Behaviour & Information Technology

ISSN: 0144-929X (Print) 1362-3001 (Online) Journal homepage: https://www.tandfonline.com/loi/tbit20

A typology of human computation games: an analysis and a review of current games

Ei Pa Pa Pe-Than, Dion Hoe-Lian Goh & Chei Sian Lee

Behaviour & Information Technology, 2015
Vol. 34, No. 8, 809–824, http://dx.doi.org/10.1080/0144929X.2013.862304

A typology of human computation games: an analysis and a review of current games

Ei Pa Pa Pe-Than*, Dion Hoe-Lian Goh and Chei Sian Lee

Wee Kim Wee School of Communication & Information, Nanyang Technological University, 31 Nanyang Link, SCI Building, Singapore 637718

(*Received 16 January 2013; accepted 19 October 2013*)

Human computation games (HCGs) utilise human intelligence for the purpose of problem solving through games. Recently, HCGs have increased in popularity within various application areas. A proper system of describing such games is necessary in order to obtain a better understanding of the current body of work and identify new opportunities for future research directions. To date, however, there is little research conducted in understanding how such games could be classified in a systematic manner. Past classification schemes rely either on quality or entertainment aspects of these games. In order to have a deeper understanding of such games, this study argues that they should be classified via different dimensions. As a starting point, this paper reviews existing literature on HCGs and then develops a typology consisting of 12 dimensions and related strategies. This typology provides a holistic view of HCGs, and provides a common understanding of the domain among researchers and designers. The typology will help researchers to better understand the nature and consequences of HCGs, and designers better identify strengths and weaknesses of different strategies for each dimension, thus facilitating the creation of entertaining and effective HCGs.

Keywords: human computation games; crowdsourcing games; social computing; survey; typology

1. Introduction

In recent years, the popularity of online gaming has dramatically increased in terms of usage and amount of time spent because of its accessibility, variety, affordable prices, and engaging gameplay. According to a report by the Entertainment Software Association (ESA), as of 2011, 72% of American households play computer or video games, and its latest report (2012) stated that 15% of the most frequent game players were online gamers. At the same time, with the proliferation of Web 2.0 applications, users generate a rich amount of useful information as a by-product of everyday activities (Law and von Ahn 2011). Therefore, games have been utilised as a motivator to encourage content creation and user participation in human computation (Goh and Lee 2011). This has led to the rise of a new paradigm called the human computation game (HCG) that utilises games as a motivator to encourage user participation.

The HCG paradigm was introduced by von Ahn and Dabbish (2004) in which humans perform computational tasks as a by-product of enjoyable gameplay. In particular, such games provide fun or enjoyment as an incentive for participation. One of the earlier examples of HCGs is the ESP game (von Ahn and Dabbish 2004) in which two unrelated players are tasked to produce matching labels for given images, with the collected labels used as descriptors to improve the accuracy of image retrieval algorithms. The success of the ESP game has been overwhelming and research has shown that hundreds of thousands of people have played this game, generating millions labels for images within a short period of time (Law and von Ahn 2009). The innovativeness of this idea and the potential to leverage human computation has further driven the emergence of different genres of HCGs to solve various computational problems in areas including music annotation (e.g. Barrington *et al*. 2009), natural language processing (e.g. Chamberlain *et al*. 2009), and location-based content creation (e.g. Casey *et al*. 2007), protein folding (Cooper *et al*. 2010), and digital circuit design (Terry *et al*. 2009). These examples represent just the tip of the HCG iceberg and illustrate some of the types of problems that have been addressed with the help of games.

The proliferation of HCGs in various domains demonstrates the potential to advance computing power through harnessing problem solving abilities of game players. However, the diversity of HCGs makes it difficult for researchers to understand the current HCG landscape and identify commonalities and differences between them (Quinn and Bederson 2011). This lack of understanding could also make it harder for researchers and designers to describe such games, and to compare and classify them systematically, resulting in hindrances to efforts in uncovering new opportunities to leverage the power of HCGs. More importantly,

*Corresponding author. Email: ei1@ntu.edu.sg

a common language of describing HCGs is needed to maximise the knowledge among parties both within industry and academia (Björk *et al*. 2003).

Despite these benefits, to the best of our knowledge, there is little work that describes HCGs comprehensively. Previous attempts are mainly based on a single dimension such as game rules (e.g. von Ahn and Dabbish 2008) or game properties (e.g. Yuen *et al*. 2009). Describing HCGs in previous studies is not sufficient in providing a comprehensive understanding of these games because they contain a multitude of features related to both gaming and human computation systems (HCSs). In other words, given the entertainment-output generation duality of HCGs (Goh and Lee 2011), a typology should cover the important features from both aspects.

Thus, the overall objective of this paper is to develop a multi-dimensional typology that provides a holistic understanding of HCGs. Specifically, this study reviews existing literature concerning HCGs to identify their salient dimensions, and strategies used to implement these dimensions. This typology provides a well-defined approach for researchers and designers to understand and identify essential differences between HCGs, as well as classify them in a precise and analytical way. Once HCGs have been classified, the investigation of the nature and effectiveness of these games may be facilitated (Brandtzæg 2010). Equipped with a detailed understanding of HCGs, researchers could uncover directions for future investigation, and designers could identify effective strategies for HCG design.

The rest of the paper is organised as follows. In Section 2, we present a review of HCSs, followed by related work on the development of HCGs and the classification of these games. A methodology to extract related articles and HCG dimensions is described in Section 3. In Section 4, we classify current HCGs along different dimensions. We discuss our findings and implications of our work, as well as possible future directions for this field in Section 5. Finally, we provide concluding remarks in Section 6.

2. Related work

2.1. Human computation systems

HCSs harness human intelligence for solving computational problems which are beyond the power of computer programs, but which can be performed well by humans (Reeves and Sherwood 2010). Typically, HCSs utilise extrinsic motivation for people to participate and carry out the computation tasks to their best abilities (Krause and Smeddinck 2011). One example is reCAPTCHA in which users are required to generate a text string for a given distorted text so that they are granted access to their intended actions on the Web (Quinn and Bederson 2011). These outputs will later be used to improve the book digitisation process. Additionally, there are some online crowdsourcing markets such as Amazon Mechanical Turk (AMT) (http://www.mturk.com/) which consists of two types of participants: requesters and workers (also known as Turkers). Workers perform tasks uploaded by requesters and get paid for their successful completion. Examples of tasks on AMT include audio transcribing, image tagging, and finding information on websites.

In summary, current HCSs draw users' attention either by granting access to certain Web resources (e.g. reCAPTCHA) or offering monetary incentives to contributors (e.g. AMT). As users in paid HCSs have been found to be highly motivated by monetary rewards, it has become important for requestors to set attractive prices so that their tasks are more likely to be chosen by workers (Ipeirotis 2010). Although paying the right price is important, offering monetary incentives may not always be guaranteed to yield a higher performance (Huck *et al*. 2012). Extrinsic rewards may discourage people who are intrinsically motivated to perform the designated tasks from participation (Deci and Ryan 2000). It is also probable that users may devise ways to earn quick money, rather than contribute desirable outputs (Law and von Ahn 2011).

In this regard, HCGs may address some of these problems in that they facilitate people to implicitly carry out human computation tasks as a side effect of gameplay. In particular, performing computational tasks could be tedious, and also depend on volunteers' willingness to spend their time and efforts on generating useful computations (Yuen *et al*. 2009). Additionally, hiring human experts to perform computational tasks could be expensive (von Ahn and Dabbish 2004). In contrast, a HCG provides fun for people and in return, they contribute useful computations as a by-product (Hacker and von Ahn 2009). Hence, recruiting online game players could be a viable alternative to leverage human computation, potentially alleviating issues related to monetary rewards.

2.2. Human computation games

In essence, HCGs are designed to make boring tasks more interesting (Krause and Smeddinck 2011) so that humans are motivated and willing to contribute high-quality outputs in the course of gameplay. They can also be seen as a value-added paradigm building upon distributed collaborative environments by means of providing enjoyable experiences through games. HCGs are also termed as Games With a Purpose (GWAPs; von Ahn and Dabbish 2004) in the past literature because they serve a purpose of generating useful outputs as a side effect of gameplay. As such, they represent a genre of crowdsourcing, and are also known as crowdsourcing games (Guy *et al*. 2011). HCGs may also be considered a type of gamification or gameful design, terms introduced in recent literature, which basically utilises game design elements (such as badges) in non-gaming contexts to increase users' motivation and retention (Deterding *et al*. 2011). Used in this context, a HCG is a type of gamified application with a focal purpose of generating useful computations. Therefore, HCGs are also termed as computainment, representing the dual purpose of generating computations while providing entertainment (Goh *et al.* 2011b). In line with most recent literature (e.g. Azzopardi *et al.* 2012), this study adopts the term 'HCG' that reflects the central concept of generating useful computation through games.

As mentioned, due to its potential, numerous HCGs have been developed in various domains. Besides the ESP game described previously, a number of other HCGs have also been developed with different gameplay styles to tackle the image labelling problem. For instance, in contrast to the ESP game which requires players' cooperation to earn points, KissKissBan (Ho *et al.* 2009) employs a competitive perspective. Specifically, in order to win the game, a pair of collaborative players needs to generate matching labels which are not the same as those created by another player called the blocker, who competes against the former.

Beyond image labelling, HCGs have been used to foster mobile content-sharing behaviour. The Gopher Game (Casey *et al.* 2007) allows players to contribute geospatial information by means of performing information creation tasks through a game agent called a gopher. Indagator (Lee *et al.* 2010) incorporates gaming elements into content-sharing activities through which players can share and browse media-rich location-based information, and earn points by rating and creating content. SPLASH (Goh *et al.* 2011a) adopts a pet-based game genre in which players feed location-based multimedia content to virtual pets.

In the area of music annotation, HCGs have been developed to address the lack of descriptive labels for music content on the Web. One example is TagATune (Law and von Ahn 2009), in which pairs of players decide whether they are hearing the same song based on other players' hints. Next, Herd It (Barrington *et al.* 2009) is a multiplayer, social game for collecting users' opinions for a given song. Players have to agree upon a label for a given song with other players to earn game points. In the natural language processing domain, Sentiment Quiz (Rafelsberger and Scharl 2009) and Jinx (Seemakurty *et al.* 2010) are used to gather sentiment information of a given statement and to solve the word sense disambiguation problem, respectively.

Moreover, FoldIt (Cooper *et al.* 2010) has shown that HCGs are useful in solving problems in the biochemistry domain. In this game, players produce protein structures that have lowest energy conformation so that they can help to determine the proteins' function. Next, HCGs have been utilised to reduce the critical path length of the field programmable gate arrays (Plummings; Terry *et al.* 2009), with the resulting outputs used to improve digital circuit design. In summary, the brief review of existing literature suggests that HCGs are versatile and that they have potential to address computation problems in various domains. This study, therefore, proposes a typology centred on the salient dimensions of HCGs to enhance the conceptual understanding of this research area.

2.3. Related classification systems

Although there is no classification system that provides a comprehensive understanding of HCGs, they have been found in generic HCSs. Yuen *et al.* (2009) classified HCSs into three categories based on their evolutionary stages: initiatory, distributed, and HCGs. Initiatory HCSs collect useful knowledge from users to improve artificial intelligence or machine learning systems, while distributed systems harvest vast amounts of information from Internet users. Thereafter, HCGs emerged to channel the tremendous amount of time spent on playing online games into solving human computation problems. Additionally, HCGs were classified using game-based properties including game structure, verification method, game mechanism, gameplay, and number of players. However, Yuen *et al.*'s (2009) classification of HCGs did not include all their essential features, for instance, incentive schemes which are an important motivator for participation (Mason and Watts 2009).

Another classification system was developed by Quinn and Bederson (2011), consisting of six dimensions of HCSs including motivation, quality, aggregation, human skill, participation time, and cognitive load. According to this system, HCGs use enjoyment as their primary motivation mechanism, and utilise different types of quality control mechanisms including redundancy, multilevel review, and automatic check. However, this system is not able to describe how a computational task is integrated into gameplay as these dimensions are deduced from existing HCSs. Next, Doan *et al.* (2011) classified online crowdscouring systems with nine dimensions including nature of collaboration, target problems, roles of human users, tasks performed by users, recruitment and retainment strategies, ways to combine user contributions, evaluation methods, architecture, and degree of manual effort. In this scheme, HCGs were classified as implicit systems where players perform computational tasks as part of gameplay. Further, this scheme suggests that players can take multiple roles including content, perspective, and component providers, as well as slaves. However, like Quinn and Bederson (2011), this system does not consider how game-based properties of HCGs are incorporated.

In the context of HCGs, three general game structure templates were proposed by von Ahn and Dabbish (2008): output-agreement, input-agreement, and inversion-problem. In HCGs that apply output-agreement, the success condition is achieved only when players produce matching outputs with others, whereas the winning condition of HCGs with input-agreement will be obtained when both players guess correctly whether or not they are given the same inputs based on their partners' outputs. In inversion-problem HCGs, the victory condition is achieved when a player can guess the correct output based on the cues or hints provided by his/her partner. Although these strategies are widely used in HCGs, there are some other verification mechanisms (e.g. binary verification) that have been successfully employed in current games. Further, as in the systems described above, other HCG dimensions are not considered.

In summary, previous work has focused on HCSs. For those few that deal with HCGs, a single dimension or a few game-based properties is considered. Classification of HCGs only from a single aspect is insufficient because these games are multifaceted, inheriting features from both HCSs and games for entertainment. Simply put, acquiring useful outputs in HCGs is as equally important as providing fun to players. To address this gap, we develop a multi-dimensional typology of HCGs which is based on the 12 most prominent dimensions of such games, and is described in the following sections.

3. Methodology

Our work includes searching articles concerning HCGs and analysing them to identify the salient dimensions for the purposes of developing a typology. This approach is consistent with similar work such as Hong *et al*. (2009) and Brandtzæg (2010). The steps involved are described as follows.

- *Step 1: Literature Search Scope and Article Selection Criteria*. First, articles were searched in online academic databases including ACM Digital Library, IEEE Xplore, Science Direct, and SpringerLink. These databases have also been used in past studies (e.g. Hong *et al*. 2009, Brandtzæg 2010). The following search terms were used: 'human computation', 'HCG', 'GWAPs', 'crowdsourcing games', and 'gamification'. The total number of articles found was 120. Next, 17 articles were excluded because they did not primarily discuss about HCGs. For instance, the article by Chamberlain *et al*. (2009) was selected as it describes how incentive schemes are structured in HCGs, but that of Chilton *et al*. (2010) was excluded as its main focus is on workers' tasks searching behaviours in a human computation market, rather than on HCGs in particular. Finally, a total of 103 articles from 2004 to 2012 were selected for this study.
- *Step 2: Analysis of the Articles.* The objective of this stage is to develop classification dimensions by analysing the HCGs in the selected articles. These dimensions were derived both inductively and deductively. First, dimensions were drawn from previous classification schemes of HCSs (e.g. Yuen *et al*. 2009, Quinn and Bederson 2011). Output quality control mechanisms, aggregation methods, roles of human users, game mechanism, gameplay synchronicity, verification mechanism, and player composition were identified. Next, the selected articles in Step 1 were classified using these dimensions. Specifically, in every HCG, we looked for these dimensions and identified how they are implemented. Next, following an inductive approach, we scrutinised every game to identify new dimensions. Once new dimensions were found, we tried to apply them to every other game. If they were applicable to most HCGs in the review, these were introduced into the typology. The dimensions that were inductively derived from the analysis include interaction mode, architecture, incentives, scoring systems, and collection. This process continued until the dimensions and associated strategies that could describe almost all HCGs were collected.
- *Step 3: Verification of Typology*. Finally, to verify our typology, two independent coders were assigned to classify the HCGs in the articles according to identified dimensions of the previous steps. Intercoder reliability was then assessed. Before the actual coding process began, a pilot study was conducted with both coders to make them familiar with the coding scheme and purpose of the study. Coders were shown how to look out for dimensions and their associated strategies in the HCGs. Both coders then performed the coding process independently. The results showed a strong agreement between coders with a Cohen's Kappa value of 0.842 (Stemler 2001).

3.1. *Description of the articles*

Most of the articles were published in conference and workshop proceedings while some of them were found as journal articles and technical reports. The articles by publication sources are categorised in Table 1.

Based on the analysis, it was found that most of the articles reviewed focused on design and deployment in various domains. Among articles selected, 72.72% were about design and development of HCGs to solve particular problems (e.g. von Ahn and Dabbish 2004), while 9.1% of the articles presented generic game design frameworks (e.g. Reeves and Sherwood 2010). Additionally, 7.79% compared the performance of different HCGs genres used to solve a specific problem (e.g. Goh *et al*. 2011b), while only 2.6% studied motivations of HCG players (e.g. Lee *et al*. 2010). As well, 7.79% of the articles described HCGs as a kind of HCSs and presented salient features of such games (e.g. Quinn and Bederson 2011). The articles used in this study are summarised according to their characteristics and depicted in Figure 1.

Articles in our analysis were from various domain areas such as image labelling, music annotation, common-sense reasoning, natural language processing, sentiment analysis, and location-based information sharing, and many others. Figure 2 shows the distribution of the number of articles by target human computation problems solved.

Table 1. Classification of articles by source.

Journal articles	Number of articles	Source
CHI Conference on Human Factors in Computing Systems	15	Conference
Communications of the ACM	2	Journal
Human Computation Workshop (HCOMP)	18	Workshop
IEEE Intelligent Systems	1	Journal
IEEE Potentials	1	Journal
Conference on Advances in Computer Entertainment Technology (ACE)	5	Conference
Conference on Artificial Intelligence (AAAI)	4	Conference
Conference on Asia-Pacific Digital Libraries (ICADL)	3	Conference
Conference on Digital Information Management (ICDIM)	3	Conference
Conference on Multimedia (MM)	3	Conference
Conference on Multimedia Information Retrieval (ICMR)	2	Conference
Conference on Music Information Retrieval (ISMIR)	2	Conference
Conference on Semantic Systems (I-SEMANTICS)	2	Conference
World Wide Web conference (WWW)	6	Conference
Journal of Information Science	1	Journal
Journal of the American Society for Information Science and Technology (JASIST)	2	Journal
Nature	1	Journal
Nordic conference on Human-Computer Interaction (NordiCHI)	2	Conference
International Conference on Information Technology : New Generations (ITNG)	2	Conference
Others[a]	22	Conference
Special Interest Group on Information Retrieval (SIGIR) Conference	2	Conference
Technical Reports	6	Technical report
Total	103	

Note: [a]Conferences that published only one article related with HCGs.

4. Classification dimensions

From the analysis of articles, 12 key dimensions of HCGs and various related strategies were uncovered. These dimensions were then grouped under three metacategories, namely gameplay mode, game structure, and data. The typology is presented in Table 2, and discussed in detail in the following sections.

4.1. Gameplay mode

The gameplay mode metacategory represents how players' actions and interactions are designed in HCGs (Salen and Zimmerman 2004), and described through what players' actions are needed, how players interact, and what roles they take. In the following, three gameplay dimensions, synchronicity, interactivity, and roles of human player, as well as associated strategies are discussed.

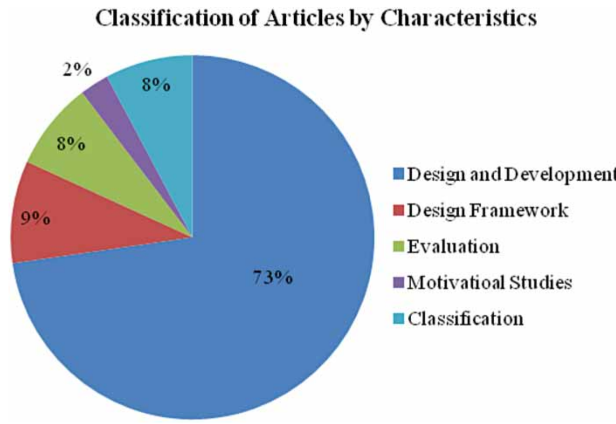

Figure 1. Classification of articles by their characteristics ($N = 103$).

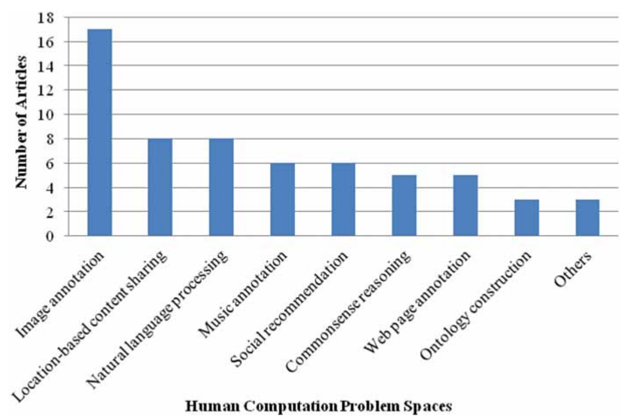

Figure 2. Distribution of article numbers by human computation problem spaces.

4.1.1. Synchronicity

This dimension describes whether or not simultaneous player action is needed during gameplay (Yuen *et al*. 2009), and can be categorised into two modes: synchronous and asynchronous. In synchronous gameplay, players need to be logged in exactly at the same time and provide real-time responses to each other (Yuen *et al*. 2009). In asynchronous gameplay, players take turns during gameplay and real-time responses are not needed (Casey *et al*. 2007).

Synchronous HCGs could be useful to promote a sense of spontaneity among players, that is, all players can see the outcomes of their contributions at the same time. It could also be suitable when there is not enough verified information to determine the correctness of outputs. In Verbosity (von Ahn *et al*. 2007), two players take different roles:

Table 2. Typology of human computation games.

Metacategory	Dimension	Description	Possible strategies		Example
Gameplay mode	Synchronicity	Types of player actions needed during gameplay	Synchronous		Matchin, Verbosity
			Asynchronous		Gopher Game, Eyespy
	Interactivity	Modes of interaction among players within gameplay	Direct		Rapport Game, Listen Game
			Indirect		Virtual Pet Game, Gopher Game
	Roles of human player	Player's roles of involvement in solving computational task	Content provider		Eyespy, PhotoCity
			Content analyzer		ESP game, Verbosity
Game structure	Game mechanics	Rules of the gameplay that players need to abide to arrive at the victory condition	Collaborative		ESP game, Herd It
			Competitive		Competitive HCG described in Goh *et al.* (2011b), KissKissBan
			Hybrid		Indagator, SPLASH
	Player composition	Player organisation in the game	Single-player		GuessWho, Dogear Game
			Two-player		ESP game, Peekaboom
			Multi-player		Gopher Game, Herd It
			Two-team		KissKissBan
			Multi-team		PhotoCity
	Architecture	Implementation and design of the game	Standalone		ESP game, Phrase Detective
			Platform-based		Sentiment Quiz, Virtual Pet Game
	Incentives	Incentive structure used to encourage participation in human computation	Social incentives		Foldit, SPLASH
			Financial incentives	Monetary rewards	Phrase Detective
				In-game currency	Mini-HCGs in SPLASH
			Score points		TagATune, GuessWho
	Scoring system	Scoring mechanisms used to reward players	Collaborative		Herd It, Matchin
			Comparative		Dogear Game, Phrase Detective
			Hybrid		GiveALink
Data	Verification mechanisms	Mechanisms used to determine correctness of the output	Social verification	Output-agreement	ESP game, Sentiment Quiz
				Inversion-problem	Peekaboom, Verbosity
				Input-agreement	TagATune, Page Match
			Binary verification		OnToGalaxy, FastTag
			Automatic check		Foldit
			Multilevel review or peer review		Gopher, Rapport Game
	Control	Methods used to ensure the quality of outputs beyond the commonly used verification methods	Redundancy		Gopher Game, GuessWho
			Player testing or ground truth seeding		Peekaboom
			Random-selection		FastTag, Jinx
			Player outcome prediction		Tag4Fun
	Aggregation	Methods used to combine individual contributions to solve the target problems	Collection		ESP game, Page Hunt
			Iterative improvement		CityExplorer, Gopher Game
	Collection	Approaches used to collect data	Open-ended		ESP game, Jinx
			Close-ended		Sentiment Quiz, Herd It

either the narrator or the guesser to produce common-sense facts. The narrator provides hints about a secret word to the guesser who concurrently has to make a correct guess to win the game.

The asynchronous approach is useful for situations when real-time matches between players are very unlikely to occur, and it may happen when players do not have same level of knowledge about a given computational task (Guy *et al.* 2011). For instance, in recommending bookmarks for a particular person in an organisation, two players may not always have the same opinions about the target person, thereby contributing different bookmarks to him/her (Guy *et al.* 2011). Eyespy (Bell *et al.* 2009) is an example of HCG in which players create photos and tags about geographic locations, and earn points when their created photos and tags are confirmed by other players at a later time.

4.1.2. Interactivity

This describes how HCGs support interaction among players. Fundamentally, interaction can be defined as two-way or multi-way exchange of data among computing agents (Keil and Goldin 2005). Our review suggests two interaction modes: direct and indirect. In direct interaction, players perform game tasks through a direct communication channel with other players. Direct interaction may be anonymous or non-anonymous. In the former, players' identities are not revealed to others players. In non-anonymous mode, players know who they are playing with because the game allows them to access profile information of other players such as name and photos (Kuo *et al.* 2009). Next, in indirect interaction, players perform game tasks through an indirect communication channel with other players, typically, via a game agent. This mode is implicitly anonymous because players have no knowledge about who receives and responds to their actions (Keil and Goldin 2005).

In non-anonymous direct interaction mode, performance is more likely to improve because players may be more cautious about their actions when they know who their partners are (Kuo *et al.* 2009). However, such games have a high probability of cheating as players could collude with others to manipulate the game. The Rapport Game (Kuo *et al.* 2009) for collecting common-sense knowledge provides direct interaction, allowing players to ask questions to players of their choice, and answer questions asked by others.

Indirect interaction is a valuable approach to integrate rich forms of social interactions in gameplay while providing partial anonymity through communication with a game agent. With this interaction, asynchronous HCG play could also be achieved. Since the interactions among players are not predefined, but dynamically constructed, decentralised coordination or self-organisation can be developed among players (Keil and Goldin 2005). The Virtual Pet Game (Kuo *et al.* 2009) employ indirect interaction in which players answer questions of their pets that are asked by other players, and ask questions to their pets which will be responded to by other players.

4.1.3. Roles of human players

This dimension describes possible roles of human players in HCGs, and can be categorised as content provider and content analyzer. Some HCGs allow players to take both roles within a single game. As a content provider, players are primarily involved in contributing content which can later be analysed and validated by others (Doan *et al.* 2011). As a content analyzer, the primary task involves analysing and making judgements about the content to produce meaningful values for them (Reeves and Sherwood 2009).

The content provider role is useful in situations where players are required to create something new or when collecting the diversity of data is important to solve the target problem. EyeSpy (Bell *et al.* 2009) is an example in which players contribute photos and text of geographic locations that can support map navigation tasks. The content analyzer role is useful for situations where a thorough analysis of a given object is needed to produce a meaningful value for it. This approach could be used to acquire in-depth opinions for a given computational task. The ESP game (von Ahn and Dabbish 2004) is an example in which players are required to analyse given images and generate descriptive labels for them.

4.2. Game structure

This metacategory describes how HCG elements are organised so that the game performs its designated purpose (von Ahn and Dabbish 2008). Five game structure dimensions of game mechanics, player composition, architecture, incentives, and scoring systems, and their variants are identified and discussed in the following.

4.2.1. Game mechanics

This dimension describes the rules of gameplay that players need to abide to arrive at the victory condition (Sicart 2008). In other words, it identifies the players' relationship within the game (Yuen *et al.* 2009). Our review suggests three different types of strategies which are widely used: collaborative, competitive, and hybrid.

From our review, most current HCGs are collaborative in nature, requiring players to work together to reach the winning condition of the game (Goh *et al.* 2011b). In contrast, some HCGs are implemented as competitive games in which players compete with or work against each other to achieve the winning condition (Ho *et al.* 2009). The mode of competition may vary depending on how the winning condition is defined, for example, it could be time-based and/or outcome-based. Hybrid HCGs possess both collaborative and competitive properties.

Collaborative games are preferable in situations where agreement among users is vitally important to produce a meaningful output. This type of HCG is suitable for collecting data at an initial stage when there is no sufficient data to gauge the quality of outputs yet. However, players tend to contribute more generic outputs in this type of HCGs (Weber *et al*. 2009). Herd It (Barrington *et al*. 2009) is an example which encourages collaboration by awarding points based on the percentage agreement among players on the tag for a given song. As in any other kind of games, cheating could exist in HCGs. Basically, cheating is defined as an activity performed by players in order to gain advantage over another players or to win the game without following the specified game rules (Thompson 2007). As such, players could take advantage of collaborative HCGs by forming a coalition with their partners to win the game, resulting in low data quality (von Ahn and Dabbish 2008), for instance, generating same output for every game round.

Competitive HCGs could be used to alleviate the cheating problem (Ho *et al*. 2009). As players are in competition, coalitions are unlikely to form. However, this type of HCG is more useful only when there is enough data or an automatic system to evaluate the correctness of the outputs. KissKissBan (Ho *et al*. 2009) is an example in which a blocker and a pair of collaborative players compete such that the blocker prevents the others from reaching agreement. Hybrid HCGs could also be used in a situation where collaboration is required to complete a game task, at the same time providing a competitive experience among players. However, this type of game may have more complex game rules than those that employ a single mechanic. PhotoCity (Tuite *et al*. 2010) is an example in which players in each team work together to own a building by uploading photos of it while two or more teams are competing for the ownership of this building.

4.2.2. Player composition

This dimension describes how players are organised in the game and includes single-player, two-player, multi-player, two-team, and multi-team. A single-player game refers to the solo game in which players play alone, whereas two players working together or opposing each other comprises the two-player game (Yuen *et al*. 2009). In multi-player games, individual players compete or collaborate to achieve the winning condition (Elverdam and Aarseth 2007). Finally, games with two or more teams refer to two-team or multi-team games, respectively (Tuite *et al*. 2010).

The advantage of single-player HCGs is that it is able to start once one player has logged in. As such, it could alleviate the 'cold-start' problem, which is caused by the lack of user participation (Dulačka *et al*. 2012). However, such HCGs could thwart the social experience during gameplay which is one of the reasons of playing online games (Yee 2006). GuessWho? (Guy *et al*. 2011) is a single-player HCG for collecting knowledge about a person from his/her peers in which the player is required to create a name of a person who is related to the given person (i.e. relationships), or tags for a given person.

The advantage of a two-player HCG is that the correctness of outputs could be determined based on the mutual agreement between players. In other words, this type of HCG is useful where consensus among players is necessary to solve the designated problem. Curator (Walsh and Golbeck 2010) is a two-player HCG in which both players create collections of two given sets of items on their own, and they are then awarded game points based on their matching collections. The results could be useful for collection recommender systems. Similar to two-player HCGs, correctness of outputs could be determined based on mutual agreement between players in multiplayer HCGs. This type of HCG can also enhance players' social experience (Barrington *et al*. 2009). The Gopher game (Casey *et al*. 2007) is a multiplayer HCGs in which players pick up and perform any game tasks of interest that are created by other players.

A two-team HCG could be utilised to foster collaboration as well as competition. This is because in such games, players need to perform game actions that complement their teammates while defeating members of another team. Put simply, players in each team work together to generate outputs while competing with another team. KissKissBan (Ho *et al*. 2009), described earlier, is a two-team HCG in which one player who is a blocker competes with a team of two players. The competitiveness of multi-team HCGs could be higher than two-team games as players encounter more opponents. However, this type of game may suffer from the 'cold-start' problem as multiple players are required to participate. Although non-player characters or bots could be utilised if there are insufficient players, it could degrade the performance of the game (von Ahn *et al*. 2006). PhotoCity (Tuite *et al*. 2010), described earlier, is an example of a multi-team HCG.

4.2.3. Architecture

The architecture dimension describes how HCGs are implemented, and can be standalone or platform-based. In standalone architectures, HCGs are developed as independent games hosted on their own servers. The popularity of social networking sites and community-based websites in recent years has led to the evolution of platform-based variants which piggyback on established systems to take advantage of their existing user bases (Rafelsberger and Scharl 2009).

Standalone HCGs could be useful to attract people who wish to meet new like-minded people through playing games. In this way, HCGs could become a new genre of user-generated content outlet, such as Wikipedia and YouTube, in which users not only create content, but also manipulate and consume them (Nov 2007). However, this kind of game needs to recruit users explicitly which may cause the 'cold-start' problem. Verbosity (von Ahn *et al.* 2007) and Phrase Detective (Chamberlain *et al.* 2009) are examples of standalone HCGs.

Platform-based HCGs could be used to alleviate the problem related to low participation by recruiting existing users on these platforms through their friends' invitations or notification features. However, this type of HCG needs to have attractive features to grab users' attention more than other games on the same platform. Further, a platform may become obsolete and users may switch to a newer one. Sentiment Quiz (Rafelsberger and Scharl 2009) is built upon Facebook, which is a popular social networking website, and utilises its notification feature to attract players.

4.2.4. Incentives

Although individuals play games mainly to seek fun or pleasure, their behaviours within gameplay can be controlled by various types of rewards (Elverdam and Aarseth 2007). Therefore, this dimension describes how different incentives schemes are utilised to reinforce players' actions and deepen engagement with HCGs. Three incentive schemes are commonly used: social, financial, and game points. Essentially, social incentives are provided by integrating various modes of social interactions as part of HCG play, while financial incentives are provided in the form of actual monetary payment or in-game currency which can be used to buy in-game items (Chamberlain *et al.* 2009). Finally, game points are an indicator of one's success in the game, and are used to rank or level-up players' within the game.

To provide social incentives, various forms of communication among players, such as in-game chat rooms, forums, and implementing games on social networking sites, have been utilised (Chamberlain *et al.* 2009). HCGs with social incentives could be useful to attract players who tend to create and maintain social relationships through gameplay (Yee 2006). For example, Foldit (Cooper *et al.* 2010) provides an in-game chat room and SPLASH (Goh *et al.* 2011a) provides a virtual room to promote socialising.

Schemes which use real money to reward participants attract players through extrinsic motivation. Phrase Detective (Chamberlain *et al.* 2009) is an example in which players are rewarded with real money by random-selection on a weekly basis, and three highest scorers on a monthly basis. However, money may reduce intrinsic motivation of players and demotivate others who play HCGs to gain social standing or contacts (Law and von Ahn 2011). Therefore, it is necessary to investigate the effects of momentary rewards on HCG players' motivation. This issue has become more important when certain groups of people around the world have started participating in computational tasks to make a living (Ross *et al.* 2010), and thus legal and ethical issues, such as labour exploitation, need to be taken into account (Fort *et al.* 2011). For instance, players may not have employee status because they perform computational task as a side-effect of gameplay (Zittrain 2008). The effects of low wages and the determination of an optimal amount of monetary incentive should also be considered in designing HCGs.

HCGs which use in-game currency schemes should have attractive virtual goods so that users will try to earn currency to purchase them by completing game tasks. Therefore, players who wish to collect and possess many objects in games are more likely to be attracted to this type of game (Bostan 2009). The mini-HCGs in SPLASH (Goh *et al.* 2011a) are examples where players earn in-game currency by providing information about locations, and this can later be utilised to decorate their avatars and virtual rooms.

Game points are a commonly used incentive scheme in HCGs, and also facilitate performance assessment among players. They can be considered as a basic incentive mechanism, typically acquired after completing certain milestones (Elverdam and Aarseth 2007). Recently, it has become a popular reward system in most gamified applications; for instance, awarding users with points for visiting new web pages and sharing links on social networking sites (Hamari and Eranti 2011). TagATune (Law and von Ahn 2009) is an example in which pairs of players are rewarded points for each successful guess about whether they are listening to the same audio clip.

4.2.5. Scoring system

This dimension describes how HCGs award game points, in-game currency, and monetary payment to players. Basically, three types are used: collaborative, comparative, and hybrid. Collaborative scoring systems reward players by allocating points based on the agreement among players, whereas in comparative scoring systems, players are awarded points for agreeing with ground truth data (Chamberlain *et al.* 2009). HCGs sometimes employ both scoring systems within a single game.

A collaborative scoring system could be utilised to encourage players to produce more matching outputs, and also useful for situations when there are no sufficient verified facts to assess the outputs to be right or wrong so as to determine whether the players win or not. One such example is Herd It (Barrington *et al.* 2009), which gives points based on the percentage agreement of a player's contributed labels of music files with others. On the other hand, a comparative scoring system could be useful when there is a large collection of standard answers or benchmarks that could be used to check against the outputs to identify the winner of the game. Dogear (Dugan *et al.* 2007) employs a comparative scoring system of checking players' selected bookmarks against the underlying Dogear database.

A HCG with a hybrid scoring system could produce more reliable and widely accepted outputs among players. In particular, the reliability of standard outputs could be verified and popular outputs could be captured based on majority confirmation. However, the difficulty level of earning rewards is increased because players need to produce matching outputs with others as well as with benchmarks. GiveALink (Weng and Menczer 2010) uses a hybrid scoring system in which player contributed tags for web pages are checked against the existing GiveALink database as well as previous players' tags to award points.

4.3. Data

This metacategory describes how HCGs deal with user-generated outputs. Specifically, it covers how such games ascertain the quality of user-generated outputs and aggregate the collected outputs, as well as what types of data collection approaches are used based on the nature of outputs required to tackle target problems.

4.3.1. Verification mechanisms

This dimension describes the mechanisms used to determine correctness of the output generated. As output quality is equally important as entertainment in HCGs, built-in verification methods are integrated as part of game rules so that players' contributions could be validated during gameplay. Generally, HCGs use four verification mechanisms: social, binary, automatic check, and multilevel or peer review.

Social verification, also known as the social agreement strategy, relies on the mutual agreement between two or more players, and contributions are not accepted until agreement is achieved (Ho and Chen 2009). Output-agreement, inversion-problem, and input-agreement are widely used (von Ahn and Dabbish 2008). In output-agreement, an output is considered correct only if players produce the same result for a given input. In inversion-problem, the output is considered correct only if a player can successfully guess the original input by using the clues given by other players. Finally, in input-agreement, the output is considered correct only if players can successfully guess whether they are given the same or different input by examining each other's output.

In binary verification, players need to determine whether an object is relevant or non-relevant to the given category, but they are not allowed to create a new category (Krause and Aras 2009). In automatic check, players' outputs are checked against those generated from other automated systems (Cooper *et al*. 2010). Finally, in multilevel or peer review, peer assessment is achieved by means of rating or voting (Law and von Ahn 2011), and this is often used as an additional quality control mechanism along with other verification mechanisms.

Social verification is more suitable when there is no limit to the number of possible outputs to be collected from users, and commonly accepted outputs are necessary to solve the target problem. The ESP game (von Ahn and Dabbish 2004) is an example using output-agreement where players win only if they agree on a label produced by them. Peekaboom (von Ahn *et al*. 2006) employs the inversion-problem mechanism in which both players win the game if the player can correctly guess the tag for an image based on parts of it revealed by his/her partner. TagATune (Law and von Ahn 2009) applies input-agreement in which user-generated tags are considered correct only if both players can make a correct guess about the song they heard as same or different by examining their partners earlier created tags.

The binary verification mechanism is useful when there is a collection of pre-generated outputs so that players can select the relevant ones from the given choices. HCGs with binary verification may produce outputs faster than those with non-binary verification mechanisms though new outputs generation is constrained. OntoGalaxy (Krause *et al*. 2010) is an example which asks players to collect all synonyms of a given word from a given list of words. Automatic check is useful for situations in which the correctness of outputs could be better determined by checking them against those automatically generated from software. For instance, in an automated planning and scheduling context, it is easy to verify the correctness of a plan through backward reasoning (Quinn and Bederson 2011). FoldIt (Cooper *et al*. 2010) is a typical example in which players' contributed protein structures are evaluated against those generated by the Rosetta algorithm.

Finally, the multilevel or peer review mechanism could be useful to identify the best outputs (Law and von Ahn 2011) in which multiple workers are tasked to perform the voting, and the one that receives most votes is regarded as the best output. This strategy is similar to output-agreement mechanism, but it allows asynchronous verification. Gopher (Casey *et al*. 2007) and Rapport Game (Kuo *et al*. 2009) are examples providing voting mechanisms by which players judge other users' contributions.

4.3.2. Control

This dimension describes how HCGs ensure quality of outputs beyond the commonly used verification methods described earlier. Quality control mechanisms are essential because players may elude built-in verification mechanisms by forming a coalition with other players (von Ahn and Dabbish 2004), and thus no useful outputs will be collected. Four different quality control mechanisms are typically used in past literature: redundancy, player testing or ground truth seeding, random-selection, and player outcome prediction.

First, in redundancy (Quinn and Bederson 2011), the output is considered as reasonably correct only if the same output is produced by multiple players. One way to accomplish this is by assigning each task to multiple players. Next, in player testing or ground truth seeding, player reliability is assessed by giving tasks with both known and unknown outputs and checking their outputs against ground truth data (von Ahn and Dabbish 2008). Random-selection is another widely used mechanism where anonymous players are randomly paired or pre-generated outputs are randomly presented to players (Krause and Aras 2009). Finally, in player outcome prediction, players' outputs are predicted, and the next task for the player is determined given a particular prediction (Seneviratne and Izquierdo 2010).

The redundancy approach could be used where real-time synchronisation among players is not available to validate the correctness of outputs. As individuals may have different perceptions about the target problem, multiple assignments of tasks could help to minimise the variations of outputs produced by them (Law and von Ahn 2011). GuessWho? (Guy *et al*. 2011) is an example of HCG with repetitive task assignment. Player testing could be useful when a sufficient number of correct outputs to assess the trustworthiness of players are available. This allows the game to decide whether or not the player's contributed outputs should be included. Peekaboom (von Ahn *et al*. 2006) is an example which checks player's reliability with seeded images.

Next, random-selection could be used to ensure that players do not receive the same sequence of inputs in every gameplay section. Here, FastTag (Krause and Aras 2009) utilises random-selection in which each of the randomly paired players is given a random order of web page fragments. Finally, player outcome prediction could be more computationally expensive than other control mechanisms, but it could improve the performance of the game by assigning suitable tasks to players based on prediction. Here, the HCG presented by Seneviratne and Izquierdo (2010) applies player outcome prediction in which prediction models, using Markov Models and Bayes theorem, are employed to select suitable images for players such as fully annotated or non-annotated.

4.3.3. Aggregation

Aggregation describes how HCGs combine contributions to solve the target problem (Quinn and Bederson 2011). Aggregation methods used in current HCGs are categorised as collection and iterative improvement. In collection, contributions are treated as discrete facts, and each contribution is considered as a solution for a target problem (Quinn and Bederson 2011). Specifically, individual contributions are considered collectively to produce meaningful results for the target problem. Thus, statistical methods, such as average or median, could be used to aggregate contributions. In iterative improvement, players need to refine their contributions repeatedly until reviewers approve (Matyas *et al*. 2008).

The collection method is useful when a given task can ONLY be solved by users collaboratively to produce a meaningful result, and consensus among users is important to obtain a significant contribution. The ESP game (von Ahn and Dabbish 2004) is an example where only consensus labels among players are considered as meaningful for the images. The iterative improvement method could be used in a situation which requires the quality of outputs to be improved incrementally by acquiring users' opinions on how to refine it rather than asking them to create new ones. CityExplorer (Matyas *et al*. 2008) is an example in which players' contributions, such as geospatial data, are approved or rejected by other players. The rejected contributions have to be corrected and resubmitted for another review.

4.3.4. Data collection approaches

HCGs use one of two data collection approaches based on the target problems and/or the nature of outputs they expect: open-ended and close-ended. The open-ended approach allows players to provide arbitrary suggestions or free text, whereas the close-ended approach allows players to choose best-fit answers from predefined sets of choices (Rafelsberger and Scharl 2009).

The open-ended approach is more appropriate when diverse opinions or new ideas are to be collected. It can also be used as an initial step of generating outputs for problems so that the most popular outputs could be fed into close-ended HCGs to further assign confidence values (Rafelsberger and Scharl 2009). Jinx (Seemakurty *et al*. 2010) is an example which allows players to contribute replacement words or phrases for the highlighted ones in a given paragraph. A close-ended HCG could be suitable for a situation where the best-fit solutions for a target computational problem are to be identified. Sentiment Quiz (Rafelsberger and Scharl 2009) uses the close-ended approach and asks players to choose one of the given options that express the sentiment of the given statement.

5. Discussion

The typology presented in this paper clearly shows the diversity of HCGs that has been designed to date. Since the classification dimensions are derived from existing HCGs in various domain areas, this typology is somewhat dependent on how these games are designed. Nevertheless, this typology provides a conceptual understanding of HCGs because it highlights essential dimensions and various strategies to implement these games. This information could serve as a springboard for discovering new opportunities to leverage human computation effectively. From the review of the articles, this study found that much prior work tends to focus mainly on structures and mechanisms of HCGs, but less on the context and environment where such games are deployed. Although HCGs were initially used to solve computationally 'hard' problems such as computer vision problems (von Ahn and Dabbish 2004), they have been recently employed in enterprise context (e.g. Guy *et al*. 2011). Therefore, it would not be surprising that the future landscape may extend to HCGs for personal use, business, education training, and capturing social behaviours.

Moreover, this study found that in most prior work, game strategies for each dimension are identified primarily based on the nature of target computational problems and expected outputs. For instance, within the problem space of image annotation, the ESP game utilises game strategies of collaboration, synchronicity and anonymity, as well as an open-ended data collection approach to garner descriptive labels for images. Although the combination of these game strategies may yield intended outputs, it may not appeal to certain individuals. For example, certain individuals may prefer competition to collaboration, and are hence likely to perform better playing HCGs according to their preferences. Thus, in order to enhance the effectiveness of HCGs, designers should not only consider the requirements of target computational problems and relevant gaming strategies, but also the needs and preferences of target players. Understanding this information could serve as a first step towards developing a model for optimal HCG design, and should, therefore, be investigated further in future work.

One of the primary benefits of the typology is that it allows us to compare HCGs easily and precisely. In particular, the uniqueness, similarities, and differences among HCGs can be identified systematically by comparing the different strategies that have been implemented. Classifying HCGs in terms of games for pure entertainment genres such as quiz or board games is not sufficient because HCGs are designed to produce useful outputs too. For instance, both the ESP game and Phrase Detective can be classified under the quiz game genre. However, this typology informs us that these two games are different in terms of verification mechanism and gameplay synchronicity. In other words, the dimensions could be used to describe HCG genres in more detail by deconstructing the specific elements within each genre.

This typology could be helpful for researchers attempting to examine the nature of experiences or players' behaviours generated as a result of HCG play. Specifically, researchers may understand how different strategies used to implement HCG dimensions influence players' experiences and behaviours. For instance, Kuo *et al*. (2009) demonstrated that the mode of interactivity used in platform-based HCGs affected players' performance of such games. By understanding the impact of specific HCG dimensions, researchers can draw insightful conclusions about the consequences of such games. Next, understanding HCG dimensions and related strategies help researchers to identify other possible human computation problem spaces that have yet to be explored. For instance, as iterative improvement could be used to aggregate outputs in HCGs, the computational problems which are suitable to use such a strategy should be examined.

This typology is also beneficial to HCG designers in different ways. First, the key dimensions serve as a helpful starting point for designers who could conceptualise HCGs by applying appropriate strategies for each of them. Thus, pertinent aspects about each game could be generated before actual implementation (Björk *et al*. 2003). The typology could also serve as guidance for designers to adapt from possible design patterns from pure entertainment games (Krause and Smeddinck 2011). For instance, designers may learn how synchronous and asynchronous gameplays are implemented in pure entertainment games, and adopt such design patterns to the HCG context.

Second, the typology helps designers to make design choices in implementing new HCGs. Once the initial concepts of the game have been formulated, it can be analysed and improved by weighting the benefits and challenges of existing strategies to create more effective designs. For instance, designers could select a more appropriate incentive structure for a given computational problem (e.g. monetary rewards or in-game currency), considering the pros and cons of existing approaches. Additionally, it could be helpful to refine the implementation of HCGs by modifying existing strategies used, and/or replacing them with other strategies. Finally, this typology may help to expand the design space of HCGs. Through a detailed analysis of existing strategies, new ones may be uncovered. Designers may also learn the consequences of existing strategies in current HCGs so that mistakes are not repeated in the development of new games. For instance, as cheating is more likely to occur in collaborative HCGs, designers need to revise cheating-proof strategies, and to consider using other strategies such as punishment to discourage undesired behaviours (Christou 2011).

5.1. Future directions

This typology provides comprehensive information about HCGs than existing frameworks because it consists of multiple dimensions and various strategies used to implement these dimensions. For example, von Ahn and Dabbish (2008) mentioned that HCG design templates are primarily based on a single dimension, that is, social verification mechanisms. In contrast, this typology suggests, for instance, what data collection methods, and quality control mechanisms are employed. In comparison with Yuen *et al.*'s (2009) classification scheme which is mainly based on game properties, this typology provides additional information about how data generated by HCGs is dealt with, such as how data is collected and aggregated, as well as how data quality is ensured as part of gameplay. Although Quinn and Bederson (2011), and Doan *et al.* (2011) presented salient features of HCSs in their classification schemes, gaming dimensions of HCGs were not included. In order to maintain the comprehensiveness of the proposed typology, future research needs to analyse new HCGs along these dimensions. This will help to validate the proposed typology and also assist in discovering new dimensions and associated values that could be integrated into it.

Although this typology provides a deep understanding of HCGs and contributes to knowledge in the HCG design process, there are some limitations that may hinder its utility in research and development work. In particular, this typology presents various strategies to be used in designing HCGs. However, the suitability of these strategies with regard to target computational problems needs to be investigated empirically. Next, HCGs, like other types of games, consist of different game dimensions, and thus some individuals may be more attracted to certain types of HCGs than others. This typology, however, does not delve into how players' attitudes and norms are shaped by HCG designs. These issues need to be considered because they may affect not only players' enjoyment, but also their contributions. Therefore, understanding of users' needs and preferences will be of great help for devising effective HCG designs. This understanding could also facilitate the development of a model for optimal HCG design that identifies suitable HCG strategies for different types of individuals. Therefore, this study proposes the following research agenda based on the developed typology.

(1) What are suitable HCG designs that can be used to solve human computation problems efficiently?
Selecting appropriate HCG designs is important to effectively fulfil their entertainment-output generation duality (Goh and Lee 2011). The strategies used to implement HCG dimensions may expose certain strengths and weaknesses under different problem types, and research is required to investigate what the most suitable strategies for HCG design in relation to target computational problem spaces are. To date, only a few studies have conducted comparative evaluations of HCGs. Among them, Kuo *et al.* (2009) found that HCGs which supported direct communication among players produce better quality outputs than those with indirect interaction in the context of common-sense knowledge acquisition. However, it mainly focused on a single dimension of this typology and a specific computational problem space. Future research is needed to investigate HCGs from a multi-dimensional perspective to obtain a clearer understanding of their individual and collective influence on the quality and quantity of outputs produced.

(2) What are suitable ways to retain players or sustain their usage?
One of the primary challenges of HCGs is how to sustain players' long-term relationship with the game (Law and von Ahn 2011). Here, the dimensions of our typology could be influential factors. The following examples presented are not exhaustive, but represent a subset of possible research directions. Games on social networking platforms may have a greater chance to recruit users (Rafelsberger and Scharl 2009), and hence studies should be carried out to examine the potential impacts of platform-based HCGs on players' behaviours. Further, as players' behaviours may be influenced by their in-game achievements (Hamari and Eranti 2011), research on investigating the impacts of game achievements on sustained usage is needed.

(3) How do individual differences impact players' experiences and behaviours of HCGs?
Since different genres of HCGs can be developed by employing different strategies for each dimension, certain groups of users may be more attracted to particular HCG genres (Rollings and Adams 2003). Past studies on games for pure entertainment (e.g. Johnson and Gardner 2010) and HCGs (e.g. Goh *et al.* 2011b) demonstrated that individual differences in motivation influenced players' in-game experiences and behaviours, as well as preferences. Thus, different individuals may have varying preferences for HCG types, for example, collaborative or competitive, and may attain more enjoyment and perform better while playing their preferred HCG genres. Therefore, further research on how individual differences affect HCG players' experiences and behaviours should be conducted, and this understanding may be beneficial to the HCG design process.

(4) How can HCG be expanded into other platforms?
As players' participation is critical to the success of every HCG (Barrington *et al.* 2009), it is worthwhile to explore how to expand these games into possible platforms beyond what has previously been developed successfully. One possible way to achieve greater participation is the development of multi-platform HCGs. For instance, the deployment of HCGs on the mobile platform including smartphones and tablets allows individuals to play games anytime and anywhere (Quinn and Bederson 2011). In addition, GPS-enabled mobile devices enable players to collaborate or compete with other players across physical locations. The latest Casual Game Association (2012) report about social games stated that over 40 million users play social games on a daily basis. The large user base suggests opportunities for attracting and sustaining usage of HCGs. Hence, more research on effectively harnessing these new platforms is needed.

(5) What are possible research methodologies that can be used to enhance the understanding of HCGs?
It is worthy to note that much prior work on HCGs has mainly focused on designing such games to yield intended outputs for target computational problems. Little research has investigated the needs of HCG players (e.g. Goh *et al.* 2011b). As mentioned, the lack of this understanding may limit the utilisation of such games. Thus, empirical data to understand the needs of the players can help illuminate new areas of knowledge in the HCG context. For instance, with regard to the first research agenda item, future research may utilise qualitative interview methods to better understand individuals' experiences and perspectives on different HCG strategies. The collected data will shed light on

how different HCG strategies influence individuals' behaviours. Likewise, focus group interviews could be employed to gather perceptions, feelings, and attitudes towards certain HCG features, and the collected data will be of value in identifying features that appeal to different individuals. Next, surveys could be another possible research method to investigate individuals' experiences and behaviours in HCGs, in accordance with the third research agenda item. Further, observations of individuals' actual behaviours in natural HCG settings may yield a better understanding of the link between individuals' experiences and behaviours. This could be achieved through an automated data collection system that records individuals' gameplay activities over a specified period of time. Finally, the integration of self-reported survey data and secondary data may uncover new knowledge about HCGs which otherwise might not be easily observable.

6. Conclusion

As HCGs hold tremendous potential to solve a variety of problems in novel and interesting ways, many variants have been developed over the past few years. It has become correspondingly harder for researchers to understand their nature without having a typology that describes key aspects of such games in a systematic manner. To the best of our knowledge, there is a dearth of work in this regard. To address this gap, this study has proposed a multi-dimensional typology of HCGs based on a thorough review of such games in current literature. On the basis of proposed typology and literature review, opportunities for future research are also discussed.

Although our study has yielded useful insights, a few limitations exist. First, there may be some relevant articles that might have been overlooked in this study. This study reviews literature from the period between 2004 and 2012. This period is chosen because the usage of HCGs has become more prevalent around year 2004 along with the emergence of ESP games (von Ahn and Dabbish 2004). Thus, this typology does not cover articles outside this period, and there may also be some relevant articles that were not explicitly stated as focusing on HCGs. Despite this limitation, the findings of this study are of value and convey useful information about HCGs as the bulk of the articles in this period have been reviewed. Next, this typology may fail to classify new games that have emerged over time as the dimensions and associated values were deduced from existing HCGs. However, as this typology is open-ended in nature which allows new dimensions and values to be added to, the entire typology should not necessarily be abandoned. Finally, the relationships between HCG strategies, and players' needs and preferences could not be illustrated in this typology because most articles in the review are mainly concerned with the design and development, rather than investigating the behaviours of HCG players.

Funding

This work was supported by the Singapore National Research Foundation Interactive Digital Media R&D Program [grant number NRF2008IDM-IDM004-012].

References

von Ahn, L. and Dabbish, L., 2004. Labeling images with a computer game. *SIGCHI conference on human factors in computing systems*, 24–29 April 2004, Vienna. New York: ACM Press, 319–326.

von Ahn, L. and Dabbish, L., 2008. Designing games with a purpose. *Communication of the ACM*, 51 (8), 58–67.

von Ahn, L., Liu, R., and Blum, M., 2006. Peekaboom: a game for locating objects in images. *SIGCHI conference on human factors in computing systems*, 22–27 April 2006, Montréal. New York: ACM Press, 55–64.

von Ahn, L., Kedia, M., and Blum, M., 2007. Verbosity: a game for collecting common-sense facts. *SIGCHI conference on human factors in computing systems*, 22–27 April 2007, San Jose. New York: ACM Press, 75–78.

Azzopardi, L., Purvis, J., and Glassey, R., 2012. PageFetch: a retrieval game for children (and adults). *35th international on research and development in information retrieval*, 12–16 August 2012, Oregon. New York: ACM Press, 1010–1010.

Barrington, L., *et al.*, 2009. User-centered design of a social game to tag music. *15th ACM SIGKDD workshop on human computation*, 28 June–1 July 2009, Paris. New York: ACM Press, 7–10.

Bell, M., *et al.*, 2009. Eyespy: supporting navigation through play. *SIGCHI conference on human factors in computing systems*, 4–9 April 2009, Boston. New York: ACM Press, 123–132.

Björk, S., Lundgren, S., and Holopainen, J., 2003. Game design patterns. *Digital games research conference: level up*, 4–6 November 2003, Utrecht. Utrecht: University of Utrecht, 180–193.

Bostan, B., 2009. Player motivations: a psychological perspective. *Computers in Entertainment*, 7 (2), 1–26.

Brandtzæg, P.B., 2010. Towards a unified media-user typology (MUT): a meta-analysis and review of the research literature on media-user typologies. *Computers in Human Behavior*, 26 (5), 940–956.

Casey, S., Kirman, B., and Rowland, D., 2007. The gopher game: a social, mobile, locative game with user generated content and peer review. *4th international conference on advances in computer entertainment technology*, 15–17 June 2007, Salzburg. New York: ACM Press, 9–16.

Casual Games Association, 2012. *Social network games, casual games sector report* [online]. Available from: http://dl.dropbox.com/u/3698805/research/2012_CGA_SocialSector.pdf [Accessed 15 December 2012].

Chamberlain, J., Poesio, M., and Kruschwitz, U., 2009. A new life for a dead parrot: incentive structures in the Phrase Detectives game. *Paper presented at the webcentives workshop*, 20–24 April 2009, Madrid, Spain.

Chilton, L.B., *et al.*, 2010. Task search in a human computation market. *16th ACM SIGKDD workshop on human computation*, 25–28 July 2010, Washington, DC. New York: ACM Press, 1–9.

Christou, G., 2011. Sanctions, punishment, and game design: designing MMORPGs for fair treatment of players. *Middle Eastern simulation and modeling conference, GAMEON-ARABIA*, 14–16 November 2011, Amman, Jordan: Arab Open University, 144–148.

Cooper, S., *et al.*, 2010. Predicting protein structures with a multiplayer online game. *Nature*, 466, 756–760.

Deci, E.L. and Ryan, R.M., 2000. The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. *Psychological Inquiry*, 11 (4), 227–268.

Deterding, S., *et al.*, 2011. From game design elements to gamefulness: defining "gamification". *15th international academic MindTrek conference: envisioning future media environments*, 28–30 September 2011, Finland. New York: ACM Press, 9–15.

Doan, A., Ramakrishnan, R., and Halevy, A.Y., 2011. Crowdsourcing systems on the world-wide web. *Communication of the ACM*, 54 (4), 86–96.

Dugan, C., *et al.*, 2007. The dogear game: a social bookmark recommender system. *2007 international ACM conference on supporting group work*, 4–7 November 2007, Sanibel Island. New York: ACM Press, 387–390.

Dulačka, P., Šimko, J., and Bieliková, M., 2012. Validation of music metadata via game with a purpose. *8th international conference on semantic systems*, 5–7 September 2012, Graz. New York: ACM Press, 177–180.

Elverdam, C. and Aarseth, E., 2007. Game classification and game design: construction through critical analysis. *Games and Culture*, 2 (3), 3–22.

ESA, 2011. *Annual report 2011* [online]. Available from: http://www.theesa.com/about/ESA_2011_Annual_Report.pdf/ [Accessed 15 December 2012].

ESA, 2012. *Essential facts about the computer and video game industry* [online]. Available from: http://www.theesa.com/facts/pdfs/ESA_EF_2012.pdf/ [Accessed 15 December 2012].

Fort, K., Adda, G., and Cohen, K B., 2011. Amazon Mechanical Turk: gold mine or coal mine? *Computational Linguistics,* 37 (2), 413–420.

Goh, D.H.-L. and Lee, C.S., 2011. Perceptions, quality and motivational needs in image tagging human computation games. *Journal of Information Science,* 37 (5), 515–531.

Goh, D.H.-L., *et al.*, 2011a. SPLASH: blending gaming and content sharing in a location-based mobile application. *3rd international conference on social informatics*, 6–8 October 2011, Singapore. Heidelberg: Springer, 328–331.

Goh, D.H., *et al.*, 2011b. Fight or unite: investigating game genres for image tagging. *Journal of the American society for information science and technology*, 62 (7), 1311–1324.

Guy, I., *et al.*, 2011. Guess who? enriching the social graph through a crowdsourcing game. *SIGCHI conference on human factors in computing systems*, 7–12 May 2011, Vancouver. New York: ACM Press, 1373–1382.

Hacker, S. and von Ahn, L., 2009. Matchin: eliciting user preferences with an online game. *SIGCHI conference on human factors in computing systems*, 4–9 April 2009, Boston. New York: ACM Press, 1207–1216.

Hamari, J. and Eranti, V., 2011. Framework for designing and evaluating game achievements. *5th international digital games research conference: think design play*, 14–17 September 2011, Hilversum, The Netherlands, 122–134.

Ho, C.-J. and Chen, K.-T., 2009. On formal models for social verification. *15th ACM SIGKDD workshop on human computation*, 28 June–1 July 2009, Paris. New York: ACM Press, 62–69.

Ho, C.-J., *et al.*, 2009. KissKissBan: a competitive human computation game for image annotation. *15th ACM SIGKDD workshop on human computation*, 28 June–1 July 2009, Paris. New York: ACM Press, 11–14.

Hong, J.-Y., Suh, E.-H., and Kim, S.-J., 2009. Context-aware systems: a literature review and classification. *Expert Systems with Applications*, 36 (4), 8509–8522.

Huck, S., Kübler, D., and Weibull, J., 2012. Social norms and economic incentives in firms. *Journal of Economic Behavior & Organization*, 83 (2), 173–185.

Ipeirotis, P., 2010. *Demographics of mechanical Turk* [online]. Technical Report, New York University, Working Paper. Available from: http://ssrn.com/abstract=1585030 [Accessed 15 December 2012].

Johnson, D. and Gardner, J., 2010. Personality, motivation and video games. *22nd conference of the computer-human interaction special interest group of Australia on computer-human interaction*, 22–26 November 2012, Brisbane. New York: ACM Press, 276–279.

Keil, D. and Goldin, D., 2005. Indirect interaction in environments for multi-agent systems. *2nd international workshop on environments for multi-agent systems*, 25 July 2012, Utrecht, Berlin. Heidelberg: Springer, 68–87.

Krause, M. and Aras, H., 2009. Playful tagging: folksonomy generation using online games. *18th international conference on world wide web*, 20–24 April 2009, Madrid. New York: ACM Press, 1207–1208.

Krause, M., *et al.*, 2010. Frontiers of a paradigm: exploring human computation with digital games. *16th ACM SIGKDD workshop on human computation*, 25–28 July 2010, Washington, DC. New York: ACM Press, 22–25.

Krause, M. and Smeddinck, J., 2011. Human computation games: a survey. *19th European Signal Processing Conference, EUSIPCO*, 29 August–2 September 2011, Barcelona, Spain. Available from: http://www.eurasip.org/Proceedings/Eusipco/Eusipco2011/papers/1569427719.pdf [Accessed 25 November 2013].

Kuo, Y.-L., *et al.*, 2009. Community-based game design: experiments on social games for commonsense data collection. *15th ACM SIGKDD workshop on human computation*, 28 June–1 July, 2009, Paris. New York: ACM Press, 15–22.

Law, E. and von Ahn, L., 2009. Input-agreement: a new mechanism for collecting data using human computation games. *SIGCHI conference on human factors in computing systems*, 4–9 April 2009, Boston. New York: ACM Press, 1197–1206.

Law, E. and von Ahn, L., 2011. Human computation. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 5 (3), 1–121.

Lee, C.S., *et al.*, 2010. Indagator: investigating perceived gratifications of an application that blends mobile content sharing with gameplay. *Journal of the American Society for Information Science and Technology*, 61 (6), 1244–1257.

Ma, H., *et al.*, 2009. Improving search engines using human computation games. *18th ACM conference on information and knowledge management*, 2–6 November 2009, Hong Kong. New York: ACM Press, 275–284.

Mason, W. and Watts, D., 2009. Financial incentives and the "performance of the crowds". *15th ACM SIGKDD workshop on human computation*, 28 June–1 July 2009, Paris. New York: ACM Press, 77–85.

Matyas, S., *et al.*, 2008. Designing location-based mobile games with a purpose: collecting geospatial data with CityExplorer. *2008 international conference on advances in computer entertainment technology*, 3–5 December 2008, Yokohama. New York: ACM Press, 244–247.

Nov, O., 2007. What motivates Wikipedians? *Communications of the ACM*, 50 (11), 60–64.

Quinn, A.J. and Bederson, B.B., 2011. Human computation: a survey and taxonomy of a growing field. *SIGCHI conference on human factors in computing systems*, 7–12 May 2011, Vancouver. New York: ACM Press, 1403–1412.

Rafelsberger, W. and Scharl, A., 2009. Games with a purpose for social networking platforms. *20th ACM conference on* *hypertext and hypermedia*, 29 June–1 July 2009, Torino. New York: ACM Press, 193–198.

Reeves, S. and Sherwood, S., 2010. Five design challenges for human computation. *6th Nordic conference on human-computer interaction*, 16–20 October 2010, Reykjavik. New York: ACM Press, 383–392.

Rollings, A. and Adams, E., 2003. *Andrew rollings and Ernest Adams on game design*. 1st ed. Indianapolis: New Riders Games.

Ross, J., *et al.*, 2010. Who are the crowdworkers? Shifting demographics in mechanical Turk. *SGICHI conference on human factors in computing systems*, 10–15 April 2010, Atlanta. New York: ACM Press, 2863–2872.

Salen, K. and Zimmerman, E., 2004. *Rules of play*. Cambridge: MIT Press.

Seemakurty, N., *et al.*, 2010. Word sense disambiguation via human computation. *16th ACM SIGKDD workshop on human computation*, 25–28 July 2010, Washington, DC. New York: ACM Press, 60–63.

Seneviratne, L. and Izquierdo, E., 2010. An interactive framework for image annotation through gaming. *International conference on multimedia information retrieval*, 29–31 March 2010, Philadelphia. New York: ACM Press, 517–526.

Sicart, M., 2008. Defining game mechanics. *Game Studies: The International Journal of Computer Game Research*, 8 (2). Available from: http://gamestudies.org/0802/articles/sicart [Accessed 15 December 2012].

Stemler, S., 2001. An overview of content analysis. *Practical Assessment, Research & Evaluation*, 7 (17). Available from: http://PAREonline.net/getvn.asp?v=7&n=17 [Accessed 11 December 2012].

Terry, L., *et al.*, 2009. Harnessing human computation cycles for the FPGA placement problem. *Internal conference on engineering and reconfigurable systems and algorithms*, 13–16 July 2009, Las Vegas. Las Vegas: CSREA Press, 188–194.

Thompson, C., 2007. *What type of game cheater are you?* Available from: http://www.wired.com/gaming/virtualworlds/commentary/games/2007/04/gamesfrontiers_0423 [Accessed 15 December 2012].

Tuite, K., *et al.*, 2010. Reconstructing the world in 3D: bringing games with a purpose outdoors. *16th ACM SIGKDD workshop on human computation*, 25–28 July 2010, Washington, DC. New York: ACM Press, 41–44.

Walsh, G. and Golbeck, J., 2010. Curator: a game with a purpose for collection recommendation. *SIGCHI conference on human factors in computing systems*, 10–15 April 2010, Atlanta, GA. New York: ACM Press, 2079–2082.

Weber, I., Robertson, S., and Vojnovic, M., 2009. Rethinking the ESP game. *SIGCHI conference on human factors in computing systems*, 4–9 April 2009, Boston, MA. New York: ACM Press, 3937–3942.

Weng, L. and Menczer, F., 2010 GiveALink tagging game: an incentive for social annotation. *16th ACM SIGKDD workshop on human computation*, 25–28 July 2010, Washington, DC. New York: ACM Press, 26–29.

Yee, N., 2006. Motivations of play in online games. *CyberPsychology & Behavior*, 9 (6), 772–775.

Yuen, M.-C., Chen, L.-J., and King, I., 2009. A survey of human computation systems. *12th IEEE international conference on computational science and engineering*, 29–31 August 2009, Vancouver. IEEE Computer Society, 723–728.

Zittrain, J., 2008. Ubiquitous human computing. *Philosophical Transition of the Royal Society A*, 366 (1881), 3813–3821.