a poster [34, 57]. Instead of assessing a freelancer’s work on established design principles, CrowdFolio provided first impressions of freelancer’s work as a whole, allowing freelancers to seek additional information to improve their portfolio, in line with expert processes of sensemaking (e.g., [34, 53]). Some might not even consider the portfolio feedback provided by CrowdFolio to be of “high quality” by the standards of project-based crowd feedback systems (e.g., [57, 62]). For example, the CrowdFolio feedback collection interface did not ask crowd workers to provide actionable feedback. Nevertheless, this feedback was sufficient to help freelancers notice gaps between their own and the crowd’s perceptions of their professional identity. In these ways, our work extends the use of crowd feedback systems beyond decontextualized, decomposed tasks to more contextualized tasks with multiple projects [33]. We also contribute to a growing body of literature on how crowds can support individuals in achieving a variety of personal goals, from writing a research paper [61] to forming healthy habits [11].

8.4 Future of Portfolio Feedback Systems as a Part of Freelancers’ Lives

In line with prior work, we believe that crowd feedback systems represent only one tool in a toolkit that freelancers use to augment their online self-presentation [25, 69]. For example, during the study, about half of participants used a separate online tool called Grammarly [2] to help them remove grammatical errors from and improve the clarity of their self-descriptions. Moreover, crowdsourced

portfolio feedback may be more helpful if the feedback providers share key characteristics of one's potential clients, particularly if users already have a clear idea of the kinds of clients they want to attract. Our interviews also suggested that crowd feedback complemented existing feedback sources (e.g., friends, family, design professionals) by providing alternative perspectives at a larger scale. However, due to the focus of the current study, we were unable to directly compare CrowdFolio to these existing practices. In future research, we will study how crowd feedback systems might guide freelancers to write more clear, succinct, and unique self-descriptions. We will also explore the effects of having control over crowdworker demographics on the trustworthiness of the feedback. In addition, we plan to investigate how crowd feedback might be applied in other domains, such as web development or engineering, and how it directly compares to other feedback sources.

9 CONCLUSION

As online platforms expand freelancers' access to clients, they must be able to develop attractive, up-to-date portfolios to convey their skills and unique style. Portfolio feedback systems have the potential to change the way freelancers improve their portfolios and our study offers lessons in how to collect this feedback. By comparing holistic and decomposed approaches to collecting feedback in a portfolio-based crowd feedback system, CrowdFolio, our study transcripts suggested that the holistic approach allowed freelancers to discover stylistic similarities in their projects, while the decomposed approach provided freelancers with helpful detail about the visual attractiveness of their projects. Regardless of approach, portfolio feedback helped freelancers recognize opportunities to improve their portfolio, helped freelancers find new ways to describe their work, and filled a need for rapid, regular, and inexpensive feedback from a diverse, non-expert crowd. Our results show that it is possible to increase the complexity of crowd feedback systems to help users reach their professional career goals.

REFERENCES

- [1] [n.d.]. Amazon Mechanical Turk. <https://www.mturk.com/>
- [2] [n.d.]. Grammarly. <https://www.grammarly.com/> Library Catalog: www.grammarly.com.
- [3] [n.d.]. stringr package | R Documentation. <https://www.rdocumentation.org/packages/stringr/versions/1.4.0>
- [4] 2018. Behance :: Best of Behance. <https://www.behance.net/>
- [5] 2018. Dribbble. <https://dribbble.com/>
- [6] 2018. Facebook. <https://www.facebook.com/>
- [7] 2018. reddit. <https://www.reddit.com/>
- [8] 2019. Upwork. <https://www.upwork.com/>
- [9] Kumar Abhinav, Alpana Dubey, Sakshi Jain, Gurdeep Viridi, Alex Kass, and Manish Mehta. 2017. CrowdAdvisor: A Framework for Freelancer Assessment in Online Marketplace. In *Proceedings of the 39th International Conference on Software Engineering: Software Engineering in Practice Track (ICSE-SEIP '17)*. IEEE Press, Piscataway, NJ, USA, 93–102. <https://doi.org/10.1109/ICSE-SEIP.2017.23>
- [10] Robin S Adams, Jennifer Turns, and Cynthia J Atman. 2003. Educating effective engineering designers: the role of reflective practice. *Design Studies* 24, 3 (May 2003), 275–294. [https://doi.org/10.1016/S0142-694X\(02\)00056-X](https://doi.org/10.1016/S0142-694X(02)00056-X)
- [11] Elena Agapie, Lucas Colusso, Sean A Munson, and Gary Hsieh. 2016. PlanSourcing: Generating Behavior Change Plans with Friends and Crowds. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM. <https://doi.org/10.1145/2818048.2819943>
- [12] AIGA, Google, and Accurat. 2019. *Design Census 2019: Understanding the state of design and the people who make it*. Technical Report. AIGA. <https://designcensus.org/data/2019DesignCensus.pdf>
- [13] Paul André, Aniket Kittur, and Steven P. Dow. 2014. Crowd synthesis: extracting categories and clusters from complex data. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing (CSCW '14)*. Association for Computing Machinery, New York, NY, USA, 989–998. <https://doi.org/10.1145/2531602.2531653>
- [14] J.Scott Armstrong, William B. Denniston, and Matt M. Gordon. 1975. The use of the decomposition principle in making judgments. *Organizational Behavior and Human Performance* 14, 2 (Oct. 1975), 257–263. [https://doi.org/10.1016/0030-5073\(75\)90028-8](https://doi.org/10.1016/0030-5073(75)90028-8)

- [15] Carlie Barrett, Alex Broerman, Emmanuel Carillo, Morgen Depenthal, Kevin Doyle, Allison Dunphy, Matt Gill, Natalie Miklosic, Samantha Gustafson, Mei Hsieh, Aaron Kurosu, Rachel Lang, Andrew Leemhuis, Jon Panichella, Casey Schneider, James Sloss, Steve Speaker, Aaron Watkins, and Jacklyn Woniger. 2011. *Hire Me?! The Portfolio Handbook: A Guide to Creating Your Design Portfolio*. University of Cincinnati, Cincinnati, OH. <http://www.blurb.com/b/2516994-hire-me-the-portfolio-handbook>
- [16] Julia Cambre, Scott Klemmer, and Chinmay Kulkarni. 2018. Juxtapeer: Comparative Peer Review Yields Higher Quality Feedback and Promotes Deeper Reflection. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 294:1–294:13. <https://doi.org/10.1145/3173574.3173868>
- [17] Joseph Chee Chang, Aniket Kittur, and Nathan Hahn. 2016. Alloy: Clustering with Crowds and Computation. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 3180–3191. <https://doi.org/10.1145/2858036.2858411>
- [18] Lydia B. Chilton, Greg Little, Darren Edge, Daniel S. Weld, and James A. Landay. 2013. Cascade: crowdsourcing taxonomy creation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1999–2008. <https://doi.org/10.1145/2470654.2466265>
- [19] Kwangsu Cho and Christian D. Schunn. 2007. Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education* 48, 3 (April 2007), 409–426. <https://doi.org/10.1016/j.compedu.2005.02.004>
- [20] Bruno Coelho, Fernando Costa, and Gil M. Goncalves. 2015. HYRE-ME - Hybrid Architecture for Recommendation and Matchmaking in Employment. In *Information and Software Technologies (Communications in Computer and Information Science, Vol. 538)*, Dregvaite, G and Damasevicius, R (Ed.). 208–224. https://doi.org/10.1007/978-3-319-24770-0_19
- [21] Jacob Cohen. 1968. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin* 70, 4 (1968), 213–220. <https://doi.org/10.1037/h0026256>
- [22] Maria Daltayanni, Luca de Alfaro, and Panagiotis Papadimitriou. 2015. WorkerRank: Using Employer Implicit Judgements to Infer Worker Reputation. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15)*. ACM, New York, NY, USA, 263–272. <https://doi.org/10.1145/2684822.2685286> event-place: Shanghai, China.
- [23] Deanna Dannels, Amy Gaffney, and Kelly Martin. 2008. Beyond content, deeper than delivery: What critique feedback reveals about communication expectations in design education. *International Journal for the Scholarship of Teaching and Learning* 2, 2 (2008).
- [24] Michael A DeVito, Darren Gergle, and Jeremy Birnholtz. 2017. “Algorithms ruin everything”:# RIPTwitter, Folk Theories, and Resistance to Algorithmic Change in Social Media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3025453.3025659>
- [25] Tawanna R. Dillahunt, Jason Lam, Alex Lu, and Earnest Wheeler. 2018. Designing Future Employment Applications for Underserved Job Seekers: A Speed Dating Study. In *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18*. ACM Press, Hong Kong, China, 33–44. <https://doi.org/10.1145/3196709.3196770>
- [26] Shoshana R. Dobrow and Monica C. Higgins. 2005. Developmental networks and professional identity: a longitudinal study. *Career Development International* 10, 6/7 (Oct. 2005), 567–583. <https://doi.org/10.1108/13620430510620629>
- [27] Steven Dow, Elizabeth M Gerber, and Audris Wong. 2013. A pilot study of using crowds in the classroom. *Proceedings of the 31st Annual ACM Conference on Human Factors in Computing Systems* (April 2013), 227–236. <https://doi.org/10.1145/2470654.2470686>
- [28] Steven P Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L Schwartz, and Scott R Klemmer. 2010. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Transactions on Computer-Human Interaction* 17, 4 (Dec. 2010), 1–24. <https://doi.org/10.1145/1879831.1879836>
- [29] Brooke Erin Duffy, Urszula Pruchniewska, and Leah Scolere. 2017. Platform-Specific Self-Branding: Imagined Affordances of the Social Media Ecology. In *Proceedings of the 8th International Conference on Social Media & Society (#SMSociety17)*. ACM, New York, NY, USA, 5:1–5:9. <https://doi.org/10.1145/3097286.3097291> event-place: Toronto, ON, Canada.
- [30] Matthew Wayne Easterday, Daniel G Rees Lewis, and Elizabeth M Gerber. 2015. The logic of the theoretical and practical products of design research. *Australasian Journal of Educational Technology* 32, 4 (Feb. 2015), 1–21. <https://doi.org/10.14742/ajet.2464>
- [31] Nicole Ellison, Rebecca Heino, and Jennifer Gibbs. 2006. Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment. *Journal of Computer-Mediated Communication* 11, 2 (Jan. 2006), 415–441. <https://doi.org/10.1111/j.1083-6101.2006.00020.x>
- [32] Ethan Fast, Binbin Chen, and Michael S. Bernstein. 2016. Empath: Understanding Topic Signals in Large-Scale Text. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4647–4657. <https://doi.org/10.1145/2858036.2858535>
- [33] 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 2017 CHI Conference on Human Factors in*

- Computing Systems (CHI '17)*. Association for Computing Machinery, Denver, Colorado, USA, 4454–4467. <https://doi.org/10.1145/3025453.3025791>
- [34] Eureka Foong, Darren Gergle, and Elizabeth M. Gerber. 2017. Novice and Expert Sensemaking of Crowdsourced Design Feedback. *Proceedings of the ACM on Human-Computer Interaction* 1, 45 (Dec. 2017). <https://doi.org/10.1145/3134680>
- [35] The Python Software Foundation. 2020. difflib – Helpers for computing deltas – Python 3.7.7 documentation. <https://docs.python.org/3.7/library/difflib.html>
- [36] Alessandro Gandini. 2016. Digital work: Self-branding and social capital in the freelance knowledge economy. *Marketing Theory* 16, 1 (March 2016), 123–141. <https://doi.org/10.1177/1470593115607942>
- [37] Michael D Greenberg, Matthew W Easterday, and Elizabeth M Gerber. 2015. Critiki: A scaffolded approach to gathering design feedback from paid crowdworkers. In *the 2015 ACM SIGCHI Conference*. <https://doi.org/10.1145/2757226.2757249>
- [38] Jason Haberman and David Whitney. 2007. Rapid extraction of mean emotion and gender from sets of faces. *Current Biology* 17, 17 (Sept. 2007), R751–R753. <https://doi.org/10.1016/j.cub.2007.06.039>
- [39] Jason Haberman and David Whitney. 2009. Seeing the mean: Ensemble coding for sets of faces. *Journal of experimental psychology. Human perception and performance* 35, 3 (June 2009), 718–734. <https://doi.org/10.1037/a0013899>
- [40] Bruce Hanington and Bella Martin. 2019. *Universal Methods of Design Expanded and Revised: 125 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*. Rockport Publishers. Google-Books-ID: SFnBDwAAQBAJ.
- [41] Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage, Chris Callison-Burch, and Jeffrey P. Bigham. 2018. A Data-Driven Analysis of Workers’ Earnings on Amazon Mechanical Turk. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 449:1–449:14. <https://doi.org/10.1145/3173574.3174023>
- [42] Rebecca D. Heino, Nicole B. Ellison, and Jennifer L. Gibbs. 2010. Relationshopping: Investigating the market metaphor in online dating. *Journal of Social and Personal Relationships* 27, 4 (June 2010), 427–447. <https://doi.org/10.1177/0265407510361614>
- [43] David Hesmondhalgh and Sarah Baker. 2010. ‘A very complicated version of freedom’: Conditions and experiences of creative labour in three cultural industries. *Poetics* 38, 1 (Feb. 2010), 4–20. <https://doi.org/10.1016/j.poetic.2009.10.001>
- [44] David Hesmondhalgh and Sarah Baker. 2013. *Creative Labour : Media Work in Three Cultural Industries*. Routledge. <https://doi.org/10.4324/9780203855881>
- [45] Julie Hui, Amos Glenn, Rachel Jue, Elizabeth Gerber, and Steven Dow. 2015. Using anonymity and communal efforts to improve quality of crowdsourced feedback. In *Proceedings of the Third AAAI Conference on Human Computation and Crowdsourcing*. <https://doi.org/10.1108/TG-09-2013-0035>
- [46] Behance Inc. 2012. 6 Steps To Creating A Knockout Online Portfolio. <https://99u.adobe.com/articles/7127/6-steps-to-creating-a-knockout-online-portfolio>
- [47] Behance Inc. 2012. A Portfolio for Clients: A Beginner’s Guide to Getting More Gigs. <https://99u.adobe.com/articles/52662/a-portfolio-for-clients-a-beginners-guide-to-getting-more-gigs> Library Catalog: 99u.adobe.com.
- [48] Mohammad Hossein Jarrahi, Will Sutherland, Sarah Beth Nelson, and Steve Sawyer. 2020. Platformic Management, Boundary Resources for Gig Work, and Worker Autonomy. *Computer Supported Cooperative Work (CSCW)* 29, 1 (April 2020), 153–189. <https://doi.org/10.1007/s10606-019-09368-7>
- [49] Hyeonsu B. Kang, Gabriel Amoako, Neil Sengupta, and Steven P. Dow. 2018. Paragon: An Online Gallery for Enhancing Design Feedback with Visual Examples. ACM Press, 1–13. <https://doi.org/10.1145/3173574.3174180>
- [50] Joy Kim, Sarah Sterman, Allegra Argent Beal Cohen, and Michael S Bernstein. 2017. Mechanical Novel. In *the 2017 ACM Conference*. ACM Press, New York, New York, USA, 233–245. <https://doi.org/10.1145/2998181.2998196>
- [51] Aniket Kittur, Jeffrey V Nickerson, Michael Bernstein, Elizabeth M Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The future of crowd work. In *Proceedings of the 16th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, New York, New York, USA, 1301–1318. <https://doi.org/10.1145/2441776.2441923>
- [52] Aniket Kittur, Boris Smus, Susheel Khamkar, and Robert E. Kraut. 2011. CrowdForge: crowdsourcing complex work. In *Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11)*. Association for Computing Machinery, Santa Barbara, California, USA, 43–52. <https://doi.org/10.1145/2047196.2047202>
- [53] Gary Klein, Jennifer K Phillips, Erica L Rall, and Deborah Peluso. 2007. A data-frame theory of sensemaking. In *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making*. http://books.google.com/books?hl=en&lr=&id=GQK8mccWIVoC&oi=fnd&pg=PA113&dq=A+data+frame+theory+of+sensemaking&ots=cXazVII4PI&sig=Amx9bbxRp9g_J6eXwJzcY6LyKUU
- [54] 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, CSCW (Dec. 2017), 1–18. <https://doi.org/10.1145/3134695>

- [55] Richard D Lennox and Raymond N Wolfe. 1984. Revision of the Self-Monitoring Scale. *Journal of Personality and Social Psychology* 46, 6 (1984), 1349–1364. http://www.communicationcache.com/uploads/1/0/8/8/10887248/revision_of_the_self-monitoring_scale.pdf
- [56] Harold Linton. 2012. *Portfolio Design*. W.W. Norton & Company. Google-Books-ID: ZEQppwAACAAJ.
- [57] Kurt Luther, Jari-Lee Tolentino, Wei Wu, Amy Pavel, Brian P. Bailey, Maneesh Agrawala, Björn Hartmann, and Steven P. Dow. 2015. Structuring, Aggregating, and Evaluating Crowdsourced Design Critique. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, New York, NY, USA, 473–485. <https://doi.org/10.1145/2675133.2675283> event-place: Vancouver, BC, Canada.
- [58] Karen S. Lyness and Edwin T. Cornelius. 1982. A comparison of holistic and decomposed judgment strategies in a performance rating simulation. *Organizational Behavior and Human Performance* 29, 1 (Feb. 1982), 21–38. [https://doi.org/10.1016/0030-5073\(82\)90240-9](https://doi.org/10.1016/0030-5073(82)90240-9)
- [59] Jennifer Marlow and Laura Dabbish. 2014. From Rookie to All-star: Professional Development in a Graphic Design Social Networking Site. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 922–933. <https://doi.org/10.1145/2531602.2531651>
- [60] David Navon. 1977. Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology* 9, 3 (July 1977), 353–383. [https://doi.org/10.1016/0010-0285\(77\)90012-3](https://doi.org/10.1016/0010-0285(77)90012-3)
- [61] Michael Nebeling, Alexandra To, Anhong Guo, Adrian A. de Freitas, Jaime Teevan, Steven P. Dow, and Jeffrey P. Bigham. 2016. WearWrite: Crowd-Assisted Writing from Smartwatches. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, San Jose, California, USA, 3834–3846. <https://doi.org/10.1145/2858036.2858169>
- [62] Tricia J. Ngoon, C. Ailie Fraser, Ariel S. Weingarten, Mira Dontcheva, and Scott Klemmer. 2018. Interactive Guidance Techniques for Improving Creative Feedback. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press, Montreal QC, Canada, 1–11. <https://doi.org/10.1145/3173574.3173629>
- [63] Craig Oldham. 2017. *Oh sh*t what now?: honest advice for new graphic designers*. Laurence King Publishing Ltd, London. OCLC: 1029212644.
- [64] O*NET OnLine. 2020. Summary Report for: 27-1024.00 - Graphic Designers. <https://www.onetonline.org/link/summary/27-1024.00>
- [65] D. Royce Sadler. 2009. Transforming Holistic Assessment and Grading into a Vehicle for Complex Learning. In *Assessment, Learning and Judgement in Higher Education*, Gordon Joughin (Ed.). Springer Netherlands, Dordrecht, 1–19. https://doi.org/10.1007/978-1-4020-8905-3_4
- [66] Niloufar Salehi and Michael S. Bernstein. 2018. Ink: Increasing Worker Agency to Reduce Friction in Hiring Crowd Workers. *ACM Transactions on Computer-Human Interaction* 25, 2 (April 2018), 10:1–10:17. <https://doi.org/10.1145/3177882>
- [67] Donald Schön. 1983. *The reflective practitioner*. Temple Smith, London. http://scholar.google.com/scholar?q=related:im1nTiraZZA:scholar.google.com/&hl=en&num=20&as_sdt=0,5
- [68] Anselm Strauss and Juliet M. Corbin. 1997. *Grounded Theory in Practice*. SAGE. Google-Books-ID: TtRmolAapBYC.
- [69] Will Sutherland, Mohammad Hossein Jarrahi, Michael Dunn, and Sarah Beth Nelson. 2019. Work Precarity and Gig Literacies in Online Freelancing. *Work, Employment and Society* (Nov. 2019), 0950017019886511. <https://doi.org/10.1177/0950017019886511> Publisher: SAGE Publications Ltd.
- [70] Maryam Tohidi, William Buxton, Ronald Baecker, and Abigail Sellen. 2006. Getting the right design and the design right: Testing many is better one. *Proceedings of the 24th Annual ACM Conference on Human Factors in Computing Systems* (April 2006), 1243–1252. <https://doi.org/10.1145/1124772.1124960>
- [71] Maarten W van Someren, Yvonne F Barnard, and Jacobijn A C Sandberg. 1994. *The Think Aloud Method*. Academic Press, London. http://books.google.com/books?id=lnp9AAAAMAAJ&q=intitle:THE+THINK+ALoud+METHOD+A+practical+guide+to+modelling+cognitive+processes&dq=intitle:THE+THINK+ALoud+METHOD+A+practical+guide+to+modelling+cognitive+processes&hl=&cd=1&source=gbs_api
- [72] Anbang Xu, Shih-Wen Huang, and Brian Bailey. 2014. 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 (CSCW '14)*. ACM, New York, NY, USA, 1433–1444. <https://doi.org/10.1145/2531602.2531604> event-place: Baltimore, Maryland, USA.
- [73] Anbang Xu, Huaming Rao, Steven P Dow, and Brian P Bailey. 2015. 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 Press, New York, New York, USA, 1637–1648. <https://doi.org/10.1145/2675133.2675140>
- [74] Yu-Chun Grace Yen, Steven Dow, Elizabeth M Gerber, and Brian P Bailey. 2016. 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 Press, 773–784. http://www.google.com/intl/en_us/mail/help/about.html

- [75] Haojiang Ying, Edwin Burns, Xinyi Lin, and Hong Xu. 2019. Ensemble statistics shape face adaptation and the cheerleader effect. *Journal of Experimental Psychology: General* 148, 3 (2019), 421–436. <https://doi.org/10.1037/xge0000564> Place: US Publisher: American Psychological Association.
- [76] Alvin Yuan, Kurt Luther, Markus Krause, Sophie Vennix, Steven P Dow, and Björn Hartmann. 2016. Almost an expert: The effects of rubrics and expertise on the perceived value of crowdsourced design critique. In *the 19th ACM Conference*. https://www.kurtluther.com/pdf/CrowdCrit_CSCW_2016_camera.pdf
- [77] Haoqi Zhang, Edith Law, Rob Miller, Krzysztof Gajos, David Parkes, and Eric Horvitz. 2012. Human Computation Tasks with Global Constraints. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 217–226. <https://doi.org/10.1145/2207676.2207708> event-place: Austin, Texas, USA.
- [78] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. Association for Computing Machinery, San Jose, California, USA, 493–502. <https://doi.org/10.1145/1240624.1240704>

Appendices

A FORMATIVE DESIGN RESEARCH PROCESS

A.1 Understanding Best Practices for Portfolio Development

To understand best practices for developing portfolios, we observed several portfolio design courses and reviewed resources for assembling a professional online portfolio. One of the authors observed three 80-minute classes in two portfolio design courses taught by the same instructor at a large Midwestern university in the US. The classes were targeted at engineering majors, but included students from diverse fields such as journalism and computer science. Students were guided through the process of creating online portfolio websites through weekly exercises, such as generating and curating a list of work experiences, critiquing existing portfolio websites, learning how to design a website, and presenting one’s website to others. We also reviewed class materials (e.g., syllabi, lectures) and conducted semi-structured interviews with 1) the course instructor, who had spent more than 20 years working as a graphic designer, and 2) a faculty mentor who had directed a design program for more than four years (both male).

To better understand best practices for developing *graphic design* portfolios, we reviewed several popular resources, including two books on assembling professional design portfolios [56, 63] and several online guides we found using search terms (e.g., “design,” “create,” “portfolio,” “tutorial”) and recommendations from a product manager with 14 years of experience working with graphic designers [15, 47, 47]. We used inductive coding methods [68] to analyze our notes and understand common advice and challenges freelancers face when developing an online portfolio. We focused on graphic design portfolios in this study because of the ease of sharing graphic design projects in a digital online format and the popularity of graphic design projects in creative online communities [47].

The portfolio design guides we reviewed suggested that the content of one’s portfolio (i.e., personal description, projects displayed) should convey one’s unique skills (e.g., illustration, logo design) and aesthetic style (e.g., modern, minimalist) [46, 47, 56]. Creators should also display only the most recent projects in their portfolio that represent the kinds of work they would like to do [46, 56, 63]. For example, if a freelancer is interested in designing more illustrations for children’s books, they should choose to include their newer illustration projects over their older typography projects in their portfolio. Hence, the portfolio design guides we reviewed encouraged freelancers to regularly curate and consider their current skills, styles, and career goals when assembling the content of their portfolio, as opposed to a building a comprehensive history of work.

Our observations of the portfolio design courses corroborated findings from the design guides. Here, we found structured self-reflection and holistic portfolio feedback can help portfolio creators

stay focused on refining the content of their portfolios. The course instructor explained that students were often tempted to “get started in designing [their] website right away,” rather than develop the actual content of their portfolio. Hence, the instructor used class exercises to encourage students to first take inventory of their skills, reflect on their career goals, and research their target audience, before helping them learn how to manage the visual design of their websites. Receiving holistic portfolio feedback can further help creators refine the content of their portfolios. For example, during an in-class feedback session, the faculty mentor we spoke with noticed a common theme in a journalism’s student’s portfolio: experience working with teams of engineers. The faculty mentor then encouraged the student to showcase this unique skill and highlight projects that represent this collaborative work. The course instructor explained to us that he invited faculty mentors to provide feedback during class so that students would produce higher quality portfolios and because it would be difficult for students to regularly receive feedback outside of school. While learning environments are set up for regular feedback, once individuals leave these environments, it can be difficult to collect the feedback they need.

The image shows a digital form for a self-reflection tool. It is divided into several sections:

- Question 1:** "What 3 skills are necessary to demonstrate to attract the type of clients you want to work for?"
- Skills Boxes:** Three boxes with skill descriptions and a 5-point rating scale (1-5):
 - Marketing strategy: 1 (orange), 3 (green)
 - Asking the right questions to get to your right brand: 3 (green), 4 (blue)
 - Fine arts skills, hand drawn to digital skills: 2 (yellow), 5 (grey)
- Question 2:** "What 3 words describe how you want to come across in your portfolio that would make you unique from another designer?"
- Word Boxes:** Three boxes with words and a 5-point rating scale (1-5):
 - Thoughtful: 1 (orange), 2 (yellow), 3 (green), 4 (blue), 5 (grey)
 - Driven: 1 (orange), 2 (yellow), 4 (blue)
 - Passionate: 1 (orange), 2 (yellow), 3 (green), 4 (blue), 5 (grey)
- Projects List:** A list of five projects, each with a colored label and a name:
 - Project 1 (orange): Journe Branding Project
 - Project 2 (yellow): Angers Illustration
 - Project 3 (green): Sweet Iron Waffles Marketing Materials
 - Project 4 (blue): Aligned company widget
 - Project 5 (grey): Permanent Zine
- Bottom Right:** "Your top 3 projects" section showing three selected projects: Project 1 (orange), Project 2 (yellow), and Project 3 (green).

Fig. 6. Formative Prototype 1: Self-Reflection Tool. In this formative digital prototype, participants were asked to choose up to five projects in their portfolio to reflect on (right). Then, they were asked to type three skills they would need to demonstrate through their portfolio and three words that described their ideal portfolio (left). Participants reflected on which of their five projects portrayed these skills and words by dragging them to the appropriate boxes on the left. After doing this exercise, participants chose the top three projects in their portfolio that portrayed these skills and words the best by dragging a copy of each project to the bottom right. All of these activities took place on the same screen in the digital prototype.

A.2 Designing and Evaluating Formative Prototypes of Portfolio Feedback Systems

Based on this exploratory research, we identified two main requirements for a portfolio feedback tool. Such a tool should 1) allow designers to evaluate whether a portfolio furthers their career goals and 2) help designers reflect on and pinpoint their professional skills and unique style. With this in mind, we designed and evaluated two formative prototypes of potential portfolio feedback tools.

A.2.1 Formative Prototype 1: Self-Reflection Tool. Based on one of the activities we observed in the portfolio design courses, our first design attempt was to create a self-reflection tool to help freelancers take inventory of their design skills. We decided to first explore the benefits of a self-reflection tool given the challenges of seeking rapid feedback from others outside of a classroom setting. This formative prototype was created using Google Slides, and acted as a mockup of an interface for a web application that prompts freelancers to reflect on how their projects convey different skills and qualities (e.g., aesthetic styles) in their portfolio (Figure 6). We evaluated our first formative prototype with eight freelancers (four female, four male) who were recruited from

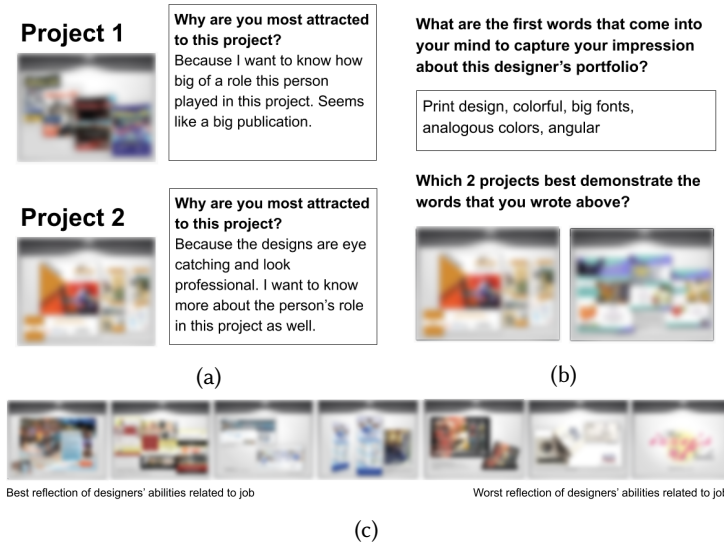


Fig. 7. Formative Prototype 2: Crowd Feedback Tool - Feedback Collection Interface. In the second formative prototype, we collected feedback on a designer's portfolio with three feedback collection tasks: (a) feedback providers were asked to select two projects in the portfolio they were most attracted to and explain their choices, (b) feedback providers were asked to write the first words that came to mind when viewing the portfolio and select two projects that best capture those words, and (c) feedback providers were shown a sample graphic design job description and rank projects based on how well they reflect the designers' abilities related to the job. Project images have been blurred to maintain the confidentiality of the designer.

a popular online freelancing platform [8] and a local tech meetup in a large metropolitan city. A member of the research team showed participants, either remotely or in person, a blank slide template and watched them edit the slide as they were guided through the formative prototype. Participants were asked to list up to five projects in their portfolio, choose three skills and three words to describe their professional work, then assign projects to the different skills and words they chose. Participants were asked to think aloud as they used the prototype [71]. Five study sessions were conducted remotely over video call, while three were conducted in person.

Although participants found this formative prototype helpful for curating their portfolio, participants explained that, to remain competitive, they still need to understand whether their online portfolio could capture the limited attention of a potential client. For instance, one participant mentioned the challenge of understanding which projects were the most visually attractive to clients. This participant used Google Analytics to understand which projects were viewed the most on their portfolio website, as well as demographic details of visitors to their website to better understand who their potential clients might be. However, Google Analytics could not help this participant understand why some people were attracted to parts of their portfolio, nor who was the perceived target audience for the portfolio (e.g., children, if the designer primarily made projects for children). As a result of the findings from this first design attempt, we added a third design requirement for a portfolio feedback tool: it should help designers understand their portfolio's appeal to potential clients and the perceived target audience through rapid feedback from new, diverse, and readily-available feedback providers.

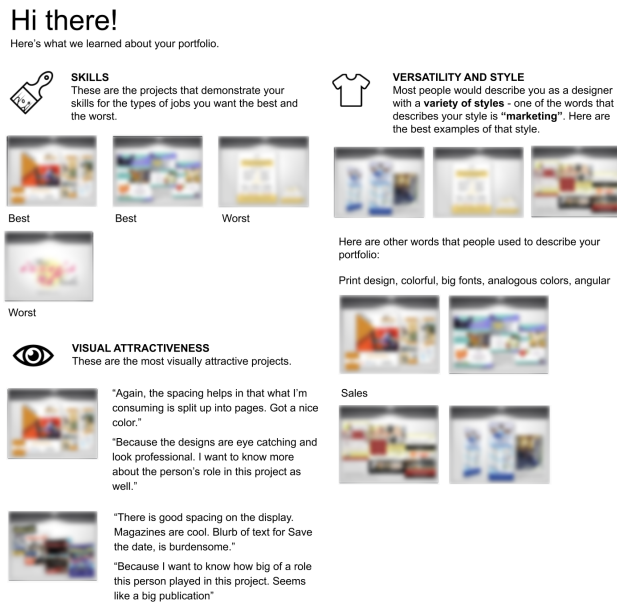


Fig. 8. Formative Prototype 2: Crowd Feedback Tool - Feedback Presentation Interface. We displayed portfolio feedback on the second prototype to participants. On the left, we displayed the projects that best demonstrated the designers' skills and were the most visually attractive. On the right, we displayed the projects that best conveyed style words feedback providers associated with the portfolio. For the purposes of the formative study, we showed all participants feedback on the same designer's portfolio. Project thumbnails have been blurred to maintain the confidentiality of the designer.

A.2.2 Formative Prototype 2: Crowd Feedback Tool. In our second design attempt, we used wizard-of-oz methods [40] to understand how helpful freelancers would find a crowd feedback tool that could simulate the experience of a potential client and provide rapid feedback. To do this, we created a second formative prototype consisting of two interfaces: a feedback collection interface and a feedback presentation interface. These interfaces were created using Google Slides, and were used as a mockup for the interface of a potential web application.

We first tested our feedback collection interface (Figure 7). We selected a graphic designer's portfolio from Behance [4], a popular online community for graphic designers, to use as the example portfolio depicted in the view. We selected this designer's portfolio because it was representative of other portfolios we had seen on Behance. The feedback collection interface consisted of a web survey with three feedback collection tasks that asked about: 1) the visual attractiveness of different projects in the portfolio, 2) the perceived styles or aesthetics in the portfolio, and 3) the projects that best portrayed a designer's graphic design skills, following the design requirements we identified from our formative research described earlier. To simulate the experience of a client with limited attention, we displayed only thumbnail previews of each project in the portfolio to feedback providers. Feedback was provided by five people, including two members of our research team and three university students (three female, two male) who served as proxies for a crowd of novice feedback providers. We showed participants the Google Slides mockup, asking them to drag project thumbnails on the screen to the appropriate boxes and type their answers to each survey

question. We asked participants to think aloud [71] as they interacted with the formative prototype in person with our research team.

Second, to test our feedback presentation view (Figure 8), we recruited three freelance graphic designers (two female, one male) with online portfolios in the US via an online freelancing platform and scheduled remote video conference calls with these participants. The presentation view displayed the projects from the example portfolio and the feedback we collected from the participants above. During each remote video call session, we sent participants a link to the formative digital prototype in Google Slides, where they saw all of the feedback on a single page. Participants were asked to think aloud as they imagined using the prototype for their own portfolio. We also asked participants whether this type of feedback would be useful to them for improving their portfolio.

Overall, participants found both the content of the collected feedback and the feedback presentation interface helpful. Participants said that the perceived styles feedback would help them come up with new ways to describe and tag their work. This was important to freelancers as it could help them more accurately attract new clients to their online portfolio looking for specific styles. Similarly, participants said that the visual attractiveness feedback would help them quickly identify which project thumbnails were the most attractive so that they could propose changes to those that were less appealing. However, participants did not trust that an online crowd could assess *how well* individual projects demonstrated graphic design skills. Instead, participants suggested that crowd feedback regarding perceived target audience and the primary design skill shown in the portfolio would be more trustworthy than having a novice crowd judge the quality of their work. These findings informed the design of the CrowdFolio system, which is described in the next section.

In addition to supporting designers' career goals and collecting feedback rapidly, we identified three aspects of a portfolio that would be most appropriate for crowd feedback:

- Visual Attractiveness (which projects capture the most attention from potential clients?),
- Styles (how would other people describe the portfolio?),
- and Target Audience (what type of audience are the creative projects designed for and what primary design skills does the portfolio convey?).

While these early findings were promising, we still needed to understand the practical challenges to collecting crowd feedback on online portfolios and the effects of decomposed (i.e., reviewing one project at a time) and holistic approaches (i.e., reviewing multiple projects at once, as in the portfolio design courses) to collect feedback. In our final system, CrowdFolio, we build on the second formative prototype by 1) including an additional section for collecting feedback on a freelancer's perceived target audience and single primary design skill (e.g., illustration, digital advertising) and 2) employing actual crowdworkers to provide feedback.

Received June 2020; revised October 2020; accepted December 2020