Bridging the Utilitarian-Hedonic Divide
in Crowdsourcing Applications

Mark Melenhorst[1][2], Jasminko Novak[1][3], Isabel Micheel[1], Martha Larson[2], Martin Boeckle[3]

[1] European Institute for Participatory Media
[2] Multimedia Computing, Department of Intelligent Systems, Delft University of Technology
[3] University of Applied Sciences Stralsund

m.s.melenhorst@tudelft.nl, j.novak@eipcm.org, i.micheel@eipcm.org,
m.a.larson@tudelft.nl, martin.boeckle@fh-stralsund.de

ABSTRACT
This paper introduces a novel perspective on the gamification of crowdsourcing tasks by conceptualizing it as the introduction of hedonic quality into the solution of utilitarian tasks and into the design of corresponding systems. We demonstrate how such a conceptualization can enable crowdsourcing applications to involve new kinds of crowds in everyday contexts that cannot be reached with existing models. We illustrate its application with the design of TrendRack, a gamified crowdsourcing application in the domain of fashion. We then discuss the results from a first evaluation, suggesting successful engagement of fashion customers in everyday contexts.

Categories and Subject Descriptors
H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval

General Terms
Design, Experimentation, Human factors

Keywords
Crowdsourcing, motivation, hedonic quality, GWAPs, user study, mobile application development

1. INTRODUCTION
The people that carry out crowdsourcing tasks are known as "crowdworkers". This designation often leads the designers of crowdsourcing platforms to an "all work and no play" attitude. The idea that crowdsourcing can also be fun is relegated to the domains that explicitly involve games. In this paper, we present a mobile crowdsourcing application and the results of an evaluation study that allow us to demonstrate the importance and value of fun (hedonic value). We argue that the hedonic value of

crowdsourcing tasks is important for two different reasons. First, working conditions: in the crowdsourcing community considerable debate has emerged on labor conditions and the extent to which crowd workers should be treated as regular employees (e.g. [2], [12]). If this were the case, then in the design of crowdsourcing tasks attention must be paid to employee satisfaction and, consequently, to the enjoyment crowd workers gain from performing a task. Second, following the saying "A happy worker is a productive worker" we point out that the labor conditions are not only relevant from an ethical perspective, but also from the perspective of data quality. An increase in user satisfaction is expected to result in higher levels of engagement with the crowdsourcing task and as a result an increase in data quality and a reduction of cheating behavior.

While the importance of the hedonic value of crowdsourcing tasks has not been completely overlooked, it is often the case that crowdsourcing tasks are presented as either hedonic (i.e., games with a purpose) or utilitarian (e.g., goal- and task-driven). In contrast, in this paper we present and evaluate a crowdsourcing system that bridges this "hedonic-utilitarian divide". For this purpose we introduce a mobile app that combines gamified elements with a user interface that is designed for both appeal and productivity.

After classifying crowdsourcing systems in terms of worker incentives, we explain the distinction between hedonic and utilitarian crowdsourcing tasks and elaborate our argument for a third class of crowdsourcing tasks that bridges the hedonic-utilitarian divide.

In conceptualizing crowdsourcing tasks in terms of the value they deliver, we first define crowdsourcing. In this paper, we consider crowdsourcing as an open call for contributions from members of a voluntary crowd to solve a given problem or carry out tasks requiring simple human intelligence judgments, often in exchange for micro-payments, social recognition or entertainment value [14]. Two main classes of crowdsourcing systems include: 1) task-oriented crowdsourcing platforms that distribute small work assignments (e.g., Amazon Mechanical Turk) and 2) gamified crowdsourcing systems. Crowdsourcing platforms incentivize users by offering small monetary rewards for microtasks. In contrast, gamified crowdsourcing systems take advantage of game-like motivational mechanisms (e.g., fun, joy of use, competition). Gamification refers to the use of game elements and game mechanics in non-game contexts [1]. Two main approaches include 1) designing a game that seamlessly integrates pragmatic tasks (e.g., learning, information acquisition, data validation) into

game actions, and 2) adding different kinds of game-like reward elements (e.g., points, badges, leaderboards) to an otherwise non-playful application context, in order to stimulate engagement. In gamified crowdsourcing, the first approach is reflected in games-with-a-purpose (GWAPs) as specialized game designs where crowd-tasks are solved as a by-product of play ([1]). Examples include the classical ESP game [1] for labeling images, and Sketchness, in which players identify clothing items in fashion images from contours drawn by their game partners [7]. Alternatively, in the second class of approaches, microtasks (work assignments) are typically enriched with game-like rewards (such as points, badges and leaderboards).

While the development of GWAPs has been the subject of extensive research and with clear conceptualizations of different strategies (e.g. goals, rules, tasks), the latter type of gamified crowdsourcing has been much less investigated. Hence, a clear conceptualization that could guide the design of effective gamified crowdsourcing systems is still missing. Recent contributions for instance have questioned the applicability of straightforward "additive" gamification, suggesting that gamifying a work assignment requires a holistic approach, involving a fundamental redesign of the crowd-task, including its conceptual metaphor and the way it is presented to the users [12].

In this paper, we propose that the challenge of gamifying crowdsourcing tasks can be better understood (and designed for) by conceptualizing it from a utilitarian vs. hedonic systems point of view. We argue that the gamification of crowdsourcing tasks can be perceived as the introduction of hedonic quality into the solution of utilitarian tasks and into the design of corresponding systems. We illustrate how this conceptualization has been put in practice in the design of a concrete gamified crowdsourcing application in the domain of fashion and discuss results from a first evaluation in a real-world context. Finally, this real-world example demonstrates how the introduction of hedonic elements to otherwise utilitarian tasks can be used to involve new kinds of crowds in everyday contexts that cannot be reached with existing models.

Although crowdsourcing has been mentioned previously in the literature as involving both a utilitarian and hedonic component, to our knowledge, this is the first work that has specifically targeted conceptualizing and evaluating a system on the basis of its combined hedonic and utilitarian quality.

2. THE UTILITARIAN-HEDONIC DIVIDE IN CROWDSOURCING

The contrast between crowdsourcing systems that have joy of use at the center of attention (GWAPs), and systems that focus on economically viable, efficient and high-quality output, can be explained by the distinction between utilitarian and hedonic systems. Utilitarian systems address tasks and activities where user motivation in using the system is driven by the expectation of an external reward or benefit [10]. In contrast, hedonic systems serve activities in which users are intrinsically motivated by benefits stemming from the interaction with the system as such [10]. Hedonic systems typically provide stimulation by their challenging or novel character, or identification by communicating important personal values to others [9].

Traditional crowdsourcing systems with microtasks can be perceived as utilitarian systems, while gamified crowdsourcing applications in form of GWAPs correspond to the class of hedonic systems. Systems that only serve utilitarian objectives have limited usefulness for tasks where low monetary incentives do not suffice to motivate workers to produce high-quality data or when higher payments are not feasible. In fact, how to capture the attention of workers to support them to deliver high-quality data is an important topic in crowdsourcing research [8][14][18][20]. Computational solutions to this have been proposed and tested (e.g. [11]), often without addressing user satisfaction. In contrast, GWAPs, as gamified crowdsourcing applications, are designed to deliver hedonic quality to their users. Research in this area primarily focuses on the user's sense of enjoyment, while the (camouflaged) utilitarian objectives (e.g. the task outcomes) impose requirements on the game design, but are not at the center of attention. In terms of Hassenzahl [9], GWAPs offer hedonic stimulation by means of personal rewards and competitive game-like elements such as points, badges and leaderboards. Note that the disguised utilitarian objectives in GWAPs is topic of debate under the label of the 'puppet master problem': an invisible force that exercises control over players' behavior without a priori revealing the game plot or, in this case, the utilitarian purposes of the game [15].

Even though in the literature the distinction between hedonic and utilitarian systems is frequently presented as a dichotomy, some work has pointed out that a class of systems could exist that inherently combine the two aspects [17]. We argue that such a more balanced view on the distinction between utilitarian and hedonic systems can also be applied to gamified crowdsourcing systems. In this case gamification elements are used to enrich an otherwise utilitarian crowd-task (work assignment). However, systems of this new class of gamified crowdsourcing tasks are not pure hedonic systems as they still offer work assignments that are not a source of enjoyment by nature. Similarly, when solving the task does not yield a system-external (monetary) reward, such systems cannot be seen as pure utilitarian systems either. Thus, to motivate users, such gamified crowdsourcing systems (non-GWAPs) need to deliver both utilitarian and hedonic value to the users (crowd workers). As a result, they should extend the notion of gamification elements to include design elements beyond traditional game mechanics. Rather, a number of issues from user interface design known to relate to both hedonic (e.g. user experience, satisfaction, appeal) and pragmatic aspects of system use (e.g. ease of use) also need to be considered. This supports the proposition that the gamification of crowdsourcing tasks requires a holistic, integrative approach that encompasses a fundamental redesign of the crowd-task, including its conceptual metaphor and its presentation to the users [12]. The resulting system can be expected to not only offer stronger motivational affordances, but also a positive impact on task performance.

Two important underexplored areas of research emerge from this perspective. First, few systematic studies have been carried out on the user interface design of crowdsourcing systems in general and on the impact of specific design elements in particular [19]. Previous research in HCI has also shown that hedonic aspects can play an important role for the user acceptance of a system ([9], [17]), which can also influence user effectiveness in task performance [17]. Secondly, existing evaluations of gamified crowdsourcing tasks tend to focus on task performance, while little attention is paid to user satisfaction and impact on motivation (e.g., [5], [6]). Similarly, few insights exist on the impact of specific gamification elements on the extrinsic and

intrinsic motivation of crowd workers [20]. In particular, little is known about their relation to hedonic quality stimulation.

The next sections present the first results of the design and evaluation of a gamified crowdsourcing app. The application provides a conceptual metaphor that relates the task to users' personal context and employs hedonic aspects and principles of gamification to increase engagement, while allowing for a quick an efficient task performance. The proposed application design and first evaluation results give an example of how this class of systems can be developed and how they could be used to engage novel kinds of user groups with a type of crowdsourcing tasks that require the involvement of particular target groups as crowd-workers (e.g., customers of a specific type of products) in real-world contexts.

3. DESIGN OF A HYBRID UTILITARIAN-HEDONIC CROWDSOURCING APP

The hybrid utilitarian-hedonic application that we developed, called 'TrendRack', engages a fashion-aware crowd to judge the similarity between different garments. The similarity judgments were used within a trend analysis system for fashion SMEs to compare an SME's image of a garment to a stream of social images crawled from e.g. Twitter. They were particularly interested in similarity between their fashion image and social fashion images with respect to human-defined criteria (e.g., casualness, trendiness of garments) important to fashion consumers. Such complex judgments cannot be performed purely automatically, while judgments by a general crowd (as on platforms like AMT) are insufficient: to be relevant they require awareness of specific notions of fashion and need to be performed by existing or potential customers of the fashion companies.

As opposed to AMT workers, a fashion-aware customer crowd cannot be motivated by means of traditional micropayments. First, building a community of reliable, readily available crowd workers with the required expertise is a time-consuming and expensive task. Second, the aforementioned trend analysis system is targeted at SMEs who cannot afford to continuously pay crowd workers for the microtask.

While PCs and laptops primarily afford usage for pragmatic goals, smartphones are considered to be appropriate devices for both pragmatic and – possibly even more important – hedonic purposes. For that reason, a mobile crowdsourcing app was developed to reach the target crowd quickly, intuitively and when available (leisure time).

The user interface of the app is presented in Figure 1. The crowd task itself is of a mainly utilitarian nature. A reference image posted by the fashion company (a new garment design) is shown to users. They should then identify up to five images of clothing they consider similar to the reference image. The five images can be selected from a stream of social fashion images, which is the result of an automatic image similarity search. The reference image is associated to a clothing category (e.g., skirts) to avoid confusion in case images display multiple garments. Users are requested to judge similarity based on a criterion that is pre-selected by the fashion company (e.g., cut). The similarity criteria were derived from a user study ran on Amazon Mechanical Turk, which was presented at CrowdMM last year ([16]).

To make the task easier to understand for the target crowd, an interaction metaphor was chosen that resembles the natural actions users undertake with respect to clothes. Users select similar images from the stream by dragging and dropping them onto hangers on a virtual clothing rack as if sorting real clothing items. The stream can be played automatically, paused,, or navigated with a more intuitive swipe left and right gesture.

To increase the level of engagement of workers with the task, hedonic elements are included to gamify the user experience. For each selection, users are awarded points and stimulated with audio-visual feedback both during and after each task: A chime-like sound is played for each selection and upon task completion, where it is combined with visual "fireworks" (see Fig. 1. center). With the collected points, users earn fashion-related badges associated with a status (from "Trend Observers" to "Fashion Goddesses", Fig. 1. right). To stimulate competition, crowd users can compare themselves with others on a leaderboard.
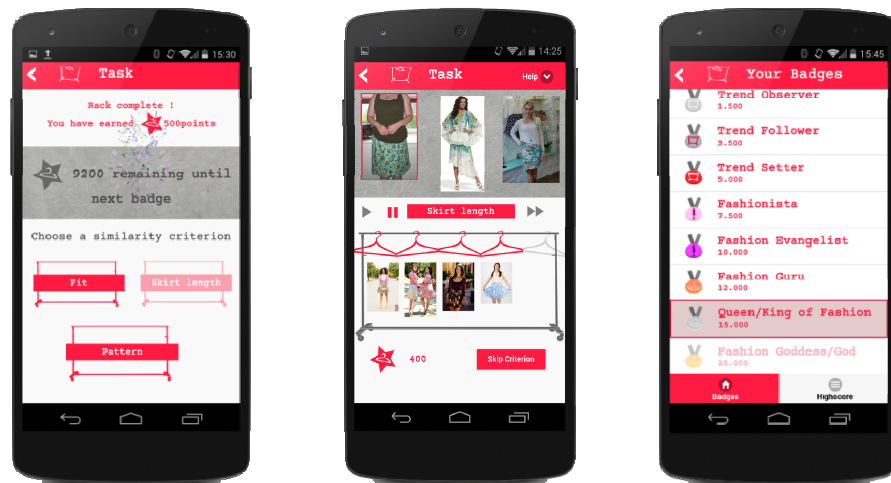

Figure 1. TrendRack hybrid utilitarian-hedonic crowdsourcing app. Left: Crowd-task based on human-centred similarity criterion. Center: Audio-visual feedback for task completion. Right: Social reputation badges leaderboard.

To summarize, the TrendRack app offers hedonic value by means of the interaction metaphors, the audiovisual system feedback, and the addition of gamified elements. The intuitive interaction simultaneously contributes to the utilitarian value in the sense that they allow for an efficient task execution. In other words, user interface design of crowdsourcing tasks can contribute to both the hedonic value and the utilitarian value of crowdsourcing tasks.

4. EVALUATION DESIGN

A field experiment was designed to evaluate the perceived hedonic quality and pragmatic quality users attribute to the gamified crowdsourcing tasks in the mobile app. The evaluation was carried out with two fashion-aware crowd groups in two settings: with anonymous participants in a (mobile) online experiment and with user study participants in a workshop setting. Both groups consisted of crowd members of the fashion crowdsourcing start-up "13 Dresses", which currently counts around 300 members. To avoid including only those users that are already active as fashion crowd members, additional users for both experiments were recruited via social media channels, based on the prerequisite that they should have a more than average awareness of or interest in fashion and could be seen as potential consumer crowd members by fashion SMEs. An additional incentive for participants was provided by offering three shopping vouchers, which were awarded to randomly selected participants. To avoid influencing the incentive structure we explicitly stated that participants' app usage did not influence the chance on getting a voucher.

The *online crowd experiment* was conducted over 15 days. To gain insight into user behavior in real-world-like conditions, three tasks were posted daily, to which participants responded as they wished. Participants' activity was logged with respect to their response times and number of task responses. After the experiment, participants needed to fill in an online questionnaire as a prerequisite to win a fashion voucher.

The *offline user study* was set up to investigate the dynamics of use in a realistic mobile setting. The study was presented to the participants as an informal after-work test event. The main objective was to get qualitative feedback from participants, and to

observe them responding to tasks based on incoming notifications in two typical contexts of use: 1) while occupied with other activities on their smartphone (phone in front of them, as e.g. when commuting or passing time), 2) while occupied with offline activities, as they e.g. would at home or while waiting at the doctor's office (phone close-by but not in use). At the end, participants filled out a questionnaire, followed by a group discussion.

In both the online experiment and the user studthe questionnaires focused on hedonic quality, pragmatic quality, and appeal. Hedonic quality and pragmatic quality were measured using Hassenzahl's AttrakDiff2 questionnaire [9]. Their scale consists of three sub scales: hedonic quality-stimulation (HQ-S), hedonic quality-identification (HQ-I), and pragmatic quality (PQ). All sub scales are measured with seven seven-point semantic differentials, ranging for instance from typical to original (HQ-S). HQ-S was used to measure the hedonic stimulation of the app, PQ was used to assess the utilitarian dimension (e.g., perceived effectiveness, and efficiency of use). Hedonic identity (HQ-I) was not evaluated since the system does not provide means for the user to communicate one's personal identity to other people. The appeal of TrendRack's gamified elements was assessed with a five-point Likert scale, ranging from very unappealing (--) to very appealing (++), preceded by the following question: "How appealing did you find the following aspects of the application?"

5. EVALUATION RESULTS

5.1 Hedonic and utilitarian perception of the mobile crowd app

A total of 17 users participated in the evaluation of the TrendRack mobile app, 9 in the online crowd experiment and 8 in the user study. In each of the two settings 8 participants filled out the questionnaire. In the user study all 8 participants were female, in the online experiment 5 were female and 3 male.

In the *online crowd experiment* a total of 35 tasks were submitted, each containing three fashion similarity criteria to solve (e.g. skirt length, pattern, cut, etc.) On average each participant solved 10 tasks and 29 similarity criteria. On average, only one criterion
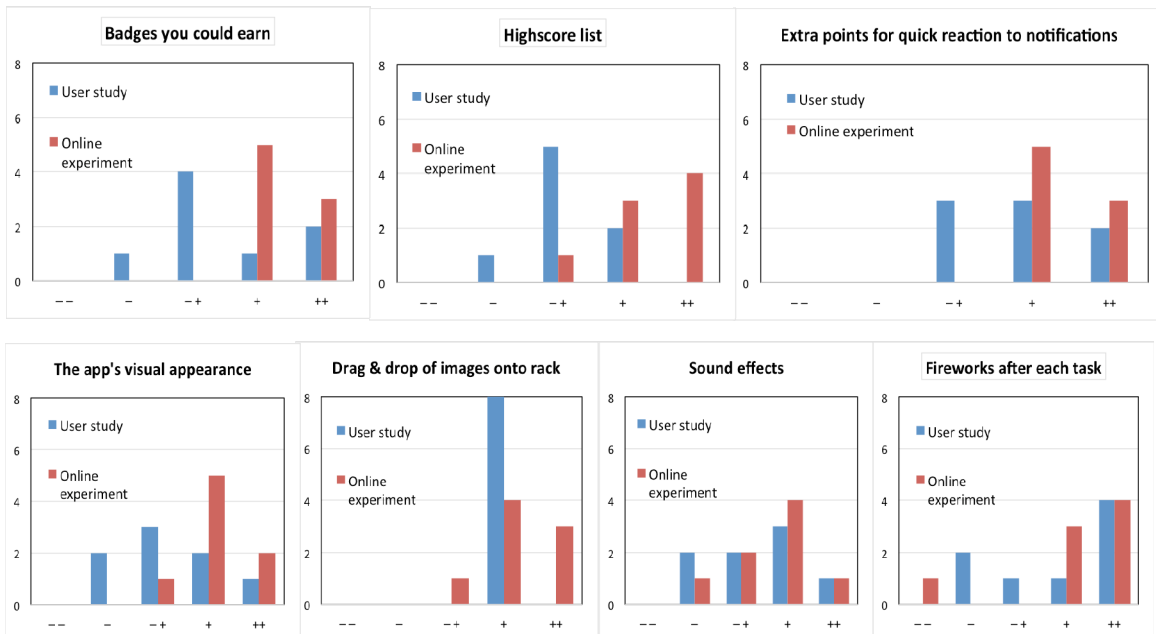

Figure 2. Appeal of gamification elements

was skipped by each user over the whole duration of the experiment, suggesting a high level of engagement with the tasks. In the open questions in the questionnaire, participants stated that they solved tasks mainly as a way to pass time while on the road and while they had their smartphone at hand. One participant stated that they solved tasks mainly while being on the balcony for a smoke and doing different things on their smartphone like chatting or browsing, another user said to have solved tasks mainly while watching TV in the evenings. This suggests that the developed crowdsourcing model succeeded in engaging the participants in their natural environment and context of use.

During the *offline user study*, 9 tasks (with a total of 27 similarity criteria) were available for participants to solve, of which each participant in this group solved 4 tasks and 11 similarity criteria on average. Here, the setting was to simulate different (assumed) situations in which users might be susceptible to registering and solving tasks (while sitting on a sofa browsing print magazines, browsing through an online shop on their smartphone or performing online searches with their smartphone). The results suggest a notably lower engagement in this setting than in the online experiment, which is attributable to the lab setting and the correspondingly tendency towards lower usage of the competition elements (e.g., the scoring and leaderboards; Figure 2).

HQ-S and PQ inform us about the value the TrendRack app offers to its users and about the balance between the hedonic value and the utilitarian value. We computed the hedonic quality-stimulation (HQ-S) and pragmatic quality (PQ) scale scores for both the offline user study and the online crowd experiment. The averages scores for both HQ-S and PQ scales are shown in Table 1, divided by the two study settings.

Table 1 Average hedonic and pragmatic quality scores divided by setting

Hedonic quality sub scale	Offline user study	Online crowd experiment
HQ-S	4.2 (.7)	4.9 (.6)
PQ	5.1 (.7)	5.5 (.4)

Note. Standard deviations in parentheses

As can be seen from Table 1, the HQ-S and PQ quality scores are moderately high in both settings, which according to [9] classify the app as 'balanced'. Due to the small sample size, non-parametric tests need to be used to assess the differences between the settings of the study and between the HQ and PQ scores. The Mann-Whitney U test was used to test for differences between the settings. For HQ-S, no significant differences were found between the offline user study and the online crowd experiment (U=11.0; p=.08; r=.45). The same applied to the pragmatic quality scores (U=21.5; p=.46; r=.19). A Wilcoxon signed ranks test revealed that PQ was significantly higher than HQ-S (Z=-3.08; p<.01; r=.85). These results suggest that across settings the application yielded hedonic quality without compromising the utilitarian purpose for which the crowdsourcing application was developed.

In both settings, participants were asked to indicate how appealing the gamified elements were to them. The results are displayed in Figure 2. The figure shows that elements not depending on social dynamics (e.g. audio-visual effects) were perceived rather positively in both the user study and the online experiment.

A Mann-Whitney U test revealed that participants in the online crowd experiment found the highscore list significantly more appealing than participants in the offline user study (U=7.50; p<.01; r=1.88). The distribution of values across scalepoints suggests a similar tendency for the other gamified elements. However, these differences were non-significant (U>15,50; p>.05).

In the online experiment, the number of times participants have viewed the highscore page (showing their own score with their status badge and a high-score list of all users) provides an additional indication of how users perceive the gamified elements. In the online experiment, participants viewed the score page on average 12 times during the 15 days the experiment lasted (with a maximum of 44 views and a minimum of 2 views). The almost daily view of the highscore page suggests that participants were interested in their rank on the high score list.

Participants in the offline user study also expressed explicitly during the discussion that they enjoyed playing with the app. Especially the idea to extend the gamification approach by offering fashion shopping vouchers for high-scoring players was popular among the participants, suggesting possibilities for further integration of hedonic and utilitarian aspects of the engagement model.

Apart from the gamification elements also the utilitarian value of the TrendRack app is important, both for the users and for the resulting image classifications. Participants were asked whether the interaction was clear and understandable on a five-point Likert scale, ranging from 'completely disagree' to 'completely agree'. Across the two settings, all but one participant answered either 'rather agree' or 'completely agree' on that statement. On a similar Likert scale, the clothing rack metaphor was perceived as rather or very appropriate by all but two participants across settings. As a confirmation of the positive perception, users stated that it was "quick, easy and fun", that they liked the "idea, design and practicability" and the "fashion topic".

These results suggest that the conceived design combining utilitarian and hedonic aspects succeeded in motivating the participants to work on the tasks assigned, over a period of time, with both a high perception of hedonic stimulation and pragmatic quality.

6. CONCLUSIONS

Even though TrendRack was evaluated with a relatively small number of participants, the first results suggest a strong potential for a class of gamified crowdsourcing systems that combine utilitarian system objectives with hedonic quality for end-users. In particular, in real-life contexts, participants confirmed the engaging role of different elements of hedonic stimulation resulting in a high level of engagement with the otherwise utilitarian tasks. The results support the importance of an integrative, holistic approach to gamifying crowdsourcing tasks, encompassing a redesign of the conceptual task metaphor and the overall user experience in addition to common gamified incentives. A systematic comparison between hybrid hedonic-utilitarian task designs and the baseline (e.g. the utilitarian-only tasks often found on AMT) is necessary to draw more final conclusions about the potential of hybrid hedonic-utilitarian crowdsourcing tasks.

The development and evaluation of the app also holds promise for businesses to engage new crowds of potential customers with specific characteristics (e.g., fashion awareness) for the execution

of utilitarian tasks, without (exclusively) relying on monetary incentives. Future research should experimentally evaluate the contribution of different design elements to the utilitarian and hedonic value, including individual gamified elements, interaction metaphors, and visual designs, resulting in guidelines for application designers. More work is also needed to investigate the interplay between hedonic quality and pragmatic quality [8] in the design of hybrid utilitarian-hedonic crowdsourcing tasks.

ACKNOWLEDGMENTS

REFERENCES

[1] Ahn, L.v. and Dabbish, L., 2008. Designing games with a purpose. *Communications of the ACM*, *51,* 8, 58-67.

[2] Bergvall-Kåreborn, B. and Howcroft, D., 2014. Amazon Mechanical Turk and the commodification of labour. *New Technology, Work and Employment 29,* 3, 213-223.

[3] Deterding, S., Dixon, D., Khaled, R., and Nacke, L., 2011. From game design elements to gamefulness: defining "gamification". In *Proc. of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (Tampere, Finland, 2011), ACM, 9-15.

[4] Djelassi, S. and Decoopman, I., 2013. Customers' participation in product development through crowdsourcing: Issues and implications. *Industrial Marketing Management 42*, *5 (7)*, 683-692.

[5] Dumitrache, A., Aroyo, L., Welty, C., Sips, R.J., and Levas, A., 2013. "Dr. Detective": combining gamification techniques and crowdsourcing to create a gold standard in medical text. In *Proc. of the 1st International Workshop on Crowdsourcing the Semantic Web* (Sydney, Australia, 2013).

[6] Eickhoff, C., Harris, C.G., Vries, A.P.d., and Srinivasan, P., 2012. Quality through flow and immersion: gamifying crowdsourced relevance assessments. In *Proc. of the 35th international ACM SIGIR conference on Research and development in information retrieval* (Portland, Oregon, USA, 2012), ACM, 871-880.

[7] Galli, L., Fraternali, P., Martinenghi, D., Tagliasacchi, M., and Novak, J., 2012. A Draw-and-Guess Game to Segment Images. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Conference on Social Computing*, 914-917.

[8] Goodman, J.K., Cryder, C.E., and Cheema, A., 2013. Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples. *Journal of Behavioral Decision Making 26*, 3, 213-224.

[9] Hassenzahl, M., 2004. The Interplay of Beauty, Goodness, and Usability in Interactive Products. *Human-Computer Interaction 19*, 4, 319-349.

[10] Heijden, H., 2004. User acceptance of hedonic information systems. *MIS Quarterly. 28*, 4, 695-704.

[11] Hirth, M., Hoßfeld, T., and Tran-Gia, P., 2013. Analyzing costs and accuracy of validation mechanisms for crowdsourcing platforms. *Mathematical and Computer Modelling 57*, 11-12, 2918-2932.

[12] Irani, L. and Silberman, M., 2013. Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. In Proceedings of the CHI (Paris, France2013), ACM, 611-620.

[13] Jacobs, M.. Gamification: Moving from 'addition' to 'creation'. In *Proceedings of the ACM CHI 2013 Workshop on Designing Gamification: Creating Gameful and Playful Experiences*, 2013.

[14] Kazai, G., 2011. In Search of Quality in Crowdsourcing for Search Engine Evaluation. In *Advances in Information Retrieval*, P. Clough, C. Foley, C. Gurrin, G.F. Jones, W. Kraaij, H. Lee and V. Mudoch. Eds. Springer Berlin Heidelberg, 165-176.

[15] McGonigal, J. (2006). The Puppet Master Problem. Design for Real World, Mission Based Gaming. In. Second Person, P. Harrigan, N. Wardrip-Fruin. MIT Press.

[16] Melenhorst, M., Menendez-Blanco, M., and Larson, M., 2014. A Crowdsourcing Procedure for the Discovery of Non-Obvious Attributes of Social Images. In *Proceedings of the Proc. of the 2014 International ACM Workshop on Crowdsourcing for Multimedia* (Orlando, Florida, USA 2014), ACM, 45-48.

[17] Novak, J. and Schmidt, S., 2009. When Joy Matters: The Importance of Hedonic Stimulation in Collocated Collaboration with Large-Displays. In *Human-Computer Interaction – INTERACT 2009*, T. Gross, J. Gulliksen, P. Kotzé, L. Oestreicher, P. Palanque, R. Prates and M. Winckler (Eds). Springer Berlin Heidelberg, 618-629.

[18] Peer, E., Vosgerau, J., and Acquisti, A., 2014. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavioural Research Methods 46*, 4, 1023-1031.

[19] Rahmanian, B. and Davis, J.G., 2014. User interface design for crowdsourcing systems. In *Proc. of the International Working Conference on Advanced Visual Interfaces* (Como, Italy, 2014), 405-408.

[20] Redi, J. and Povoa, I., 2014. Crowdsourcing for Rating Image Aesthetic Appeal: Better a Paid or a Volunteer Crowd? In *Proceedings of the 2014 International ACM Workshop on Crowdsourcing for Multimedia* (Orlando, Florida, USA2014), ACM, 25-30.

[21] Richter, G., Raban, D., and Rafaeli, S., 2015. Studying Gamification: The Effect of Rewards and Incentives on Motivation. In *Gamification in Education and Business*, T. Reiners and L.C. Wood (Eds). Springer International Publishing, 21-46.

[22] Venkatesh, V., Morris, M., Davis, G., and Davis, F., 2003. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly 27*, 3, 425-478.