Game taskification for crowdsourcing

A design framework to integrate tasks into digital games.

Anna Quecke

POLITECNICO
MILANO 1863

SCUOLA DEL DESIGN

Master of Science
Digital and Interaction design

Author
Anna Quecke
895915

Academic year
2019-20

Supervisor
Ilaria Mariani

Table of Contents

List of figures

Abstract

Italiano

I sistemi di crowdsourcing game-based sono una categoria ampia. Ne fanno parte tutti i sistemi che richiedono ai propri utenti di completare una attività per crowdsourcing ed incentivano la partecipazione attraverso giochi e tecniche di game design. La categoria include quindi sia serious games che artefatti gamificati. Questi sistemi cercano di coinvolgere gli utenti con piacevoli esperienze di gioco per sostenere e rafforzare la loro volontà a contribuire.

Costruendo su questo concetto, questo studio indaga l'uso dei giochi e del game design come catalizzatori per stimolare la partecipazione a iniziative di crowdsourcing. Lo studio passa in rassegna un'ampia letteratura interdisciplinare per collegare i temi del gioco e del crowdsourcing e analizza lo stato dell'arte concentrandosi su casi studi rilevanti. Nel mentre è discusso come i sistemi di crowdsourcing game-based siano strumenti ideali per incentivare e sostenere l'innovazione sociale.

In particolare, è apparso interessante e promettente il tema della taskification come area di ricerca, che ancora necessita di indagine scientifica.

La taskification è il processo di integrazione di attività che mirano ad ottenere risultati concreti all'interno di videogiochi. Questo approccio risponde al bisogno del crowdsourcing di coinvolgere un grande numero di participanti. Infatti, i grandi giochi di intrattenimento raggiungono bacini di utenti estremamenti ampi. L'integrazione fluida di attività che mirano ad ottenere risultati concreti all'interno di un gioco può spingere quei bacini di utenti a contribuire completando le richieste del crowdsourcing.

Abstract

Riconoscendo le potenzialità della taskification e la scarsa conoscenza che c'è al riguardo, questa ricerca mira a definire un framework per la progettazione della taskificazione di giochi per applicazioni di crowdsourcing. Sono stati analizzati vari ambiti per sviluppare uno strumento interdisciplinare, poi testato con utenti con strumenti qualitativi. Il contributo finale di questa ricerca è composto dalla discussione del posizionamento di taskification in relazione a gamification e serious games, un insieme revisionato di linee guida per la progettazione di sistemi di crowdsourcing game-based e il framework per la taskificazione di giochi con attività per il crowdsourcing.

English

Game-based crowdsourcing systems are a wide category. It contains all the systems that crowdsource a task and incentivize participation through games and game design, hence including serious games as well as gamified products. Those systems aim to engage users with pleasurable gaming experiences to support and enhance their willingness to contribute.

Building on this, this study enquiries the use of games and game design as catalysts for participation in crowdsourcing projects. It reviews a wide interdisciplinary literature to bridge game and crowdsourcing design, and it analyses the state of the art focusing on relevant case studies. In doing so, it recognizes and discusses that game-based crowdsourcing systems are excellent tools to foster social innovation.

In particular, the topic of taskification emerged as an interesting and promising area of enquiry, still in need of scientific investigation.

Taskification is the process of integrating purposeful activities in entertainment gaming contexts. This approach encounters and is aligned to the need of crowdsourcing to engage a wide number of participants. Indeed, great entertainment games have extremely large user bases. The seamless integration of a purposeful activity into a game can redirect those user bases to participate and contribute to the task.

Recognizing the potentiality of taskification and the lack of knowledge regarding it, this research aims to provide a framework to design game taskification for crowdsourcing. Knowledge from different fields was collected and summarized to return a cross-disciplinary tool. The framework was hence tested with users and data was collected through qualitative methods. The final contribution of this research includes the positioning of taskification with regard to gamification and serious games, a review of guidelines for game-based crowdsourcing systems and the framework to tasking games with crowdsourcing activities.

Introduction

Gamers play in many different situations: to relax at home, to take a break during work, to fill a waiting time, to entertain themselves. Whatever is the reason, they invest a lot of time on this activity and the sum of the hours spent on video games is enormous. Data shows that there are about 2.5 billion gamers worldwide (Narula, 2019), around 30% of the world population, and they spend more than six hours each week playing (The State of Online Gaming 2020, 2020)[1]. Together they spend around 15 billion hours on video games each week.

This huge "workforce" has achieved great objectives in their game worlds. In April 2009, *Halo 3* players spent 565 days fighting the third and final campaign in the fictional Great War to collect over 10 billion kills against their virtual enemy, the Covenant (McGonigal, 2011). This has inspired the idea that gamers could complete objectives of that scale even in the real world.

> Take *World of Warcraft*, for example—the most successful MMORPG ever. Currently, with more than 11.5 million subscribers, each averaging between sixteen and twenty-two hours a week playing the game, that's 210 million participation hours spent weekly on just a single MMORPG. And the number of WoW subscribers is almost exactly the same as the number of registered contributors to Wikipedia.

1 https://it.limelight.com/resources/white-paper/state-of-online-gaming-2020/

> Based on Clay Shirky's estimate that all of Wikipedia took 100 million hours to create, the *WoW* community alone could conceivably create a new Wikipedia every three and a half days.
> (McGonigal, 2011, p. 231)

This thesis arises from this suggestion. It investigates the idea of using games and games design as incentives to involve gamers in prosocial activities through crowdsourcing. This research hence identifies the category of game-based crowdsourcing systems as matter of inquiry and explores them through several points of view.

Chapter one describes the game-based crowdsourcing systems. It introduces the concept, explaining why and how games can support crowdsourcing projects. To do so, it analyzes both crowdsourcing and games from a theoretical point of view, provides their definition and explains their touchpoints. It finally illustrates that game-based crowdsourcing systems may be particularly fit to address social innovation issues. It discusses the impact of such systems on society, covering both the concept of "Game for Impact" and ethical matters. It explains that the major obstacle to the spread of game-based crowdsourcing systems for social innovation is the lack of connection and collaboration between the game industry and research. It hence suggests and envisions that this lack will be addressed and the phenomenon will take off.

Chapter two digs into the way to harness games power. It discusses the concept of "fun" and how to convey it through different methods – by gamifying an activity, developing a serious game or taskifying a game. It also explains that fun is not the only motivator for participants in game-based crowdsourcing systems. It presents the engagement cycle of citizen science volunteers and unveils some critical concepts to understand how users decide to get involved. This investigation points out two main user groups for game-based crowdsourcing systems: volunteers and gamers. Hence, gamers' motivators are explored as well from a psychological point of view. The area of research emerges at the end of the chapter.

Chapter three outlines the research methodology of this thesis. It defines the leading research question and the qualitative methods employed to address it. It then illustrates the whole research process by mapping and explaining it.

Chapter four reports the design and testing of a framework purposed to design game taskification for crowdsourcing. It presents the theoretical foundations of the tool, its design, and the testing phase. The tool validation involved nine game designers who were asked to taskify a game with a citizen science project, relying on the framework for guidance during this process. Data were collected conducting a participant observation of the design process and focus groups with the designers involved.

Chapter five contains the research discussion. It recollects the data and analyzes them systematically to understand if and how the framework supported the design of taskification. Many insights were drawn from the field tests; moreover the feedback reported allowed to review the activities, looking for recurrent patterns and underlying meanings. The chapter also states what this research provides to the knowledge of game-based crowdsourcing and taskification. It summarises clearly the contributions to these topics and traces them in the text. Finally, it outlines possible trajectories for future research, to provoke and keep the discussion on game-based crowdsourcing and taskification going.

1 Moving crowds to achieve valuable social innovation through games

Since the advent of Web 2.0, the internet has become more and more participatory (Jenkins, 2006; Jenkins et al., 2013). Web 2.0 was innovative as it did not have particular technical demands, but it is rather focused on the design and use of websites. In particular, its "architecture of participation" (O'Reilly, 2005) has supported user's online involvement and contribution. The internet became easily shapeable for common people by participating in various activities such as uploading, commenting, voting, tagging and so on (Jenkins et al., 2013). These easy-to-use and captivating designs are what encourages people to get involved and keep them hooked to certain websites.

Among this wide context of innovation that was the establishment of the Web 2.0, the phenomenon of crowdsourcing started to rise at the beginning of 2000s. Briefly described, crowdsourcing is a model in which an organisation outsources assets from an unknown crowd on the web (detailed in par. 1.2). Crowdsourcing is «a type of participative online activity» (Estellés-Arolas and González-Ladrón-de-Guevara, 2012, p. 197) and hence it is clearly one of the outcomes of the peculiar participatory culture of the Web 2.0 (Brabham, 2013). It helps reach to online contributors willing[2] to solve a task, participating in a project without any obligation to do so. Crowdsourcing applications are wide and its usage has spread thanks to its benefits. From the organisation's point of view, crowdsourcing is useful either to harness crowd intelligence to solve complex problems or to reduce costs by distributing the work (Wang et al., 2020).

2 It is important to not misread "willing" in this context. The discussion is referring to the availability of crowdsourcees of undertaking a task, and not their desire to contribute to a specific cause. This thesis do not address crowdsourcees' motivations later in par. 1.2.1, 2.2, 2.2.1 and 2.2.2.

1.1 From participatory culture to citizen science and game-based crowdsourcing systems

Brabham (2013) states that conceptually it is possible to explain the nature of crowdsourcing both through 1) the processes of problem-solving and innovation or 2) the group phenomena of collective intelligence. He argues that crowdsourcing is a problem-solving model because organizations that have a problem and aim at a goal can use crowdsourcing to scale up the task environment, namely the set of features of the physical environment that can affect the possible different ways of solving a problem. At the same time, Brabham suggests that problem-solving and innovation may be synonyms in the field of research and development (R&D) programs or product development. On the other hand, Brabham reminds that crowdsourcing is frequently regarded as part of the studies on collective intelligence; where the latter identifies a «form of universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skills» (Lévy, 1995). A great part of collective-intelligence studies focuses on the concept that Surowiecki (2004) called the "wisdom of the crowd", a phenomenon that describes how groups of people can outperform even the best individuals or experts under specific conditions. Surowiecki states that the wisdom of crowds relies on the aggregation of the individual outputs of the members of the group rather than the averaging of their collective work. Therefore independence of the individuals and diversity of the group are essential, while too much cooperation, communication, and negotiation among members may undermine a crowd's ability to become wise.

As mentioned, crowdsourcing offers also the benefit of reducing costs by distributing labour (Wang et al., 2020). For example, Amazon Mechanical Turk[3] allows crowdsourcers to pay the people even just a few cents to complete a task. Especially in the field of scientific research, where fundings is often low to sustain the activities of researchers, crowdsourcing is a useful model to contain costs. In particular, crowdsourcing has boosted the phenomenon of *citizen science*, a kind of research that enlists the public in collecting and/or processing data as part of a scientific enquiry (Bonney et al., 2009; Silvertown, 2009). Scientists incur often in large amounts of data that require too much time for a single research to analyze and that computers are not able to correctly process either, due to technical limitations. In these cases, scientists can ask for help from common people to address these difficulties.

However, lowering payments is not the only option. Crowdsourcing studies pose particular attention to investigating people's motivations to better compensate them and have discovered a plethora of reasons why people participate in these initiatives. In the case of citizen science, many people participate for the sake of the research and do not demand additional compensation. Among these motivations, it is noteworthy that fun and enjoyment are recurring and important ones (Brabham, 2008; Ye and Kankanhalli, 2017). Although the specific topic will be further discussed in par. 1.2.1, it is hereby necessary to anticipate the role of pleasure in activating people.

Games started to be employed as a support to crowdsourcing systems to further lower the investment of crowdosurcers,

3 https://www.mturk.com/

who did not have to provide even small payments but just fun in exchange to crowdsourcees' performances. Game development has its costs, but games are also to involve and engage crowds and that makes them a valuable option to monetary incentives. This approach has demonstrated its value as it widely spread. The platform Citizen Science Games collects many different projects[4], dating back even to 2008, and shows many more already in the works. Citizen science games are a kind of crowdsourcing game, which build the game upon the task that the researchers propose to the crowd. However, in the field of citizen science, developing full-fledged games is not the only option to harness games' power. Alongside, many citizen science systems are gamified (Skarlatidou et al., 2019), as to say that they mount game mechanics on activities that are other than games (Zichermann and Cunningham, 2011). Broadly speaking, gamification is actually a popular approach in designing crowdsourcing systems (Hamari et al., 2014; Seaborn and Fels, 2015). Gamification, further discussed in par. 2.1.1, is the practice involving game design elements in non-game contexts (Deterding et al., 2011).

Gamified artefacts are fundamentally different from full-fledged games, as it will be explained in parr. 2.1.2 and 2.1.4. However, similarly to developing purposeful games, the aim of gamification is to exploit games and game design to motivate people to pursue a certain objective. Crowdsourcing games and gamified crowdsourcing are henceforth described as game-based crowdsourcing systems, since they are systems that attempt to harness the power of games and game design to ensure crowdsourcing sustainability by delivering fun and enjoyable experiences to their users.

Building on this reasoning, this thesis investigates these systems to understand their peculiar features by analyzing the literature and some outstanding case studies. Particular attention is paid to the their use for social innovation purposes, considering mostly citizen science games but not only.

4 https://citizensciencegames.com/games/

1.2 Analyzing the nature of crowdsourcing

> The term crowdsourcing was quickly adopted by the popular press and bloggers. Suddenly, new media examples that structurally had nothing to do with crowdsourcing—such as Wikipedia, YouTube, Flickr, Second Life, open-source software, and blogs—were all called crowdsourcing. Historical examples (such as the Alkali Prize in the 1700s and the Oxford English Dictionary in the 1800s) and marketing gimmicks (such as DEWmocracy and Mars's contests to choose new colors of M&Ms) were all conflated with the term. Soon anything that involved large groups of people doing anything was called crowdsourcing. Many of these loud but misguided voices—including Forbes, BusinessWeek, and countless social media gurus—spread a confusing message about what exactly crowdsourcing was.
> (Brabham, 2013, p. XVIII, XIX)

Crowdsourcing is a term often misused. Because of its fascinating nature, it has been associated with different fields, products and systems with which it had nothing to spare. According to Brabham (2013), new media technologies have redesigned both human-to-human and human-to-organizations relationships. He admits that underlying long-standing problem-solving and collaboration concepts of crowdsourcing have existed for centuries, but its form relies on this redesign. He argues that crowdsourcing

came into being thanks to the widespread adoption of the internet in the late 1990s, and the following spread of high-speed connectivity and online participatory culture in the 2000s. He states that it has values of its own and should not be downsized as an instance of an online community, a concept that has been around since before the Internet, open-source production, a new word for traditional market research, and more.

Crowdsourcing is a composite word that mixes "crowd" and "outsourcing". It was first used in an article by Jeff Howe in the *Wired journal* (2006). Howe described a trend that suddenly gave hobbyists, part-timers and dabblers the chance to sell their content, creations, even skill, and on the other side companies found cheap or free labor and the latent talent of the crowd. After its introduction, the term became quickly viral. A week later to its inception, searching "crowdsourcing" on Google already produced 180,000 results (Brabham, 2013, p. XVIII).

Later, Estellés-Arolas and González-Ladrón-de-Guevara (2012) collected all the definitions in the scholarly literature to settle a clear and shared meaning of the term. Their research brought to this definition:

> Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken.
> (Estellés-Arolas and González-Ladrón-de-Guevara, 2012, p. 197)

This definition emphasises eight crucial interconnected aspects of crowdsourcing. There are two clearly recognizable actors, **1) the crowd** and **2) the crowdsourcer**, who interact through **3) the internet**. The crowdsourcer sets **4) an open call** to assign **5) a task with a clear goal** to the crowd, establishing **6) an online assigned process of participative type**. Both crowdsourcer and crowd are meant to receive defined benefit from the system, either **7) some kind of compensation** for the former or **8) a solution to the task** for the latter.

Brabham (2013) considers Estellés-Arolas and González-Ladrón-de-Guevara's definition quite wordy but complete. He stresses the relevance of the concept of "mutual benefits" in particular. According to Brabham, the locus of control in crowdsourcing must reside between the organization and the community to maximize both the benefits of top-down, traditional management and those of bottom-up, open creative production (Brabham, 2013, p. 4). He argues this sharing of power distinguishes crowdsourcing from similar creative processes.

For example, Wikipedia and open-source software projects focus their locus of control in the community, so their structure is strictly bottom-up. Thus, none of them is technically crowdsourcing. On the other hand, marketing contests such as "choose the next flavour" ones are neither crowdsourcing. They set the locus of control primarily within the organization and do not employ the community's talents or labour.

Instead, crowdsourcing initiatives combine both approaches. E.g., *Threadless*[5] sells clothing designed by the crowd (bottom-up) but it has its contest rules and requirements for submission (top-down). *InnoCentive*[6] enables companies to propose scientific challenges that have clear rules and solution parameters (top-down) to the crowd that can creatively solve them (bottom-up). *Amazon Mechanical Turk*[7] lets the crowd respond to open tasks (bottom-up) that serve the specific requester's needs (top-down).

5 https://www.threadless.com/
6 https://www.innocentive.com/
7 https://www.mturk.com/

Crowdsourcing is a model that blends a bottom-up, open, creative process with top-down organizational goals and ensures a mutually beneficial outcome. In the aforementioned examples, the benefits can be revenues for the service and rewards for the designers (*Threadless*), patents or similars for the company and a compensation for the solver (*InnoCentive*), a completed task for the requester and a small amount of money for the workers (*Amazon Mechanical Turk*). This symbiotic relationship needs the efforts of both parties to flourish. Although crowdsourcer's gains are usually evident, crowdsourcees' ones are not always manifest as well. For this reason, the next paragraph will discuss what encourages participation in crowdsourcing projects from a psychological point of view and it introduces a relevant concept for this thesis: the relevance of fun and enjoyment in crowdsourcing.

1.2.1 The relevance of fun and enjoyment in crowdsourcing

The definition from par. 1.2 clarifies that crowdsourcing «always entails mutual benefit» (Estellés-Arolas and González-Ladrón-de-Guevara, 2012). As for volunteering, crowdsourcing appears to be free labour when it doesn't provide an economic income, but people attending it have some kind of return anyway. This return, however, is what motivates people to participate. As crowdsourcing depends on the achievement of a critical mass and the maintenance of sufficient rates of participation to obtain results (Morschheuser et al., 2017), it is crucial to know why people would be willing to contribute and how to stimulate them.

A famous study on the composition of the crowd at iStockphoto discusses these topics. iStockphoto is a platform that gathers a community of amateur photographers — "iStockers" — who upload their stock images, animations, and video clips to the Web site to be sold. Anyone can download any content they want from the platform, and profits are divided between the photographers and iStockphoto. Brabham (2008) demonstrates how the opportunity to earn money is the most popular reason to be part of the community (89.8%), but not the only one. People reported thinking of iStockphoto (Fig. 1) as an aid to improve their skills (79.1%), a creative outlet (76.9%), and also as a fun activity (71.9%).

Table 1: Percent of iStockers responding "yes" to statements on the question about why they participate at iStockphoto.

Replies	Percentage
The opportunity to make money	89.8
It helps me improve my photography, video, and/or illustration skills	79.1
It is a creative outlet for me	76.9
It is fun	71.9
The opportunity to earn a reputation as a good photographer/artist	49.8
I can produce photos, videos, and illustrations that I like	48.2
It is a better way to make stock photography, video, and illustrations	38.3
To build a network with other creative people	36.7
It passes the time when I am bored	21.2
To build a network of friends	16.6

Fig. 1 Results from the investigation of iStockers' motivations to participate (Brabham, 2008).

Other than Brabham, several other researchers have investigated the motivation effective in pushing people to contribute. Various theories have been applied to understand crowdsourcees' behaviour. Ye and Kankanhalli (2017) have summed these researchers (Fig. 2).

The theories shed light on continuance intention (Sun et al., 2012, 2011), participation intention (Zheng et al., 2011), content-adding (Geri et al., 2017) and level of activity (Boons et al., 2015). of crowdsourcees. Enjoyment, either hedonic value which generates satisfaction or intrinsic motivation, affect both continuance and participation intention. Monetary reward, recognition, prestige, reciprocity and pride are able to influence crowdsource-

Table 1
Theories used to explain solvers' participation in crowdsourcing.

Theories	Proposition	Results	Reference
Value theory	People are motivated to act by judging the value of the action. There are two types of value, i.e., utilitarian and hedonic (enjoyment)	Hedonic value → satisfaction; Hedonic value, satisfaction → continuance intention	Sun et al. (2011)
Value expectancy theory	Individuals' actions are related to their *subjective value* of behavioral outcomes and the *expectancy or probability* to conduct the behavior successfully and achieve outcomes. As expectancy increases, effect of value on behavioral intention increases	Extrinsic motivation (monetary reward), intrinsic motivation (enjoyment) → continuance intention	Sun et al. (2012)
Motivation theory	People are driven to achieve their goals by their motivations. Motivation can be intrinsic and extrinsic	Extrinsic motivation to gain recognition, intrinsic motivation (enjoyment) → participation intention	Zheng et al. (2011)
		Awareness of rewards, prestige, reciprocity → content-adding	Geri et al. (2017)
Social identity theory	Individuals' identity is based on their group membership. The group which an individual belongs to is an important source of pride and self-esteem	Pride → level of member activity	Boons et al. (2015)
Value sensitive design theory	Human values should be accounted for in the design of computer technologies	Nine values shared by solvers on Amazon's Mechanical Turk are uncovered	Deng et al. (2016)

Fig. 2 Ye and Kankanhalli's sum of theories applied to crowdsourcing platforms to explain users participation (2017).

es' participation and behaviour as well. Finally, an interpretative field study by Deng and colleagues (2016) unveils nine shared values of crowdsourcees on Amazon's Mechanical Turk:

1. access: crowdsourcing allows those who are unable to conform to traditional workplace (invalids) or are going through a period of unemployment to make incomes anyway;
2. autonomy: crowdsourcing provides flexibility and freedom in making job-related decision (what, how, when, where);
3. fairness: crowdsourcees express need for fair payments, evaluations and feedbacks;
4. transparency: clear brief job description, instructions, time requirements, and payment amount are essential as well as transparent process and feedback job performance;
5. communication: direct and open communication with job requesters helps clarify a task and get feedback about their outcomes, reducing potential rejections and disputes;
6. security: assurance, safety, low work disruptions and task scamming are all desirables features of a desirable crowdsourcing environment;
7. accountability: actions of people or institutions should be traced uniquely to individual workers and job requesters so that they be held responsible for their work and behaviour, in particular when they result unethical;
8. making an impact: crowdsourcees desire to contribute to the community and have a positive impact on other people's lives or academic researches;
9. dignity: crowdsourcees value a sense of pride and respect.

Ye and Kankanhalli (2017) decide to expand this knowledge on crowdsourcing participation through social exchange theory. Social exchange theory posits that individuals act to maximize benefits and minimize costs in social exchanges and accept the exchange only when they perceive rewards exceed costs. People expect some sort of future return to their favours. However, they do not have a clear idea of what to expect, because there are no explicit rules or agreement. This means people simply believe there will be some return and justify the exchange based on that. Therefore, social exchanges are somehow bound to long-term relationships of interest, differently to one-off exchanges.

Ye and Kankanhalli's intention was to complement benefits with costs and include trust effects in a comprehensive model. They illustrate benefits as monetary reward, skill enhancement, peer reputation, enjoyment and work autonomy. Alongside, they determine cognitive effort for knowledge contributors and loss of knowledge power as two possible costs that hold people from participating in crowdsourcing. Moreover, they expect monetary reward, cognitive effort and loss of knowledge power to influence trust, which in turn may affect participation. Fig. 3 shows Ye' and Kankanhalli's model with the expected correlations, both positives and negatives, among all the elements.

Ye and Kankanhalli gather data surveying users on the large Chinese crowdsourcing site TaskCN to test the model. The results confirm most of the expected correlation. Skill enhancement, enjoyment and work autonomy affect positively participation in crowdsourcing. Monetary reward and loss of knowledge power influence trust, the former positively, the latter negatively, and trust impacts positively participation. Monetary reward directly affects participation as well, while loss of knowledge power does only indirectly through trust. In the same way of loss of knowledge power, peer reputation shows no effect on participation. Cognitive effort affects participation negatively, but was unrelated to trust.

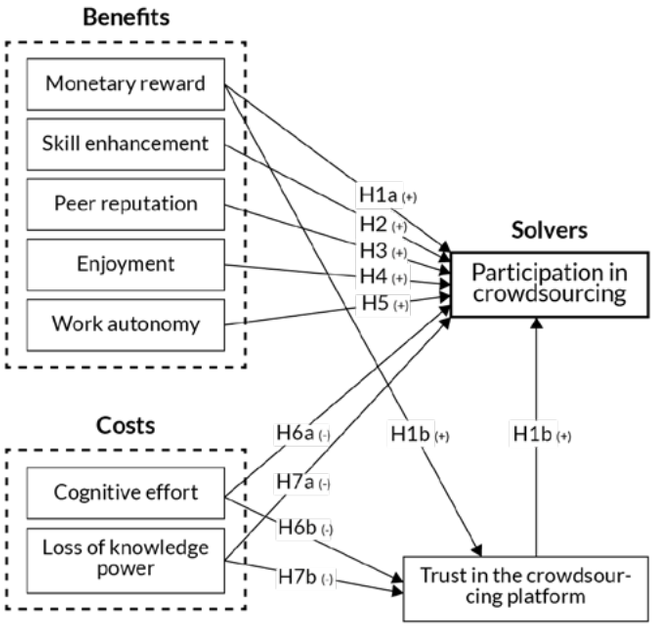


Fig. 3 The research model for solvers' participation from Ye and Kankanhalli (2017). H3, H6b, and H7a are not supported.

Ye and Kankanhalli argue peer reputation was not significant due to the competition-based model of TaskCN. They discover that since peers learn from others' successful contributions, a higher peer reputation comes with visibility that can attract others to imitate those contributions. Crowdsourcees have no interest in the competitors to improve their contribution through imitating another successful one as they would "steal" them the monetary reward. Therefore they prefer avoiding visibility. Regarding loss of knowledge power not affecting participation, Ye and Kankanhalli explain it might be due to the presence of plagiarism detection mechanisms in TaskCN. Crowdsourcees tend to trust the platform more because the mechanism decreases their perceived loss of knowledge power since nobody can copy their contributions. Lastly, cognitive effort might not influence trust when it is perceived more as a factor related to the self rather than to the environment (the crowdsourcing platform contextually).

Ye and Kankanhalli' study assesses previous knowledge and draws a comprehensive model of what drives to or retain from crowdsourcees' participation. To assure participation, costs and benefits need to be carefully balanced to make the second appear grater. As explained in social exchange theory, it is necessary that perceived rewards exceed costs for the exchange to occur. Careful design can aim to push people accepting the exchange. One way is minimizing the cost, the other is enhancing benefits, always accounting users' perception and trust.

Enjoyment and fun are clearly relevant motives to participate and contribute since they appear as a recurrent theme across many studies (Brabham, 2008; Sun et al., 2012, 2011; Ye and Kankanhalli, 2017; Zheng et al., 2011). It is no surprise many researchers and practitioners focus their attention on the possibility that fun and enjoyment could become the main bargain chip to catch people's interest.

1.2.2 The rise of Games with a Purpose

Although many researches have demonstrated the value of fun in crowdsourcing, the idea of using it to harness the power of the crowd was actually tested way before. Indeed, it was the early 2000s when Luis von Ahn conceived the *ESP Game* at Carnegie Mellon University, the first Game With A Purpose, GWAP in short. GAWPs are applications of Human-Based Computation (HBC), that is «a paradigm for utilizing human processing power to solve problems that computers cannot yet solve» (von Ahn, 2005), that engage participants through games.

As Von Ahn's GWAPs gained fame, they became the best practice to look at when developing game-based crowdsourcing. Though they are surely a great example of what games could achieve when they access a huge crowd, GWAPs are not the only kind of games that have exploited human capabilities to reach a goal. Nevertheless, it has become an umbrella term for some (Schrier, 2016), even if it was actually thought to define a really narrow niche.

The concept of HBC does not imply crowdsourcing, even though the two terms share some spirit in exploiting human skills to complete tasks. Actually, it is not unlikely that a system can employ both HBC and crowdsourcing, which causes confusion between the two terms. While they may overlap at some extent, they are two different processes.

The main difference lies in what is replaced in the first place: indeed, whereas HBC substitutes computers for humans, crowdsourcing substitutes traditional human workers for members of the public (Quinn and Bederson, 2011). Therefore, an application combines HBC and crowdsourcing if it replaces computer labour outsourcing its task to a crowd. However, the two terms define different paradigms that do not necessarily have to overlap, so an HBC application isn't always a crowdsourcing application too, and vice versa.

It is crucial that the difference between the two systems is shown and clear. GWAPs should always be associated with a computational problem, thus generating input-output behaviour (von Ahn and Dabbish, 2008). On the other hand, crowdsourcing can harness phenomena as serendipity and the wisdom of crowds that go beyond an input-output behaviour (Geiger and Schader, 2014). Researchers and practitioners need to acknowledge these facets to handle the matter properly for their objectives. Navigating these nuances and distinguishing them can improve the awareness, therefore the effectiveness, of all the possible design choices undertaken during the development of a game-based crowdsourcing system.

The next paragraph describes the *ESP Game*, the first game designed to engage common people in solving a computational problem and milestone for the definition of the category of game-based crowdsourcing.

ESP Game / Google Image Labeler

Back at the beginning of the century, web images labelling presented a major technological challenge. Despite several applications, like image search engines, requiring accurate descriptions, the techniques of that time could not provide sufficient performance. Computer vision was not precise enough, and common techniques usually paired the images to the text appearing in their same web page, which is often scarce and can be misleading or hard to process. Therefore, the best choice was to rely on human manual labelling, which clearly is a demanding activity. Acknowledging this, von Ahn and his team developed the *ESP Game* to stimulate people to do the "dirty work". The idea was to persuade people to create accurate descriptions merging this task in a pleasurable gaming experience.

The *ESP Game* was a cooperative game in which two players have to "think like each other" without the possibility to communicate to reach a higher score possible in 2.5 minutes. The game shows both players the same image and asks them to describe it with a word: the moment they choose the same word to describe the subject, the game assigned them points and displayed a new image, and so on until players run out of time. Indeed, ESP stands for "Extra Sensory Perception" as the game asked the players to "perceive" the word that the other player has chosen to describe the image.

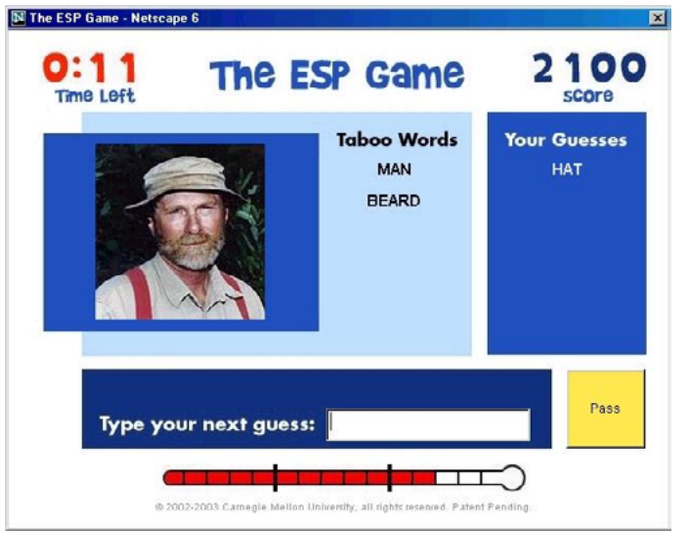


Fig. 4 The interface of the ESP Game. *The thermometer at the bottom measures how many images partners have agreed on. When filled, players gain a bonus point.*

To make the game harder, each image displays eventually up to six different words that players were not allowed to enter, called taboo words, which were previous players' guesses. A thermometer at the bottom of the screen kept track of the number of images they agreed on, upon which was calculated the score. The thermometer was designed to reinforce players' feeling of incremental success and, as well as the score displayed, established a clear relation between action and feedback which encouraged them to continue playing (von Ahn and Dabbish, 2004). Players could opt-out or skip difficult images if both agree on skipping it. Repeated passing on the same image by different players notifies the system that an image has too many taboo words or it is too complex to be used in the game. In both cases, players were unable to find a common word so the image was removed from the game database. Those images were reinserted in the game after some time because their meaning may have changed and players could label them.

To avoid communication, the two players were randomly paired. Communication was a threat to good data retrieval as it left room for cheating: indeed, players could agree on using the same word for describing every image in the session, thus making the results useless as labels. Another threat to data accuracy was misspelling. To avoid it, the game checks players' submission through a 73,000-word English dictionary and notifies errors turning the misspelt word in yellow.

Eventually, players could be paired with bots if the number of people playing is odd. Bots were actually a recorded set of actions from an earlier game session involving two people. Those gameplays were not useless to the overall collection of labels. If the bot and the player agree on a word that was already registered in the database, the word proves to be a good label. Moreover, if the player chooses a word that matches with the registered ones that the bot reproduces, then a new label would be added to the database.

Though players could not notice it, the project is mostly aimed at identifying if the guesses entered are good labels for the images. The team checks the quality of the labels generated by participants through a series of evaluations, namely:

- the precision of game-generated labels in search queries on 10 labels;
- comparison between game-generated labels and labels generated by experimental participants on 20 images;
- experimental participants evaluation of game-generated labels on 20 images.

Fig. 5 The home page of gwap.com. A brief overview of the motive of the games is displayed: "You play the games, computers get smarter, everyone benefits!"

Fig. 6 A screenshot from the ESP Game *on gwap.com. The game was polished and designed according to the whole website image.*

Those tests result in extremely high precision of searching, with 83% of correspondence between gamers and experimental participants' labels and 85% of labels valued as useful by experimental participants (von Ahn and Dabbish, 2004).

The game ran over the Web for four months equipped with just 350,000 images, 293,760 of which were labelled with 1,271,451 words. Meanwhile, the game fun was estimated according to the time people (13,630 users overall) spend on the game. Over 80% of the them played on more than one occasion and, among them, 33 users played more than 50 hours (von Ahn and Dabbish, 2004).

The development of the *ESP Game* pushed von Ahn building other systems that through human play were able to process inputs to obtain outputs, which ultimately led him to conceptualize the broader concept of "Games with a Purpose". Games with a Purpose, shorten GWAPs, describe systems in which people perform tasks that computers are unable to perform, through what they perceive merely as games (von Ahn and Dabbish, 2008). He collected those games on gwap.com (Fig. 5), a web platform where people could register themselves and help compute through many different games.

Von Ahn e Dabbish (2004) reported that labelling all images indexed by Google back then would require just some weeks if 5,000 people were playing the game for 24 hours a day. They stressed that, at the time, most popular games on the Web could have more than 5,000 players at once. If the *ESP Game* had succeeded becoming as popular, it would have reasonably achieved enough players to accomplish its objective, i.e. label all the images on Google (von Ahn and Dabbish, 2004). After a couple of years, the game reached 1 million registered users and though only a fifth of them played it regularly (Saini, 2008), it had already produced 50 million labels through their contributions (von Ahn and Dabbish, 2008).

Significantly, the industry valued this research and its possible real-world impact. Well-labelled images could improve many technologies, like an image search engine. A service enhancement could have raised the number of users that employ it, hence increasing the revenues of the service provider itself. Indeed, Google itself bought a licence from von Ahn to create its own version of the

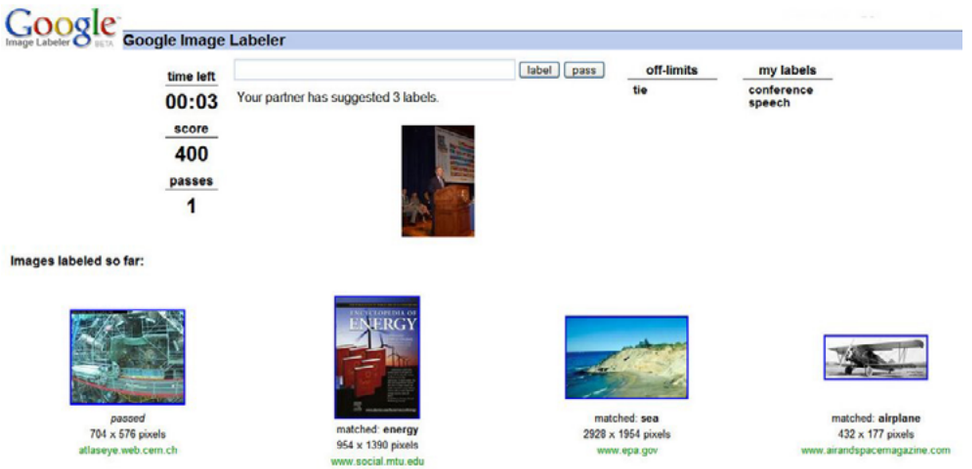


Fig. 7 The Google Image Labeler*: it had all the features of the* ESP Game *plus the pass counter and the list of the images labelled during the session.*

game. The company launched *Google Image Labeler* in 2006 and it helped improve search results for online images until 2011 when it was shut down due to internal reorganization (Eustace, 2011).

It was accessible again from 2016 through *Crowdsource*, a crowdsourcing platform developed by Google, though it deeply changed. Indeed, the current *Image Labeler*[8] is no longer a game: it lets users choose a category (i.e. "birds") and asks if the image displayed contains that element ("the image shows birds?"), without any competition involved.

8 https://crowdsource.google.com/imagelabeler/category

1.3 Games as productive systems

Implementing games appears to be the natural option to boost fun and enjoyment, which in turn stimulates participation in crowdsourcing activities. On the other hand, games do not seem to be productive systems but leisure products which distract people from real work (McGonigal, 2011). They may appear as improper to match crowdsourcing, which is a productive method to outsource resources and rely on crowdsourcees' work. To clarify this issue, it is meaningful to discuss the definition of *game* and investigate its intrinsic features. It would be easy to define a game as a tool used to play; where *play* identifies the activity of playing a game. However, this sentence would raise questions about what "play" means. To understand what is a game, it is needed to define the activity of interacting with it first. Hence, the paragraph starts discussing the definition of play to end with that of game.

In his essay *Homo Ludens*, Dutch anthropologist and historian Huizinga thoughtfully investigates the nature and significance of what is "play" and how it is intertwined with everyday life. He interprets human culture in relation to play and affirms the concept of "homo ludens" to assess the relevance of play in cultural and social human development. Indeed, he considers games «a necessity both for the individual–as a life function–and for society by reason of the meaning it contains, its significance, its expressive value, its spiritual and social associations, in short, as a culture function» (Huizinga, 1980 [1938], p. 9). Huizinga does not provide an explicit definition of play, but he sums its formal characteristics to identify it. He affirms:

> Summing up the formal characteristics of play we might call it a free activity standing quite consciously outside "ordinary" life as being "not serious", but at the same time absorbing the player intensely and utterly. It is an activity connected with no material interest, and no profit can be gained by it. It proceeds within its own proper boundaries of time and space according to fixed rules and in an orderly manner. It promotes the formation of social groupings which tend to surround themselves with secrecy and to stress their difference from the common world by disguise or other means.
>
> (Huizinga, 1980, p. 13)

This statement reveals main features of play, which may be summed in seven points:

1. it is a free and deliberate activity;
2. it stands outside "ordinary" life;
3. it is aware of standing outside "ordinary" life;
4. it absorbs the player intensely and utterly;
5. it is not connected with any material interest or profit;
6. it has own proper boundaries of time and space;
7. it follows fixed rules.

French sociologist Caillois revises and expands these features in his famous *Man, Play and Games* (2001). Caillois opens his book praising Huizinga's work and his remarkable achievement, i.e. the relevance of play in culture development itself. He values Huizinga's work and relies deeply on it, following his choice to describe play through features rather than outlining a definition and borrowing most of the features presented from his essay. Caillois establishes six core characteristics of play:

1. free;
2. separate;
3. uncertain;
4. unproductive;

5. governed by rules;
6. make-believe.

Play is free as it is not obligatory. Forcing play makes it lose its attractive and joyous quality. It is separate in a sense that it is limited in a pre-defined space and time. It is uncertain because neither the course or the result should be determined beforehand and the players' initiative should be able to influence the play. It is unproductive because it does not create goods, wealth nor new elements of any kind, and even when there is an exchange of property among the players, the end situation is identical to that prevailing at the beginning of the game. Play is governed by rules that set aside the ordinary laws and conventions and become the only ones that count during the activity. Finally, it is make-believe in a sense that players are aware that it is a kind of second reality or free unreality which appears opposed to real life.

It is remarkable to notice that both authors consider unproductivity of play as a major feature. Games are usually conceived as a leisure activity and hence it is in contrast with working time. However, game researcher and designer McGonigal counters this common belief in her book *Reality is Broken*. She affirms that games are actually well-designed, satisfying works. She states: «Compared with games, reality is unproductive. Games give us clearer missions and more satisfying, hands-on work» (McGonigal, 2011, p. 55). According to her, work is satisfying when it presents us with both clear, immediately actionable goals and direct, vivid feedback. Good games display clear goals with actionable steps, they boost people's motivation and assure their progress. They visualize the results of players' actions, manifesting them the power of their capability and their agency. To her, games can be associated with work as well but delivered to users differently, so that they are actually even more productive. McGonigal counters the idea that games are the opponents of work and even disclaims Huizinga and Caillois' concept of game as unproductive in terms of material interest or profit. She collects several case studies in her book that address real-world issues. She affirms games stems such gratification that is an infinitely renewable resource to sustain crowdsourcing initiatives. Moreover, games drive players' experience all along. They are ambitious and set awe-inspiring goals that appear to be impossible to accomplish. However, games help players to achieve them together, concentrating their efforts and improving their collaborative skills[9] (McGonigal, 2011).

McGonigal argues that games should not be viewed as separate or a distraction from real lives and work which leads to human civilisation decline. On the contrary, games fill people's lives with positive emotions, activities, experiences and strengths which might be directed to catalyze human reinvention. In her book, she describes crowdsourcing and social participation games as practical examples of how games can lead to these changes. She does not regard games as intrinsically unproductive as Huizinga and Caillois did, which is why she does not rely on their definitions on games, but rather on American philosopher Suits' one. In his book *The Grasshopper: Games, Life and Utopia*, the author states that «playing a game is the voluntary attempt to overcome unnecessary obstacles» (Suits, 2005, p. 41). This definition has become fundamental in Game Studies and is usually considered the baseline definition of play. It is highly interesting the focus of this definition. Suits stresses two core elements: "voluntary attempt" and "unnecessary obstacles". In line with Huizinga and Caillois' ideas, he regards play as a free and deliberate activity. Players voluntarily submit to the rules of a game in order to play. Those rules create unnecessary obstacles to augment the challenge of the task required by the game, and therefore the enjoyment which bears its accomplishment. Suits clearly explains his point of view discussing the example of golf. Golfers have to get the ball into a set of small holes very far from them, hitting the ball with a bat and with a lower number of shots of their opponents in order to win. Removing all the unnecessary obstacles of golf, the activity would result in the golfer putting the ball into the hole by hand.

9 These concepts are explained in chapters 3, 11, 12 and 13 of McGonigal, J., 2011. *Reality is broken: why games make us better and how they can change the world.* Penguin Group, New York.

In their *Rules of Play*, Salen and Zimmerman (2004) argue that, although insightful, Suits' definition does not actually refer to game but to the act of playing a game. Hence, they collect several definitions from the literature and elaborate them to produce theirs. They state that «a game is a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome» (Salen and Zimmerman, 2004, p. 80). The authors stress that their definition relies on six core elements:

1. system;
2. players;
3. artificial;
4. conflict;
5. rules;
6. quantifiable outcome.

According to Salen and Zimmerman, games are intrinsically systemic and can be understood as systems. They define a system as a set of parts that interrelate to form a complex whole. The complex whole that is a game provides contexts for interaction through spaces, objects, and behaviours. Players interact with the system of a game by exploring, manipulating and inhabiting it in order to experience the play of the game. A game is meant to be played actively by its participants, who are essential to it. They immerse themselves in the artificial time and space of the game that is clearly separated from "real life", despite the fact that games happen in the real world. In this artificial environment, they have to deal with a contest of powers, a conflict. It may take many forms: cooperation, competition, solo or multiplayer, all games contain a kind of conflict and it is central to them. Players interact accordingly to the rules of the game, which restrict players' actions and provide the structure out of which play emerges. Finally, players end the game either winning or losing or receiving some kind of numerical score: in short, they get some kind of quantifiable outcome. Games stand out from other less formal play activities because of this quantifiable outcome.

Salen and Zimmerman's definition appear to be quite handy for this research. Consider games as systems helps understand how crowdsourcing can be integrated with them. Systems correlate many elements and can be expanded to add new elements or modify afterwards. They are flexible and adaptive. DLCs and mods exemplify this feature of game clearly. A DLC, short for downloadable content, is extra content provided by the game publisher and downloadable from the Internet, released for an already released video game. It can be cosmetic content (skins) or new in-game content (characters, levels, modes, ...) or a mix of both that offers a continuation of the base game. A mod, short for "modification", is an alteration of a video game performed usually by its players or fans that modify one or more aspects of it. The changes may be small tweaks or complete overhauls, affect only the game's look or even behaviours. Replay value and interest of the game can both benefit from mods.

Games are so flexible that it is easy to adjust their system or expand them. Their flexibility allows integrating crowdsourcing as well. Moreover, as explained, their usage does not undermine people productivity, as games are actually a form of work but designed to provide actionable goals and evident feedback to push people overcoming unnecessary obstacles. Researchers have demonstrated that crowdsourcees' work is actually sustained and not negatively affected by the presence of playful or gameful interaction. Next paragraph discusses this further.

1.3.1 Evidence of the positive effects of gamifying a crowdsourcing system

It has been discussed how games are systems which can be easily adapted even to become productive. However, par. 1.3 coincides the discussion from a theoretical point of view. This paragraph describes how those concepts reveal to actually have roots in real life as well.

Game-based crowdsourcing systems show to have the power to reach many people and achieve the critical mass needed for their specific purpose. The *ESP Game* reached 13,630 players on a four-month period (von Ahn and Dabbish, 2004), the Guardian's

Investigate Your MP's Expenses reached 26,774 players on a one-year period (Flew et al., 2010), *Foldit* reached about 57,000 players on a two-year (Cooper et al., 2010a), and *Project Discovery* reached even 322,006 players on a three-year period (Sullivan et al., 2018). These games are notable examples of game-based crowdsourcing and are further explained in par. 2.1.1, 2.1.2 and 2.1.3. However, it may be questioned that these effects are not accountable to the game-like features and aspects.

Gamification studies are deeply connected to behaviouralism studies and so they have investigated how gameful experiences may affect participation and others behaviours in crowdsourcing. Morschheuser and colleagues' (2017) review on gamification application in crowdsourcing domain discusses the proven effects that it provokes by summarizing many studies on the topic.

Based on Morschheuser and colleagues' (2017) review on gamification application in crowdsourcing domain, it is discussed how gamification is a worthy strategy for crowdsourcing.

Gamification is defined by Deterding as «the use of game design elements in non-game contexts» (Deterding et al., 2011, p. 9). However, Morschheuser and colleagues deepen the reasoning with a behavioural perspective, and describe it as a «design that seeks to, first, increase the motivation of users or participants to engage in an activity or behavior and, second, to increase or otherwise change a given behavior» (2017, p. 27). Therefore, in the context of crowdsourcing, they see it as an attempt to turn crowdsourcees' motivations from being purely rational gain-seeking into becoming self-purposeful and intrinsically driven. The use of gamification in crowdsourcing is further explored in par. 2.1.1, 2.2.1 and 2.2.2.

Morschheuser and colleagues examine empirical papers which analyze psychological and behavioural outcomes. Gamification affects motivation, attitudes, fun/enjoyment, engagement, and even more. Fig. 8 shows the papers classified upon the kind of psychological outcome that they investigate and confirm.

Studies on psychological outcomes use mostly simple questionnaires or qualitative observations, or even the observations of how participants behaved were used as a proxy for psychological

Psychological outcome	Literature	#
Motivation	(Altmeyer et al., 2016; Bowser et al., 2013; Eickhoff et al., 2012; Itoko et al., 2014; Kawajiri et al., 2014; Kobayashi et al., 2015; Liu, et al., 2011a, 2011b; Machnik et al., 2015; Massung et al., 2013; Nose and Hishiyama, 2013; Preist et al., 2014; Prestopnik and Tang, 2015; Roengsamut et al., 2015; Runge et al., 2015; Tinati et al., 2016)	15
Attitudes	(Bowser et al., 2013; Dergousoff and Mandryk, 2015; Itoko et al., 2014; Kobayashi et al., 2015; Martella et al., 2015; Preist et al., 2014; Prestopnik and Tang, 2015; Roengsamut et al., 2015; Runge et al., 2015; Tinati et al., 2016)	10
Fun/Enjoyment	(Altmeyer et al., 2016; Bowser et al., 2013; Choi et al., 2014; Dumitrache et al., 2013; Kobayashi et al., 2015; Lee et al., 2013a; Melenhorst et al., 2015; Prandi et al., 2016; Prestopnik and Tang, 2015; Roengsamut et al., 2015; Runge et al., 2015; Sheng, 2013; Tinati et al., 2016)	13
Engagement	(Altmeyer et al., 2016; Bowser et al., 2013; Liu, et al., 2011a, 2011b; Snijders et al., 2015)	4
Other (e.g. appeal, interest, immersion)	(Cucari et al., 2016; Kobayashi et al., 2015; Melenhorst et al., 2015; Prestopnik and Tang, 2015);	4

References in bold refer to studies in which empirical results about gamification have been reported.

Fig. 8 Psychological outcomes reported in the literature reviewed by Morschheuser and colleagues (2017).

aspects. Morschheuser and colleagues report that only four studies employ validated psychometric measurement instruments.

Morschheuser and colleagues affirm that several studies (see Fig. 9) that compared gamified and non-gamified approaches report that gamification leads to positive improvements, such as increases in (long-term) participation, output quality and reduction in cheating compared to traditional paid crowdsourcing. Only three studies reported more negative effects than positive.

However, Morschheuser and colleagues argue that gamification is not an approach which ensures greater participation in any case. A study shows very small differences between a group tested on gamification and a control group without gamification. Another discusses that simple gamification approaches like points and le-

aderboards are not able to replace financial incentives. Considerations like those remind that gamification is a difficult approach which needs to be designed carefully and, in particular, it is not unfailing.

Ten studies report positive results which rely on users' perceptions of the gamified crowdsourcing system or their engagement. Morschheuser and colleagues stress that although these studies do not show the effects of gamification per se, they represent positive indicators for the acceptance of gamification in the context of crowdsourcing.

Considering all the studies in Fig. 9, more than 90% of them report positive or mainly positive outcomes of gamification in crowdsourcing. Most cases reported positive effects on quantitative contributions, as shown in Fig. 10, but also that positive effects on qualitative contribution and long-term engagement are achievable. Indeed, various studies indicate that gamification affected these parameters as well, but they strongly depend on the context and concrete implementation of gamification affordances.

All this considered, it is clear that gamification is able to have desirable effects in the context of crowdsourcing. It can increase (long-term) participation, engagement, output quality, reduce cheating behaviours and is well-accepted.

Nevertheless, gamification is not the only solution to develop a game-based system. Crowdsourcing systems can exploit the positive effects of games on users in various ways. It is possible to mix games and crowdsourcing with various approaches and obtain different outcomes. It is important to know these approaches and choose the proper one to reach one's specific aim. Par. 2.1 and its subparagraphs explains this in detail.

Results	Compared a gamified approach with a non-gamified one	No comparison (interviews, user feedback, perceptions, time series analysis, influence of context factors)	Comparisons between different gamification designs	#
Quantitative *-inferential*	(Eickhoff et al., 2012; Nose and Hishiyama, 2013; Dergousoff and Mandryk, 2015)	(Melenhorst et al., 2015)	(Choi et al., 2014; Ipeirotis and Gabrilovich, 2014; Lee et al., 2013b; Runge et al., 2015)	8
Quantitative *-descriptive*	(Carlier et al., 2016*; De Franga et al., 2015; Dumitrache et al., 2013*; Kobayashi et al., 2015; Liu, et al., 2011a, 2011b; Simões and De Amicis, 2016; Sørensen et al., 2016; Talasila et al., 2016)	(Pothineni et al., 2014; Roengsamut et al., 2015)	(Feyisetan et al., 2015; Packham and Suleman, 2015*)	12
Qualitative	(Kacorri et al., 2015; Martella et al., 2015)	(Machnik et al., 2015; Saito et al., 2014; Tinati et al., 2016)	(Preist et al., 2014; Prestopnik and Tang, 2015)	7
Mixed-*inferential*	(Altmeyer et al., 2016; Vasilescu et al., 2014)	(Bowser et al., 2013; Itoko et al., 2014)	(Kawajiri et al., 2014; Massung et al., 2013; Prandi et al., 2016)	7
Mixed-*descriptive*	(Goncalves et al., 2014)	(Lee et al., 2013a; Snijders et al., 2015)		3
Total	More positive (14) / negative (2)	More positive (10)	More positive (10) / negative (1)	37

* Studies that reported negative effects of gamification, for instance compared to paid crowdsourcing or non-gamified approaches

Fig. 9 Results on gamified crowdsourcing from Morschheuser and colleagues' (2017) review.

Outcomes	Literature	#
Positive effects on the quantitative contribution / willingness to contribute	(Altmeyer et al., 2016; Bowser et al., 2013; De Franga et al., 2015; Dergousoff and Mandryk, 2015; Eickhoff et al., 2012; Feyisetan et al., 2015; Ipeirotis and Gabrilovich, 2014; Itoko et al., 2014; Kawajiri et al., 2014; Kobayashi et al., 2015; Lee et al., 2013a; Lee et al., 2013b; Liu, et al., 2011a, 2011b; Martella et al., 2015; Massung et al., 2013; Nose and Hishiyama, 2013; Pothineni et al., 2014; Prandi et al., 2016; Preist et al., 2014; Prestopnik and Tang, 2015; Roengsamut et al., 2015; Simões and De Amicis, 2016; Snijders et al., 2015; Talasila et al., 2016; Tinati et al., 2016; Vasilescu et al., 2014)	26
Positive effects on the qualitative contribution	(Dergousoff and Mandryk, 2015; Eickhoff et al., 2012; Feyisetan et al., 2015; Goncalves et al., 2014; Ipeirotis and Gabrilovich, 2014; Kawajiri et al., 2014; Kobayashi et al., 2015; Lee et al., 2013b; Massung et al., 2013; Prestopnik and Tang, 2015; Runge et al., 2015; Simões and De Amicis, 2016; Sørensen et al., 2016)	13
Positive effects on continued work / long-term engagement	(Itoko et al., 2014; Kawajiri et al., 2014; Kobayashi et al., 2015; Lee et al., 2013b; Massung et al., 2013; Prestopnik and Tang, 2015)	6

Fig. 10 Positive effects of gamification in crowdsourcing reported in Morschheuser and colleagues' (2017) review.

1.4 Design between social innovation and game-based crowdsourcing

The previous paragraphs discussed the touchpoints between crowdsourcing and games. In particular, it has been addressed how fun and enjoyment are relevant factors in crowdsourcing (par. 1.2.1) and what makes games productive systems, rather than mere leisure ones (par. 1.3). Therefore, crowdsourcing systems can employ games to empower the enjoyment of the activity and without undermining their productivity. Par. 1.3.1 supports this concept by proving evidence about the positive effects of gamifying crowdsourcing systems.

The following vets into the idea that game-based crowdsourcing systems can be valuable in the field of social innovation for their features. Manzini defines social innovation as «a process of change emerging from the creative re-combination of existing assets (from social capital to historical heritage, from traditional craftsmanship to accessible advanced technology), the aim of which is to achieve socially recognized goals in a new way» (2014, p. 57). Social innovation occurs when new ideas meet social goals (Mulgan et al., 2007; Murray et al., 2010) and improve society's capacity to act by creating new social relationships or collaborations (Murray et al., 2010).

Top-down or bottom-up processes can both produce social innovation, but often they combine in hybrid processes (Manzini, 2014). Murray and colleagues affirm that «most social change is neither purely top-down nor bottom-up» but «involves alliances between the top and the bottom» (Murray et al., 2010, p. 8). As described in par. 1.2, Brabham (2013) explains that crowdsourcing blends bottom-up and top-down processes. Hence, it is a model inherently suitable for being applied in the context of social innovation. Governments have applied it to solve societal problems by co-creating or co-designing with citizens, changing the way they deliver societal value (Randhawa et al., 2019).

Game-based approaches can further sustain the process of social innovation. Bayrak (2019) claims that video games and game design can catalyze design for social innovation and offer a space for collective exploration of problems and solutions. He applies a theoretical analysis to identify seven merits of games to design for social innovation (Fig. 11).

The game world and reward system encourage active participation, hence games are **participatory** (Murray, 2017). Games persuade in a **procedural** way, i.e. through direct and indirect in-

Fig. 11 Merits of games to design for social innovation and their explanation (Bayrak, 2019).

teractions that become expressive in their process (Bogost, 2007; Murray, 2017). They are artefacts that **reach** various ages, demographics, social and ethnic backgrounds (Bayrak, 2019). They allow playing in an **expressive** way by performing, strategizing, improvising and acting (Bogost, 2007). They are **effective** in describing and reproducing complex systems (Bogost, 2007; Sutton-Smith, 2001). They appear **compelling and engaging** so that they catch interest and attention, pushing to voluntary action and participation (Bayrak, 2019). Finally, games are **systemic** because all their elements and features are interconnected (Salen and Zimmerman, 2004).

Bayrak explains that these merits can help face challenges of designing for social innovation. Bayrak explains that design for social innovation has four main challenges:

- participation;
- transferability and scalability;
- increasing complexity;
- replication and adaptability.

Fig. 12 Relations among merits of games, challenges of design for social innovation and some common approaches in the context of designing for social innovation (Bayrak, 2019).

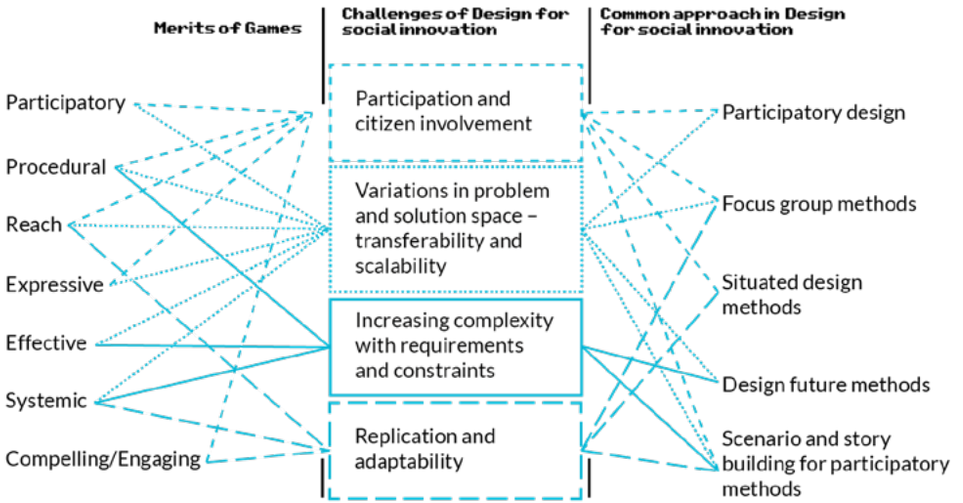


Fig. 12 shows that games merits can address these specific challenges of designing for social innovation and support approaches that are commonly implied in the context of social innovation.

Games can improve participation rates in social innovation because they reach many people and encourage their active involvement in their system thanks to their participatory nature. Games then reinforce engagement through expressive, compelling and engaging experiences which can be understood through games' procedurality.

Games can even sustain variations in problem and solution space, enabling transferability and scalability based on their systemic, effective and procedural features. Games can scale and reach even more players to push them to participate in social innovation. Those players can explore the game and social innovation at hand by interacting with it in expressive and free ways.

When complexity increases with requirements and constraints, games can sustain the design for social innovation by representing and inquiring the complexity through its system, in particular through its procedurality and effectiveness, which describe and express the specific process matter of the design for innovation.

Finally, games can tackle the problem of replication and adaptability of designs for social innovation. They can reach new user bases and engage them. At the same time, games as systems can adapt to the new circumstance and still be effective in describing the object on which the design for social innovation focuses.

This idea of using games in the context of social innovation has stimulated the branch that is known with the name of "Game for Impact". Stokes and colleagues (2016) note that the rise of this category of games was somewhat unexpected. According to them, the fact that those games emerged organically is the reason why researchers and practitioners have focused on the details of the phenomenon, rather than studying and analyzing it as a whole. For this reason, the term is still highly unclear and debated. Next paragraph explores this issue in depth.

1.4.1 Which impact deserves recognition? Disputes on Game for Impact definition

Samuel Liberty, game designer, lecturer for Northeastern University's College of Arts, Media & Design, collaborator of the Engagement Lab at Emerson College and co-founder of the games-for-impact consultancy Extra Ludic, considers a social impact game any game that creates an impact on the outside world (Liberty in Hundal, 2017). «Whether it be benevolent or nefarious, purposeful or unintentional, when the medium is used to create an impact on the outside world beyond the scope of the game, it becomes a social impact game», he states. For Liberty, intent (either positive or negative, present or absent) is not necessary to describe a social impact game. He considers part of this category only games which produce real-world impact beyond the scope of the game. This is interesting because Liberty unlinks the game goal from the impact that it produces. This could be somehow divergent from the Game Awards interpretation, as their reward category "Game for Impact" is meant «for a thought-provoking game with a pro-social meaning or message», a definition which implies intent. However, Liberty extends the concept so broadly that could lose sense to make a difference at all, since he considers part of the category also games that unintentionally produce real-world impact.

Conducting an extensive literature review of empirical evidence of game impact, Connolly and colleagues (2012) analyzed 129 papers showing that serious games, game-based learning (GBL) and commercial-off-the-shelf (COTS) were all able to provoke some sort of consequence. The most common outcomes reported in the various studies were affective and motivational (33), knowledge acquisition/content understanding (32) followed by perceptual and cognitive skills (20), behaviour change (13), physiological outcomes (11) and social/soft skills outcomes (11). Since games can stimulate all these effects in players, Liberty's point of view does not provide sufficient limits to make differences between a common game and a game for impact. On the other hand, the Game Awards' definition stresses the need of designers' intention of stimulating a social impact, which clarifies a distinctive feature to define a game for impact as such.

During a talk at the Game Developer Conference 2019 Maler explains that «any creation, whether intended or not, will have a social impact on its audience» (Maler, 2019). She stresses that this kind of impact is not a choice, but a fact, hence it is an **indirect social impact**. For this reason, she argues that it is important to make inclusive games, i.e. consider the impact that each creative decision could have on society and individuals and limit potential negative impacts. However, she clarifies that inclusive games are different from activist games, which intentionally attempt to actively support a social cause instead. Activist games aim at a **direct social impact** which is both commercially and militantly risky. Maler states that games of this kind will hinder some people and may generate little sales while being hard to develop to achieve a real impact.

Although Activist Games (Kafai, 2008) are different from Games for Impact, this reasoning on the difference between direct and indirect social impact is meaningful for this discussion. Liberty seems to do not differentiate between direct and indirect social impact and account them as equal. Contrary, the Game Awards' definition of Game for Impact stress that it has to be "thought-provoking game", implying that the social influence is intentional and carefully addressed. At the same time, the Game Awards' definition may exclude some games as they conceive them as a medium for a message and not as a system which can move or support pro-social action.

Games for Impact is a term still in its early days: it arose from the Games for Change movement but the discussion on its definition has produced many conflicting opinions. Stokes and colleagues (2016) record research and practice show deviations and inconsistencies, as the field has grown organically and not from a single research program. They investigated the matter through qualitative focus groups interviews with practitioners, participants and beneficiaries, as well as in-depth interviews with experts, game designers and funders to further investigate those divergences. They conclude there is an evident fragmentation in the field and

remark this causes struggles in discussion and progression on related topics[10].

Stokes and colleagues declare a reason for this fragmentation is due to ideological interpretations of the term impact which leads to polarisation of opinions. They report this is plain in the case of *Foldit*. Although the case study is extensively presented in par. 2.1.2, this game clarifies the multifaceted nature of this branch of games. On one side, members of the learning science community argue *Foldit* did not produce impact because it did not explicitly teach players about proteins, while some practitioners with training in community organizing consider that it achieved impact by successfully harnessing collective action towards a civic goal. This different judgment comes from implicitly focusing on a single kind of impact (learning or aggregating) and denying any other kind of impact.

In this complex scenario, game-based crowdsourcing's impact is still not clearly recognized nor yet framed. In particular, considering the distinction between direct and indirect social impact, it seems proper to address the indirect social impact that game-based crowdsourcing may produce. For this reason, it is important to consider its ethical implications.

1.4.2 Ethical implications of game-based crowdsourcing systems

Game-based crowdsourcing systems raise many ethical issues because they lay in between different fields which struggle with their specific ethical matters. Indeed, they have to face the problem coming from the controversial use of crowdsourcing (Standing and Standing, 2018), games (Zagal et al., 2013; Sandovar et al., 2016), gamification (Kim and Werbach, 2016), and persuasive technology in general (Berdichevsky and Neuenschwander, 1999), as explained in the following. That said, this paragraph does not intend to be exhaustive on the topic, but rather demonstrates how complex ethics inherently belongs to the specific context of game-based crowdsourcing systems.

Game-based crowdsourcing can be described as persuasive technology, an artefact created primarily to change the attitudes and behaviours and persuade their users (Berdichevsky and Neuenschwander, 1999). In the case of game-based crowdsourcing, framing the task as a game or gamifying it are methods to convince the user to contribute, actually offering freely their labour, in exchange for an enjoyable experience. This trade is not unethical per se, but it can stumble into unethical applications.

Berdichevsky and Neuenschwander (1999) argue that the analysis of persuasive technologies' ethics has to focus on the methods employed in the persuasion itself since motivations or outcomes tend to be the same as in traditional person-to-person persuasion (Fig. 13). Therefore, the employment of games and game design to persuade is the element that has to be analyzed.

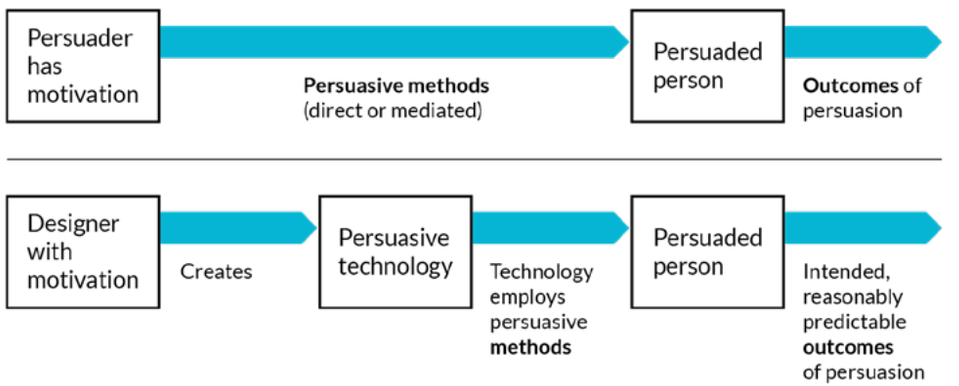


Fig. 13 Framework of persuasive technologies (top) compared to traditional persuasion (bottom) (Berdichevsky and Neuenschwander, 1999).

10 The document is available at: http://gameimpact.net/reports/fragmented-field/

Berdichevsky and Neuenschwander have set a collection of principles to design persuasive technology ethically:

1. «The intended outcome of any persuasive technology should never be one that would be deemed unethical if the persuasion were undertaken without the technology or if the outcome occurred independently of persuasion.
2. The motivations behind the creation of a persuasive technology should never be such that they would be deemed unethical if they led to more traditional persuasion.
3. The creators of a persuasive technology must consider, contend with, and assume responsibility for all reasonably predictable outcomes of its use.
4. The creators of a persuasive technology must ensure that it regards the privacy of users with at least as much respect as they regard their own privacy.
5. Persuasive technologies relaying personal information about a user to a third party must be closely scrutinized for privacy concerns.
6. The creators of a persuasive technology should disclose their motivations, methods, and intended outcomes, except when such disclosure would significantly undermine an otherwise ethical goal.
7. Persuasive technologies must not misinform in order to achieve their persuasive end;
8. **The Golden Rule of Persuasion**: The creators of a persuasive technology should never seek to persuade a person or persons of something they themselves would not consent to be persuaded to do» (Berdichevsky and Neuenschwander, 1999, p. 52).

The last principle is marked as "The Golden Rule of Persuasion" as it helps overcome issues which are not specifically addressed by the other rules, like cultural differences. While its effectiveness resides in its capacity to be relevant for all cases, the Golden Rule needs the other principles to not be twisted. Indeed, a person with an arguable opinion on a topic may find it proper to convince others to share the same idea; however, the other principles may set some restriction of the possible unethical use of the technology. Berdichevsky and Neuenschwander provide an example of this: some people might want to persuade others to abort a fetus and let be persuaded of the same. While the Golden Rule would be satisfied, this persuasive act would have to face the other principles, and Berdichevsky and Neuenschwander argue that it would have already a lot of issues with the first one.

By following the principles, it is possible to avoid unethical use of games and game design in the context of game-based crowdsourcing. Moreover, Berdichevsky and Neuenschwander claim that responsibility for the persuasive technologies' built-in motivations, methods, and outcomes falls squarely on its creators and purchasers. However, developing games is a process that puts more accountability on the creators. Game designers are usually regarded as advocates for players but sometimes their interests do not align with players' ones (Zagal et al., 2013). This discrepancy of interests might manifest in games as questionable and even unethical patterns, what Zagal and colleagues (2013) call "dark game design pattern".

They define a dark game design pattern as «a pattern used intentionally by a game creator to cause negative experiences for players which are against their best interests and likely to happen without their consent» (Zagal et al., 2013, p. 7). They claim that these patterns are intentionally and purposefully employed to evoke a given behaviour, that leads to a negative experience for the players. They argue that a dark game design pattern does not result from bad design, lead by ignorance, bad trade-offs, or lack of time and resources. Instead, a dark game design pattern is the product of designing for bad, so they deem these patterns as unethical. Some examples of dark game design patterns are grinding and playing by appointment. They make it hard for players to understand the commitment that the game requires and forces them to adapt their real routines to the gameplay.

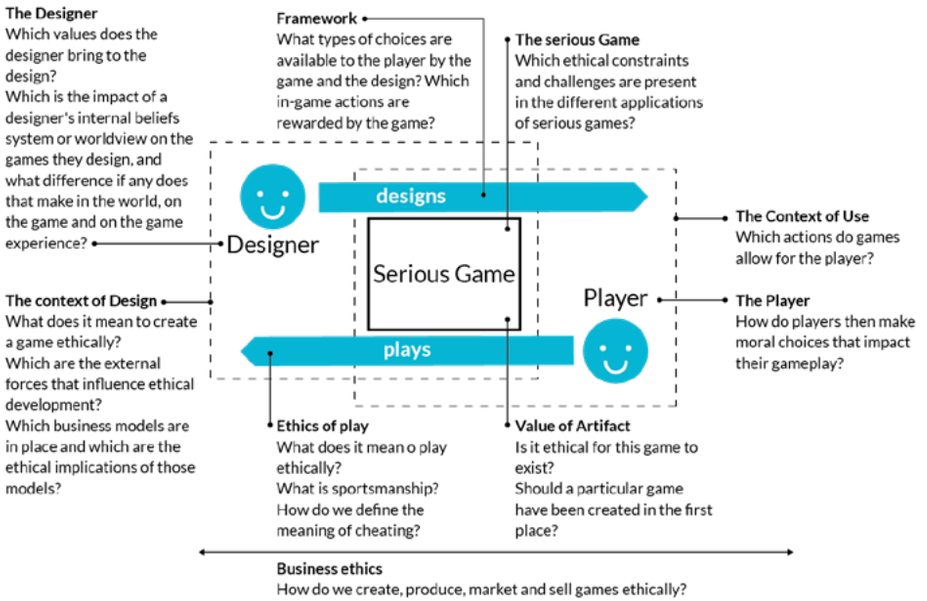
In general, Zagal and colleagues provide some suggestions to distinguish a dark game design pattern. They discuss that a pattern may become dark if it does not support a person's play style and/or manipulates them, hence it does not allow players' mindful consent. They stress the idea that unknown patterns are more likely

to have unknown negative consequences. Players with enough manipulation literacy can spot manipulative patterns and protect themselves from them, opting-out the game for example.

However, Zagal and colleagues recognize that classifying game design patterns as dark is not an easy and clear task. They stress that these patterns can vary in how strongly or effectively they cause issues for players. They are also dependent on the context in which they are used, their implementation, intended audience, and other factors. Moreover, there is a certain level of subjectivity in recognizing dark game design patterns. Hence they describe some borderline cases, like GWAPs (previously described in par. 1.2.2 and 1.2.3), namely purposeful games which aim to tackle computational problems. They claim that the games on Von Ahn's website were fair as they informed players that they were training computers by playing so that the computers became able to solve problems that could broadly benefit society. However, Zagal and colleagues note that this design could easily turn dark by omitting this information. At the same time, they discuss that if the player is unaware of the ulterior motive of the game, they would not perceive a negative experience. It would be difficult therefore to recognize this as a dark game design pattern.

When discussing the ethical use of GWAPs, Zagal and colleagues consider also serious games, or SGs in short (par. 2.1.2), games with educational purposes rather than entertainment ones. They uncover possible ethical issues in the context of SGs. They affirm that it «is one thing to invite someone to play a game and tell them that they may learn something in the process and another to try to trick them into learning something» (Zagal et al., 2013, p. 6). Once again, the idea that the player has to freely consent to play a game, knowing its ulterior motive, is crucial. However, analyzing ethics in designing SGs can cover a wide variety of elements other than this.

Sandovar and colleagues (2016) have defined the "Ecosystem for Designing Games Ethically" (EDGE, Fig. 14) to investigate ethics in the context of SGs. It clearly displays the complexity of ethical issues related to SGs. The EDGE does not focus on investigating the ethics of persuasion through games. It contains all the parts



involved in ethical matters: the serious game, the context of use, the player, the ethics of play, the value of artifact, business ethics, the framework, the designer and the context of design.

Fig. 14 The Ecosystem for Designing Games Ethically, EDGE in short (Sandovar et al., 2016).

Sandovar and colleagues discuss the use of SGs in the context of learning and military applications. In the first one, the collection of data on the performance creates grey areas on the matter of data ownership, often unclear and outside of the learner's control, and on the content access, which may be available only to students at a certain level. Moreover, sometimes the average player is likely unaware and not in a position to consider the implications of the hidden purpose of a SG due to non-transparent algorithms. Regarding military games, the authors argue that they are questionable as they may change the personality of players and affect their level of aggression while being usually open to the general public. Considering how they may impact people, it may be problematic that they are so easily accessible. The person may be responsible to properly use the tool, which is not dangerous itself. However,

Sandovar and colleagues explain that such SGs can facilitate the process of harming others, which is a highly debatable outcome. To summarize, unethical applications of SGs appear to lead mostly to manipulation, either by omitting information or by tricking the player in harmful moral and behavioural changes.

Likewise, gamified artefacts struggle with the threat of manipulating, facilitating harm or affecting negatively people's character. Indeed, Kim and Werbach (2016) suggest that the use of gamification stumble into ethical issues when it:

1. «takes unfair advantage of workers (e.g., exploitation)»;
2. «infringes any involved workers' or customers' autonomy (e.g., manipulation)»;
3. «intentionally or unintentionally harms workers and other involved parties»;
4. «has a negative effect on the moral character of involved parties».

The majority of these points have been discussed also in regard to games or SGs, but exploitation is unexplored so far. Exploitation is a very hot topic in gamification. For example, Bogost (2011a) even named gamification "exploitationware" to stress how much he considers unethical the practice. Exploitation occurs when a party takes advantage of one's vulnerabilities to push them to agree on an exchange. Gamification is usually employed to direct the user to change their behaviour in non-gaming contexts (Robson et al., 2015), hence achieving a long-term effect usually beneficial to the user. While the compensation on the spot may be mere points or badges, gamification can promise long-term benefits. For example, a gamified application designed to teach languages like *Duolingo*[11] has its reward system to keep the user engaged. At the same time, it supports a new competence acquisition, namely knowing another language. The moment the application lacks a long-term benefit, it is likely to be exploitative.

11 https://en.duolingo.com/

Kim and Werbach (2016) believe that there are two primary reasons why the business practice of gamification raises important ethical issues: 1) the overlay of virtual and real-world norms and 2) the tension between organizational and individual interests. The second point was somehow addressed by Zagal and colleagues (Zagal et al., 2013) when discussing the possibility that dark patterns emerge from a discrepancy between the creators and players' interests. Kim and Werbach claim that gamification is not per se exploitative, manipulative, harmful, or detrimental to character but the specific conditions of implementation must be considered to understand whether there are ethical issues. At the same time, they recognize that it is not possible to dismiss those objections out of hand. They map these four elements on two axes (Fig. 15), one representing the experience (real world and game world, representing the overlay of the two) and the other the actors (relational and individual, representing the tension between organizational and individual interests).

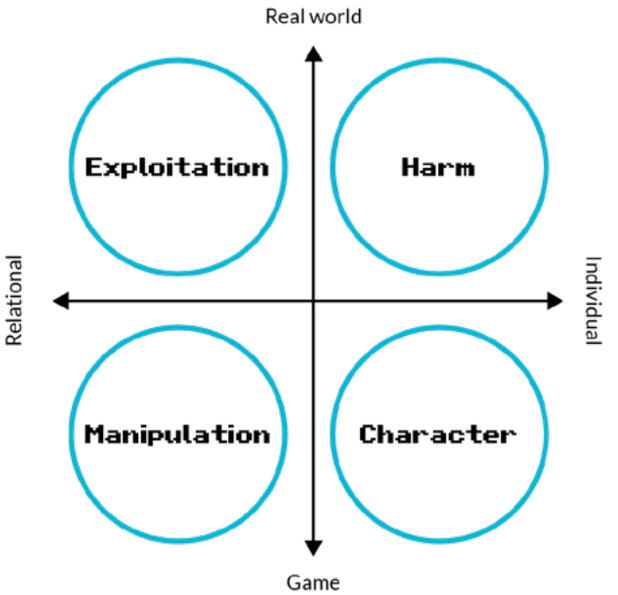


Fig. 15 Conceptual mapping of gamification ethics according to Kim and Werbach (2016).

Kim and Werbach explain that exploitation and manipulation result from the relationship between the providers and the players. When the relationship is uneven in the real world, and the providers can leverage players through their vulnerabilities, there is a exploitation issue. Instead, when the players make decisions which are beneficial to the service provider and not completely autonomous, there is a manipulation issue.

Kim and Werbach define harm and character as related to the players as individuals. When the activity prompted by the service provider produces physical or psychical injury in the real world, there is an harm issue. When there are ethical lapses, e.g. players act to satisfy the game's objectives and play indifferent to fundamental human values, there is a character issue.

Game-based crowdsourcing systems would be therefore a crowdsourcing system that exploits games and game features as a persuasive method. The aforementioned studies demonstrate the number of ethical issues that can generate from employing games and game design to influence and persuade people. However, the use of crowdsourcing per se can have ethical issues to carefully regard. Standing and Standing (2018) have drawn a set of key ethical issues (Fig. 16) in crowdsourcing and have clustered them in three macro areas: knowledge, economics and relational. The list of issues is quite wide, but Standing and Standing discuss it deeply and carfully.

Regarding the **knowledge** cluster, Standing and Standing (2018) identify five themes: manipulation, exploitation, loss of professional status, loss of ownership and unrewarded solutions.

Standing and Standing explain that crowdsourcing can incur into manipulation. For example, crowd voting is often used to evaluate generated ideas. This evaluation mechanism can establish a perception of egalitarianism and democracy while the outcomes of voting can justify management decision making. However, at the same time outcomes may not truly represent the crowd opinion if the crowdsourcer manipulates or censors the process for its own ends. The design of the voting space should avoid possibilities of manipulation or interference. while being clear, transparent and supported by crowd members and owners.

Crowdsourcing domain	Firm motivation/justification	Ethical issue	Recommendation
Knowledge	Obtaining new ideas, accessing knowledge to solve problems, and accessing expertise to evaluate proposals and products.	Crowd processes and voting can be subject to manipulation and the most highly rated ideas may not be a true representation of crowd opinion.	Firms should communicate the ground rules and guidelines for evaluation and avoid manipulating or censoring the process for their own ends.
	Firms may justify their use of crowdsourcing by arguing that consumers will benefit because there will be "better" products.	If participants' knowledge and expertise are not recognised and/or rewarded commensurate with their value it can be viewed as exploitation.	"Better products" is a weak justification and does not excuse inadequate valuation of contribution. Proper recognition of the value of ideas needs to be given.
	Firms may need highly skilled professionals and in-depth knowledge.	Crowdsourcing can undermine professional status since members of a crowd may lack identity and feel their self-worth is under threat when not properly recognised.	Reward and recognise level of knowledge and skills obtained.
	Creativity can occur through crowdsourcing by bringing together individuals and melding alternative perspectives in a collaborative task.	Knowledge from many domains can be applied in reaching a solution. Identifying and rewarding individual contributions and ascertaining ownership of ideas can be overlooked.	Firms can consider team rewards or involving the team in allocation of rewards. Firm should determine process upfront.
	Although the creativity generated by the challenge to the crowd may produce novel or innovative ideas, these do not always translate into an innovation.	Many ideas and solutions from the crowd may go unrewarded.	Consider how ideas that are not selected are to be managed, particularly those that have reached a short-listing stage because crowd members may have expended considerable effort on solutions.

Fig. 16 Ethical issues in crowdsourcing clustered in the knowledge cluster (Standing and Standing, 2018).

Standing and Standing illustrate that crowdsourcing knowledge can lead to exploitation and loss of professional status. Indeed, members of a crowd may lack identity and feel their self-worth is under threat when they are not properly recognised and rewarded. In this case, the discrepancy between the participants' knowledge and expertise and their value can be viewed as exploitation. It is necessary to avoid undermining professional status and value properly all contributions.

Standing and Standing point out also that crowdsourcing, when applied in dynamic collaborative work, can struggle in identifying and rewarding individual contributions and ascertaining ownership of ideas. Creativity in problem-solving can occur through crowdsourcing by gathering individuals from different social contexts and melding their alternative perspectives in a collaborative task. In this way, the problem will be interpreted in different ways and knowledge from many domains will converge in reaching a solution. However, this process blends and makes unrecognizable individual contributions, making it really hard to determine ownership of ideas, which can be hence overlooked.

Finally, it is noted that solutions should be carefully managed to avoid unreward the efforts of participants. Although crowd creativity may produce novel or innovative ideas, it does not mean they translate into innovation. Indeed, those ideas may incur into manufacturing, technical, or marketing barriers that make them too difficult to realise. The crowdsourcer should consider how to manage ideas that are not selected, particularly those on which crowd members may have expended considerable effort.

Regarding the **economic** cluster, Standing and Standing identify four themes: crowd pressure and employees' devaluation, poor remuneration and IP options.

The authors explain that if an organization does not like an idea, there is no pressure to use it and devalue its employees' opinion. The crowdsourcer can benefit from involving a crowd in organisational tasks both in terms of costs and reducing reliance on customer surveys and more costly marketing activities – since crowdsourcing in the product development process can quickly unveil problems and advantages of existing goods or services which

Fig. 17 Ethical issues in crowdsourcing clustered in the economic and relational cluster (Standing and Standing, 2018).

Crowdsourcing domain	Firm motivation/justification	Ethical issue	Recommendation
Economic	Minimises financial risk.	If a firm does not like an idea there is no pressure to use it—it minimises risk of financial loss associated with investing in an unused idea.	Transparent process and decision making.
	Saving money by crowdsourcing.	Firms may not pay the going rate for labour/skills.	Pay appropriate rates for skills/knowledge/labour.
	Acquiring new IP.	Participants should be unaware of IP options and be taken advantage of.	Firms should help participants be aware of IP options.
	Low cost to access crowd and low associated risk.	Internal employees may feel devalued if by-passed by the crowdsourcing process.	Internal employees should be canvassed for their ideas and insights and not be by-passed.
Relational	Co-creation process.	A power imbalance exists. The cocreation concept can provide a legitimation for "borrowing ideas" without due reward and without properly questioning unethical practice.	Transparent process and decision making. Crowdsourcing should not be used to exploit participants. Firms need to take an ethical stance at the outset.
	Transparent objectives.	Unclear objectives may mislead people into participating.	Transparency in a firm's motivations is important so that potential participants can make an informed decision about their involvement.
	Firms may draw on specialist communities too often and may see them simply as a service.	Firms need to avoid over-using communities with constant requests and need to be cognisant of a community's norms, standards, and behaviours.	Firms should be aware of the standards of acceptable behaviour and the moral code of communities.

would normally be identified through user research. It would be unethical that internal employees' ideas and insights were bypassed by the crowdsourcing process: it is necessary that the crowdsourcing organization does not rely solely on the crowd opinion.

Standing and Standing stress even that labour and skills should be paid proportionately. Remuneration for crowdsourcing tasks is variable, ranging from zero to million dollars. This issue stems also from the practice of portraying users as amateurs and by so reducing the obligation to pay at market rates. There is the perception that the participants contribute for the fun and the thrill of having their ideas chosen, while it is not always the case. Contributors should be aware of their rights and responsibilities in regard to the associated recompense and IP and copyright as well. IP protection, established to prevent the opportunistic behaviour of organization collaborators, and copyright should be carefully considered. However, IP protection systems can constrain input as they are underpinned by organisational attitudes that presuppose collaborators are self-serving and economically motivated.

Regarding the relational cluster, Standing and Standing identifies three themes: power imbalance, transparency and over-usage of communities.

They illustrate that crowdsourcing used in co-creation processes can cover power imbalance. Co-creation is a process which integrates customers in the production process and enables organizations to expand internal capabilities and keep at the forefront of the industry. The main advantages of this approach are the identification of future customer needs, a broader decision basis, increased efficiency in gathering and use of customer information, and increased customer retention. Organizations can employ crowdsourcing to incorporate virtual crowds into the production process, hence establishing a co-creation process. Despite the term co-creation implies a willing agreement of groups on a similar power stance, there is clearly a power imbalance at play which can enable "borrowing ideas" without compensating and questioning unethical practice.

Moreover, it is also discussed that unclear objectives may mislead people into participating. Standing and Standing claim that participants need to make informed decisions about their involvement in a crowdsourcing activity, therefore it is crucial that the crowdsourcer is transparent regarding its motivations. Clear communication of the required target would also make professionals understand when the aim is to engage with nonexperts. A clear objective statement from the organization would guide the decision as well. Participants may well feel there is little at stake and be less hesitant to participate if the aim of the crowdsourcer is to generate awareness or a sense of consumer belonging. On the other hand, if the organization aims to develop new products or services relying on ideas generated by the crowd, participants may spend time considering whether the firm's rewards are commensurate with the potential of the idea.

Finally, Standing and Standing claim that an organization needs to avoid over-using communities with constant requests, but it has to be aware of a community's norms, standards, and behaviours. Those are built in a community by its ability to support or reject the contributions of others: the group hence to set standards of acceptable behaviour and develop a moral code. The crowdsourcer should account community sustainment and acceptance because it empowers members of the community and leads to better quality ideas. Moreover, it is important for the community to have the possibility to put pressure on requesters to do the right thing for their workers.

This paragraph does not attempt to find a solution for ethical matters. It is a complex and wicked topic to discuss, which requires extensive analysis. Instead, the intent is providing an overview of the complexity of the theme in the context of game-based crowdsourcing. It is important to remind how tricky ethics can be on systems that converge knowledge from different fields. Game-based crowdsourcing systems emerge from such a convergence, which is clearly also their strength. Although, achieving this strength requires many different competencies that usually do not come together. Next paragraph discusses the challenge of building interdisciplinary teams and the great results that collaboration produces in the field of game-based crowdsourcing.

1.4.3 For a better collaboration between practice and research

Stoke and colleagues (2016) admit the fragmentation in the field of Game for Impact is not all bad as it exposes the efforts of a rapidly growing community and it can be beneficial at some extent. Although, they argue that too much fragmentation can seriously limit the community's potential as it will shape the design, the legitimacy of the funding assigned to these games and ultimately their impact. Stoke and colleagues set their next objective to gather perspectives and resources at the intersection of research and design. In a keynote at Game for Change 2017, Steinkhuehler considered this a crucial aspect as well. She stressed the importance of bridging the market and the research to better understand to what extent games can impact on society (Steinkhuehler, 2017).

Social impact products and service are recognized to be especially challenging as they need many competences to cooperate. Gretchen Anderson has worked on social issues in local communities as that of Detroit and Indianapolis. In her report "Designing for Social Impact" she stresses that social impact work requires even higher level collaboration then commercial projects do (Anderson, 2015).

As collaboration is hard, it may be the main barrier to the spread of the practice. Professionals in the field recognize this issue:

> Finding partners is probably one of the most difficult parts of this equation. Where does a motivated game designer go to find a scientific researcher with a good question to work on? Where does a motivated researcher go to find a game designer with the patience, adaptability and motivation to work in the sometimes messy world of research data? There's often a third partner, someone who understands crowdsourcing or citizen science, someone who understands the citizen science audience. These different people rarely stumble into each other at meetings or trade shows or conferences.
> (Interview with Jennifer Couch and Dave Miller, 2018)

Couch and Miller recognize that conferences and meetings like Games for Change and the Citizen Science Association strive to reach out and build bridges across these fields. They even suggest that game designers interested in pursuing a game aim to support a citizen science project could attend a conference where many citizen science practitioners would be, e.g. the Citizen Science Association conferences. Another option for game designers interested in working in the biomedical field might be to reach out to disease or community advocacy organizations. Couch and Miller explain that advocates often have contacts who are more motivated to work with the public and to try different approaches, such as games. Such contacts may be researchers, organizations, graduate or medical students.

Couch and Miller record a striking growth in the field. They remind that in a few years the Game Developer Conference moved from displaying a few small indie projects connected to the topic of cancer to showcasing games that deal with cancer from many points of views (treatment, diagnosis, education, anxiety-handling). They also recognize that hackathons and jams developed around scientific research or biomedical questions may be a bit of a foot in the door.

Couch and Miller believe that *Project Discovery*, a citizen science minigame in *EVE Online* (deeply described in par. 2.1.3), marked a turning point. In its first version, *Project Discovery* asked players of *EVE Online*, a great MMORPG to classify images from the Human Protein Atlas, a project that aims to map all the proteins in the human body. As they stress, *Project Discovery* has demonstrated a model to place citizen science projects in the context of a commercial game. The minigame is the result of a flourish collaboration between the Icelandic gaming company CCP Games, the Reykjavik University, the Swedish-based program Human Protein Atlas and the Swiss startup Massively Multiplayer Online Science (MMOS[12]).

12 http://mmos.ch/

MMOS appear to be particularly relevant in that partnership. MMOS is a privately held startup company specialized in citizen science and crowdsourcing. The founders of the company devise innovative solutions by combining the knowledge from many years in academia with extensive IT skills. They stated clearly their mission on their website:

> We created Massively Multiplayer Online Science to **connect scientific research and video games as a seamless gaming experience.** Research tasks completely integrated with game mechanics, narrative and visuals can open up a new channel between the gamer and the scientific community. Converting a small fraction of the billions of hours spent with playing video games will bring an enormous contribution to scientific research, and in the meantime will change how video games' expertise is perceived.

MMOS is the leading member of GAPARS (acronym for *GAmification of PARticipatory Science for training and education purposes*), a consortium that unites 8 highly reputed European universities, companies, and public institutions. The biggest European Union's Research and Innovation programme ever, Horizon 2020[13], funds GAPARS[14]. Significantly, the European Commission has given the maximum possible points to GAPARS among 12 granted projects responding to the Horizon 2020 call for the serious games industry and gamification. Therefore, the European Union endorses the strategy of the consortium, which consists of the development of non-leisure activities leveraging the existing video game industry and its technologies. GAPARS aims to support a massive increase in the number of participants and gamified contributive projects through the tools and platforms developed by its partners, thus contributing to the development of a sustainable industry.

13 https://ec.europa.eu/programmes/horizon2020/what-horizon-2020
14 https://cordis.europa.eu/project/id/732703; for further details see Sullivan et al., 2018: https://www.nature.com/articles/nbt.4225

Fig. 18 Complexity Limit competing for the World of Warcraft Race to World First on Ny'alotha, The Waking City.

Collaboration can even rise from many different organisations. In this respect, it is remarkable the partnership between Balanced Media|Technology[15], an healthcare AI company that develops GWAPs, and Complexity Gaming[16], one of America's premier and longest standing esports organizations. In January 2020 there was the latest *World of Warcraft* Race To World First, where guilds competed on the raid *Ny'alotha*, *The Waking City*. The Race to World First is a unique esports phenomenon in which at the release of every new raid in *World of Warcraft* guilds attempt to beat the final boss on the hardest difficulty, namely Mythic. Among the various competitors, Complexity line up their guild Limit.

15 https://www.balancedmediatechnology.com/
16 https://complexity.gg/

Balanced Media|Technology produced an extension that provided up to date progress, player updates and contained the online GWAP *The Omega Cluster*[17] right inside Complexity's Twitch stream. *The Omega Cluster* is meant to improve a machine learning AI used in actual cancer treatment research by isolating promising combinations of co-medication properties. Limit was able to win the race, while the extension received 6.7 million views and viewers played nearly 13 thousand rounds of the game, equivalent to 900 hours of play, almost six months worth of real world research time. It is impressive how even a single event of that relevance, and therefore highly attractive, can boost that much research.

Calls for funding and awards are increasingly publicized for social impact games (Stokes et al., 2016). Consortia as GAPARS can combine different approaches that bridge market and research, thus producing innovation. Even short-term collaborations as the one of Balanced Media|Technology and Complexity contribute to connect the gaming industry and the scientific field. «Finding partners/collaborators, finding the right opportunity or the right question for the right skills and method are always tough questions when one innovates or steps outside their own field of expertise» (Interview with Jennifer Couch and Dave Miller, 2018), but this direction is really promising for the future of social innovation through game-based crowdsourcing.

17 http://omegagame.balancedmediatechnology.com/

2 The design of game-based crowdsourcing systems

Chapter one reviewed and discussed the potentialities of game-based crowdsourcing systems. It particularly focused on explaining why games and crowdsourcing reinforce each other and what makes such systems suitable – and maybe even desirable – to support social innovation. If so far the discussion embraced a theoretical perspective, the following discourse digs into the practical implementation of game-based crowdsourcing systems which are systems nourished by fun experiences. Fun is a paramount concept when it comes to games and engagement; however it is also particularly blurry, requiring a specific discussion. It is crucial to know how to arouse and direct fun to engage in crowdsourcing activities and there are many ways to do so. At the same time, fun is not the only factor that drives participants. It is one among many in an ecosystem of motivators. It is important to understand how and when it is effective in engaging users, also in comparison with other motivators.

On the light of this introductory reasoning, this chapter investigates:

- fun and how games may be exploited to engage users and push them to participate in crowdsourcing projects (par. 2.1);
- the system of motivation for both volunteers of citizen science projects and gamers (par. 2.2).

Finally, par. 2.3 reframes the area of research of this thesis, based on the detailed discussion of those two aspects hereby presented.

2.1 What is fun and how to design for it

It has been widely discussed the relevance of fun in crowdsourcing participation (par. 1.2.1) and the idea that games can capitalize fun to sustain participation (par. 1.2.2). However, it is unclear what "fun" is and why it is so motivating. In his critical discussion on video games, their features, and their role in our society, Bogost affirms that fun is «a placeholder more than it is a description», «a surrogate term for some more complex yet unspoken sensation of gratification and satisfaction, rather than as a kind of *description* for that satisfaction» (Bogost, 2013 emphasis in original). Bogost claims that fun is related not to effect, but to structure and respect. He explains that fun is not something else that can be applied to things to make them better. He portrays fun as inherent to the activity, to its structure, and only when people submit themselves to that structure and respect it they can enjoy the activity. He claims that his point of view «shifts the form of fun from that of an experience to that of a kind of exhaust that's produced when an operator can treat a thing with dignity».

Why do people find fun in "treating an activity with dignity"? Why do they enjoy putting a lot of effort into doing something? Fun is actually a stimulus designed by natural evolution to reward and motivate. As Koster points out in his *A Theory of Fun For Game Design* (2004), fun is a moment of pleasure derived by endorphins release to reward learning. Knowledge increases survival chances, so the nervous system promotes it by compensation. Koster affirms that fun in games arises from the triumph that people feel when mastering strategies and skills to beat the game, i.d. when players have gathered enough knowledge to overcome the problem proposed in the game. By playing, people can experience hard challenges safely. Fun is such a powerful human stimulus that can overcome fear and help people face dangerous and frightening situations. Players can even be willing to complete boring and demanding activities to proceed in the game and conquer that satisfying sensation; of course, such boring and demanding activities should provide an adequate reward to the player who spent energy and time in solving them, or motivation will fall.

Fun works as an umbrella term for the varied emotions a game can produce. Game designer Marc LeBlanc, likely to Bogost, considers fun as a mere stand-in term for a more complex phenomenon. However, he digs into the concept in a different way from Bogost. In several talks at the Game Developers Conference, Le Blanc (2000) has suggested a typology to replace the general word "fun" with a more directed vocabulary. For this reason, he has developed a taxonomy which contains these categories:

1. Sensation: Game as sense-pleasure;
2. Fantasy: Game as make-believe;
3. Narrative: Game as drama;
4. Challenge: Game as obstacle course;
5. Fellowship: Game as social framework;
6. Discovery: Game as uncharted territory;
7. Expression: Game as self-discovery;
8. Submission: Game as pastime.

A game can pursue multiple typologies of fun and to varying degrees. The taxonomy is meant to support the analysis and description of games as multilayered entities where several meanings coexist (Mariani, 2016), to better understand how and why different games appeal to different players, or to the same players at different times (Hunicke et al., 2004). Salen and Zimmerman (2004) point out that many categories seem to overlap, e.g. Fantasy and Narrative. They affirm that Caillois' (1958) categories are

somehow a compact version of Le Blanc ones. Caillois defined four "fundamental categories":

1. Agôn: competition and competitive struggle;
2. Alea: submission to the fortunes of chance;
3. Mimicry: role-playing and make-believe play;
4. Ilinx: vertigo and physical sensation.

Salen and Zimmerman (2004) stress that it is not essential to choose a single typology since one of them is not necessarily better than the others. Instead, they offer a different way of thinking about fun, its many aspects and motivations and they are useful to organize observations. The authors suggest to mix and match different models and taxonomies depending on the needs of the specific game design.

Taxonomies, not only the ones mentioned here, can support the design of fun in a game. Along with comprehending and considering the various typologies of fun (Lazzaro, 2008), it is also crucial to identify a process to implement fun in the context of game-based crowdsourcing systems. Process refers to the way designers decide to harness games and game design to convey fun. A process must be strategic, hence it has to match stakeholders' resources, requirements and needs. Indeed, it is possible to identify three processes employed in the context of game-based crowdsourcing systems: gamification, SG development, and taskification. The three concepts are shortly presented in the next paragraphs along with explicative case studies which show the application of these approaches. Although these concepts are highly complex, they are addressed with the objective of comparing those approaches (par. 2.1.4). Therefore, the next paragraphs do not expand these concepts broadly but it focuses on clearly defining gamification, SG development, and taskification.

2.1.1 Gamification

Among the others, gamification became an increasingly popular approach in designing crowdsourcing systems (Hamari et al., 2014; Seaborn and Fels, 2015; Skarlatidou et al., 2019). According to Morschheuser et al. (2017), it is the intrinsic crowdsourcing dependence on a large number of people willing to contribute that stimulates the application of gamification in the field.

Despite its introduction dating back to 20 years ago, today gamification is still a highly debated term (Huotari and Hamari, 2016; Werbach, 2014). Already used at the beginning of 2000s[18], it became popular around the second half of 2010 and it is widely defined as «the use of game design elements in non-game contexts» (Deterding et al., 2011, p. 9).

Gamification has been accused of belittling game design complexity and game designers have renamed it *pointsification* (Robertson, 2010) and *exploitationware* (Bogost, 2011a). Critics' main concern is the fact that gamification moves the focus on incidental properties like points and levels (game elements, in general) as they were core features like interactions with behavioral complexity, sense of agency and competence that games are able to provide (Bogost, 2011b, 2011a; Robertson, 2010).

Despite these considerations, though, gamification began to spread and shape its own practical strategies and techniques (Zichermann and Cunningham, 2011), proving to be worthy. Then, after years of experimentations, even the academic field concluded that gamification "works" (Hamari et al., 2014). However, consistently with game designers' concern of oversimplification, Hamari and colleagues strongly pointed out 1) the role of the context being gamified and the 2) qualities of the users are key elements to consider during the design (Hamari et al., 2014). The case of

18 It is not clear who was the first to introduce the term. An interesting post on Quora reports many people as possible candidates (Who coined the term 'gamification'?, 2010).

The Guardian's *Investigate Your MP's Expenses* is a perfect example of gamification applied on crowdsourcing in the right context and with the right users.

Investigate Your MP's Expenses

In 2010, The Guardian was recognized as one of the best practices in computational journalism for digital news providers (Flew et al., 2010) and it is no surprise when analyzing *Investigate Your MP's Expenses* (The Guardian with Simon Willison, 2009).

In 2009, The Daily Telegraph of London began investigating two million leaked pages of documents related to expenses claims made by Members of Parliament (MPs) (Hicks, 2009) and posting news regarding a sample of them. As a consequence, this resulted in pressures on the House of Commons that hence released 700,000 documents concerning MP claims for household and office expenses (Rogers, 2009). It was one of the biggest scandals in British parliament history, but no institution had enough internal resources to handle the enormous number of documents. Among the press, The Guardian decided to directly ask their audience to help, launching a crowdsourcing platform on their server.

Users could report relevant documents declaring the type of document from claim, proof of claim, other document or blank page, and their level of suspiciousness among "not interesting", "interesting but known", "interesting", "investigate this!" (Fig. 20). They could add some extra comments (observations, the reasons

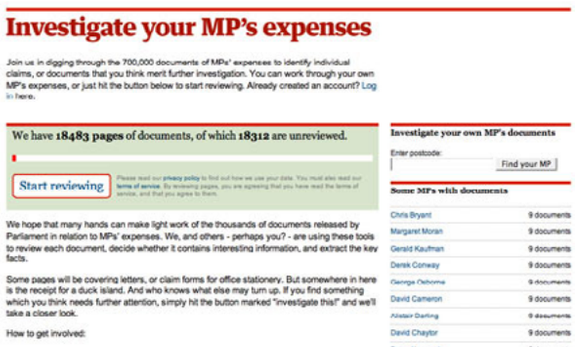


Fig. 19 The homepage of the original web application, displaying the project progression bar and the objective statement.

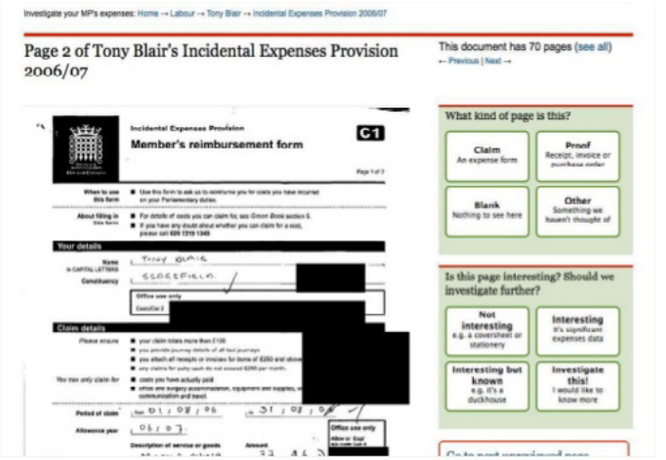


Fig. 20 The investigation interface from the original web application. Participants could report the kind of document and whether it was interesting to investigate further or not.

it needs further investigation). Once into the system, the claim was checked by internal journalists.

The website contained many sections. One displayed recently logged users and their contributions. In another section there was a progression bar of the number of documents reviewed that exhibited the progress of the collaborative effort. Each player had a personal "best individual discoveries" page that showcased their most valuable contributions. Moreover, there were two top contributor lists: one displaying users who investigated the most documents in the last 48 hours and one showing the ones that did the best from the beginning of the investigations.

Willison, an external collaborator to this project, declared in an interview «Your workers are unpaid, so make it fun» (Willison in Andersen, 2009). All those elements (the four-panel interface, progress bars, ...) were designed to provide a game-like feeling, namely to gamify the system. The easy categorization mechanic paired with great feedback and an evident community goal helped contributors maintain their motivation up.

After thehe later implementation of mugshots of each MP in the database (Fig. 21), there was an increase in participation and Willison interpreted the phenomenon as improvement in the narrative

Fig. 21 A MP's page on Investigate Your MP's Expenses, *showing their mugshot and their documents.*

of the experience. He argued that mugshots deepened the personal involvement of contributors as «everybody in the U.K. has an MP» (Willison in Andersen, 2009) and participants had the possibility to recognize their MP more clearly thanks to the images. Context and storytelling seemed to be great to persuade people to support The Guardian research, which may have led them to live-blogged the second half of the MP's expenses to keep momentum from fading (Daniel and Flew, 2010).

Investigate Your MP's Expenses was a great experimentation which achieved impressive results. After 80 hours 170,000 documents were already reviewed thanks to 20,000 users, with a participation rate of 56% (Hicks, 2009). Around a year later, in June 2010, about half of the 460,000 claims were scanned by 26,774 registered users (Flew et al., 2010).

2.1.2 Serious Games

However, combining games with crowdsourcing cannot always be considered gamification (par. 2.1.1). As Deterding and colleagues (2011) emphasize, gamification occurs only if applications embed game elements, while full-fledged games for purposes other than entertainment are so-called serious games (Fig. 22).

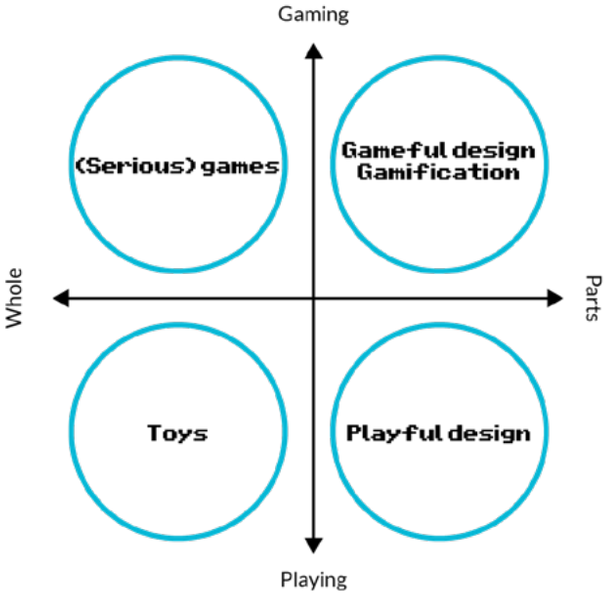


Fig. 22 "Gamification" between game and play, whole and parts (Deterding et al., 2011).

The expression "serious games" (SGs, henceforth) is way older than "gamification" and it was introduced by Abt in his namesake book. It describes games whose main goal is not entertainment but an explicit and reasoned educational purpose (Abt, 1987). Many researchers and practitioners did not appreciate the generic and conflicting message of the label SG (Ritterfeld et al., 2009), but it became an established term anyway.

Later, Susi and colleagues (2007, p. 5) suggested the definition of SGs as «games that engage the user, and contribute to the achievement of a defined purpose other than pure entertainment (whether or not the user is consciously aware of it)». They stress that SGs create 1) reliable simulations that focus on 2) problem solving through 3) elements of learning and 4) reflect natural (i.e., non-perfect) communication, while entertainment games 1) simplify simulations, 4) make communication perfect and 2) enrich experiences to 3) deliver fun. However, this not mean that SGs are not entertaining (Michael, 2006). Games are not the only en-

tertainment media that was used for other purposes: movies and books as well have been produced to convey serious messages to move their audiences (Michael, 2006). According to Ritterfeld and colleagues (2009), through their serious messages SGs can provoke three desirable outcomes in particular: learning (deliberate acquisition of skills or knowledge through practice and training), human development (psychological impact on processes of human development such as identity, attitude, emotional regulation, so on), and social change (social intervention, as political or health behavior).

However, SGs can even produce valuable social innovation. A great example of a SG as simulation of real-world situations reshaped as games which aims even to sustain scientific research is the acclaimed *Foldit*.

Foldit

Foldit[19] (University of Washington and Center for Game Science, 2008) is an ongoing citizen science game which helps scientists in protein structure prediction and design, both essential to cure diseases such as HIV, cancer and Alzheimer. Proteins can fold in a multitude of ways due to their many degrees of freedom and this makes it highly hard to predict their arrangement, even knowing they settle in the most stable state they can adopt. Possibilities are so high that calculating them requires a lot of money and time as the task is very tough for computers. On the contrary, the natural three-dimensional pattern matching capabilities of the human brain are very good at handling 3D puzzles. This consideration led to the idea of crowdsourcing protein structure simulation.

In *Foldit* people play competitively and collaboratively to fold the best proteins in 3D puzzle games. Players or groups of players try to solve proteins' puzzles or improve previous solutions to achieve better rank, which reflects the score of the best results they have found. The score for folded proteins is based on the Rosetta energy function, which relates lower energy to structures nearer the native (Cooper et al., 2010a, 2010b).

19 http://fold.it/portal/

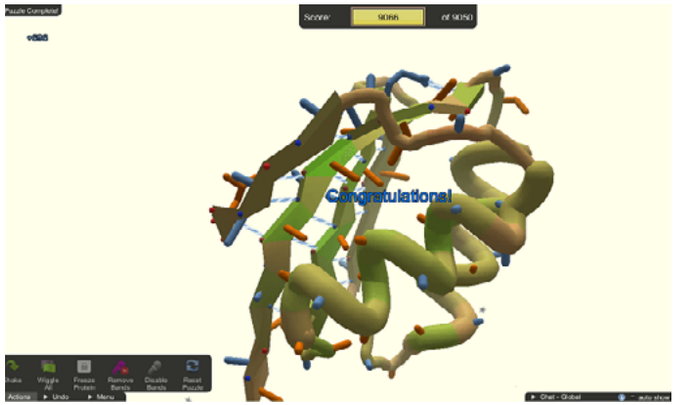


Fig. 23 After the puzzle has been completed the system rewards the player with an animation of triumph and records the score to update the player ranking.

Players interact with protein structures by direct manipulation tools and algorithms from the Rosetta structure prediction methodology. Common usage words replace technical terms so the game is more accessible for non-scientists. The game's visuals try as well to convey the impression of being approachable through its cartoonish style (Cooper et al., 2010a). Players learn the gameplay basics through a series of introductory levels and beginners puzzles, although they are not requested to complete them to try online puzzles. Game rules preserve biochemistry principles by guiding players avoiding interatomic repulsion, cavities in the structure and exposed hydrophobics. Game elements provide visual clues about those rules, to facilitate users comprehension of the structure (Fig. 24). At the same time, the game leaves enough freedom to explore different solutions, which are stimulated by intuitive and fun interactions. Those are designed from the concept of *touchability*, i.e. render the proteins so that they appear physical and moldable and players feels that they can interact with them, and keeping a balance between interactivity and results' accuracy (Cooper et al., 2010a). This balance is crucial to ensure both an entertaining experience for players through enjoyable interactions and meaningful data for scientists. Without one of these two

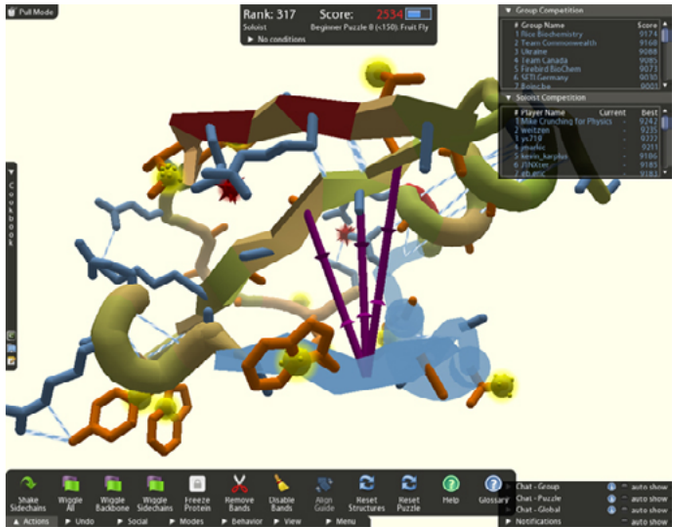


Fig. 24 An example of protein folding in the game in which all visual clues are observable.

elements, the symbiotic exchange of benefits which sustain the crowdsourcing would end.

The game simplifies visualization hiding unimportant clues until they become relevant to the gameplay to avoid information overload. Information overload is a phenomenon that occurs when one has too much information about an issue so that it becomes hard to understand it and effectively make decisions. When all clues are visible in the game (like in Fig. 24), players may be overwhelmed by the amount of information and get stuck in the game. This is a wise design choice, as clues are indeed many and very different. Spiky, rotating spheres floating between the overlapping atoms (clashes) alert atoms are unrealistically close to each other. Bubble-like objects represent voids that need to be filled. Clashes and voids are shown red as usually natural proteins should not have any. Small, pulsing yellow spheres highlight exposed hydrophobics, sidechains that would be better placed on the interior of the protein. Blue undulating bars between atom pairs show hydrogen bonds, beneficial structures that hold the protein together and therefore that should be maintained and increased in number.

The way in which *Foldit* was designed follows a clear principle: balancing players' needs and scientists' requirements through an iterative approach (Cooper et al., 2010a). Players value both the cause of sustaining scientific research and the fun which stems from playing the game. Scientists instead have to observe scientific requirements, as following fundamental physical and biological laws, to collect data that can be meaningful for their purposes. The regular interaction between the two groups on chats contributes to maintaining all parties' interests.

The community of *Foldit*, both players and scientists, is very open to discussion and listen to each other. Players can point out issues or preferences to the scientist, who in turn can communicate transparently to participants their objectives. Participants have shown a desire to help scientists and a need to have tangible outcomes, a few even had to deal with diseases caused by protein misfolding, directly or through known people (Curtis, 2015).

On the other hand, careful design of the simulation guaranteed significant outputs for scientists (Cooper et al., 2010a), both as data and tools. Indeed, the more the game is accurate in simulating the constraints of protein folding, the better are the models produced by players. Players were able to solve a long-standing protein structure determination problem, namely the M-PMV retroviral protease, which therefore provided new insights for the design of antiretroviral drugs (F. Khatib et al., 2011). Their modelling expertise also demonstrated useful in the design of artificial proteins. Through a series of puzzles, players were able to support remodelling the active site loops of a computationally designed enzyme that catalyzes the Diels-Alder reaction (Eiben et al., 2012). Players' ability to solve specific problems within the game has proved the power to outperform state-of-the-art computational methods. Scientists channelled this power through game mechanics to better understand and codify high performing players' strategies. Through in-game "recipes" to "cook" proteins, gamers have structured algorithms that achieve faster and more effective energy optimization. By studying how humans solve these puzzles intuitively in the game, researchers improved the algorithms used by protein folding softwares (F. Khatib et al., 2011).

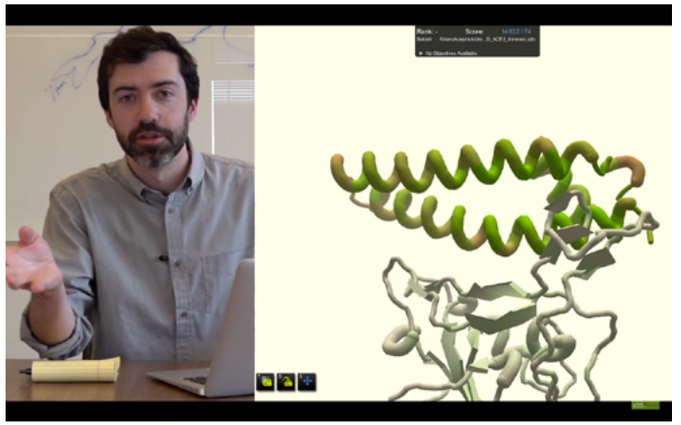


Fig. 25 One of the scientists behind Foldit (left) explaining the spike protein (right) contained in the coronavirus puzzle. See the video at: https://youtu.be/hS5g-2KhoSk.

This same year, developers introduced a new puzzle on coronavirus to support scientists in the research to face the pandemic of Covid-19 (The rundown on coronavirus, 2020; Coronavirus Spike Protein Binder Design, 2020). The puzzle required players to design a binder against coronavirus, namely a protein that binds to the coronavirus "spike" protein to block the interaction with human cells and accordingly the infection (Fig. 25). The news of the coronavirus puzzle on *Foldit* spread among the media and Google searches for *Foldit* reached a peak (Fig. 26). Although there is no data how many new users registered to the game for this precise reason, the interest that it aroused is remarkable and suggests that people are intrigued by such projects and the idea of contributing directly through a game.

Players' collaborative work forms a rich assortment of new strategies and algorithms and explores the space of possible search, besides the conformational space, which computational approaches don't (Cooper et al., 2010b). *Foldit* unveils the power of crowdsourced solutions as valuable support to scientific studies when conveyed through proper tools. Researchers (Cooper et al., 2010a, 2010b; F. Khatib et al., 2011) have acknowledged and credited players' contribution on related papers by addressing them as co-authors. In particular, the paper *Predicting protein structures with a multiplayer online game* reached 2740 article accesses and 1379 quotes on Google Scholar, which show a great media impact (Fig. 27).

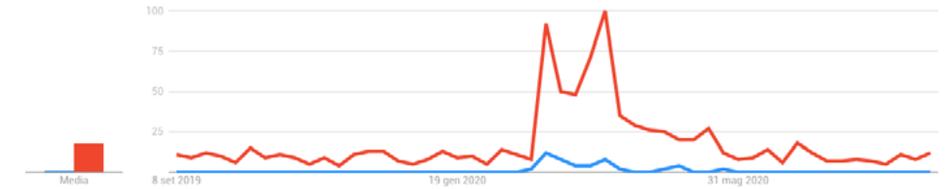


Fig. 26 Google searches trends on "Foldit coronavirus" (blue) and "Foldit" (red). The coronavirus puzzle was launched in February and ended in May.

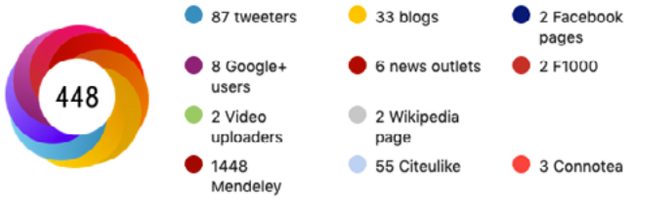


Fig. 27 Paper resonance of Predicting protein structures with a multiplayer online game *(https://www.nature.com/articles/nature09304/metrics).*

2.1.3 Taskification

Another interesting approach to fuse game and crowdsourcing was introduced by Prestopnik and Crowston as "game taskification". Opposite to gamification, taskification is the strategy of conceptualizing «the task as just one element or mechanic to be part of a larger (possibly much larger) game world» (Prestopnik and Crowston, 2012, p. 8). Therefore the designer firstly shapes the game following commercial game design principles and later embeds an external task to it. To strengthen the contrast with gamification, taskification can be defined as the use of non-entertainment tasks in game contexts. The tasks can either blend with the game and be critical to progress inside the game or be a subquest or minigame with which the player can engage freely.

Prestopnik and Crowston (2012) note though that taskification could not be as a realistic approach as gamification, however, they envisioned two possible solutions. On one hand, taskified games could be developed for profit, supporting scientific research by selling the game. On the other, mutating micro-payments system

in casual games, the tasks could be implemented as means to unlock game items, content, mechanics, or levels. *Project Discovery* is a significant case study for describing the implementation of this second approach.

Project Discovery on EVE Online

In 2015 CCP Games, the Reykjavik University, the Human Protein Atlas, and MMOS (par. 1.4.3) collaborated to release the *Project Discovery*[20], a minigame on *EVE Online*'s biggest test server (Leifsson et al., 2015). Their purpose was to complete analyzing and classifying the images in the subcellular atlas, a subpart of The Human Protein Atlas (HPA), though a citizen science game. What is peculiar about *Project Discovery* is the idea of injecting it into *EVE Online*[21] (CCP Games, 2003). It was the first time a real-world scientific task has been directly and seamlessly embedded into a mainstream MMORPG narrative (Sullivan et al., 2018).

EVE Online is a sci-fi massively multiplayer online (MMO) sandbox game. MMOs are online video games which enables a very large number of people to play simultaneously. *EVE Online*'s huge setting contains a total of 7,800 star systems that can be visited by players who can perform a plethora of in-game professions and activities, including mining, piracy, manufacturing, trading, exploration, and combat.

The game was an ideal setting for *Project Discovery* as a scientific project would not feel out of context in a science-fiction game, plus players are already used to face quite technical and complex in-game tasks (Leifsson et al., 2015). MMO games seem to be a perfect match for citizen science projects as the latter struggle with retaining contributors. On the other hand, according to Finnbogason, development manager at CCP Games, MMO games like *EVE Online* have solved the retention problem as they manage to keep people engaged. Plus, *Eve Online* community is dedicated and

20 https://www.eveonline.com/discovery

21 https://www.eveonline.com

Fig. 28 A shot during the gameplay of EVE Online.

willing to tackle demanding challenges. (Finnbogason in Barker, 2018). Moreover, the MMORPG like *EVE Online* meet many different play styles, providing social and challenging environments and enabling a great sense of immersion (Yee, 2006).

Similarly to *Foldit*, *Project Discovery* sustains scientific research on proteins. However, while in *Foldit* players have to design proteins, in *Project Discovery* the aim of the game is to correctly identify images of cells by analyzing the image's protein patterns (which are highlighted in green). Players can zoom details as true scientists to better recognize features that suggest appropriate categories to which they match. After they have selected the classes they regard as correct, they submit their responses and the game displays their result.

Since there are no right or wrong answers (the game's aim is to collect right answers from users), the feedback system displays answers highlighted in green if enough players have already agreed on the category chosen so that users perceive them as "correct". The system collects all classifications from players but only responses which overcome a certain threshold (i.e., at least

Fig. 29 The interface of Project Discovery *on* EVE Online. *On the left, there is the image that needs to be classified. On the right, there are the typologies of proteins. The ones marked in blue are those that the player selected.*

n players have to mark that answer) are deemed correct. Every submission gives rewards based on the player's accuracy rating (set at 50% as default, it can raise by classifying images correctly). Players obtain higher ranks through their actions and earn special in-game rewards (currency, resources, spaceships, badges …).

Sullivan and colleagues (2018) affirm players demonstrate to be able to notice unusual patterns. In particular, they expanded the set of known R&R proteins, which may ease professionals in understanding the biological function of this structure. They report gamer's high performance in recognizing less common classes outperformed even Loc-CAT, a machine learning application trained for the same classification task. On the contrary, Loc-CAT outperformed *Project Discovery*'s players in most common classes, especially in classes with lots of training data that provide Loc-CAT additional references gamers could not access. Even though this results in an almost equal performance of the two methods in general (accuracy level: Loc-CAT = 0.65, PD = 0.68), scientists exploited the specific players' skills to further train Loc-CAT, feeding their classification to it and increasing its performance. Experts in the HPA Cell Atlas still have higher performance than both, which suggests the system could still improve, but it is important to stress how *Project Discovery* helped scientists in improving datasets and therefore tools.

Fig. 30 The interface of Project Discovery *on* EVE Online, *showing the rewards collected after completing an image classification.*

Sullivan and colleagues (2018) record that 322,006 players of *EVE Online* played *Project Discovery*, producing around 33 million image classifications. Among these players, 59,901 had a desirable accuracy rating and contributed to 23.7 million high-quality image classifications and on average 6,846 unique players contributed each month with a retention rate of 32% and a rolling retention of 53% over the first 6 months. Overall, *Project Discovery* stood up to other in-game features over the same period and showed the power of citizen science efforts.

Project Discovery is still ongoing: in 2017 CCP Games, MMOS and the University of Geneva have launched a second iteration to help scientists hunting exoplanets. It is possible that it will be adapted in future to new citizen science projects: *Eve Online* executive producer, Andie Nordgren, seemed to imply the development team is looking for new dataset for the community to tackle (Nordgren in McAloon, 2017).

2.1.4 A comparison between gamification, SGs and taskification

The previous paragraphs have introduced the concepts of gamification, SGs development and taskification as approached to combine games and crowdsourcing systems. This paragraphs analyzes those approaches confronting them one another. This aims at defining the fundamental characteristics of each of them to support designers when deciding which approach to follow.

Taskification does not provide artifacts different from gamified systems nor from serious games. *Project Discovery* mini-game is a serious game per se. Anyway, it is worth considering the concept of "taskification" by itself as theorized by Prestopnik and Crowston (2012). All the case studies reported followed the hypothesis that fun delivered through games could attract and retain possible contributors to crowdsourcing tasks, and in this respect they achieve great results. However, starting from the same assumption, various product came out. The difference among them lies in the relation between the non-game task and the game itself. While gamification covers the whole task with a game layer and serious games development fuse the task with the game, taskification embeds the task in a small portion of the whole product, hiding in a huger environment with a great world-building.

Investigate Your MP's Expenses, *Foldit* and *Project Discovery* are representative of the strategies employed so far to mix crowdsourcing systems and games. The Guardian develop *Investigate Your MP's Expenses* to catch already highly motivated (and enraged) citizens: answering the need of active participation in "knocking down" fraudulent MPs, the system was bound to resonate. However, acknowledging the repetitive task could lead to boredom, the design team take the wise decision to shape it as a gameful experience to foster citizen's engagement. The same issue was solved similarly but in a different way by *Project Discovery*'s developing team. They had a defined user pool (*Eve Online* players) as well, but, unlike The Guardian readers, they had no direct interest in HPA's aims. Though, connecting the task to the game world, the value of participation was enhanced by in-game rewards. Even if in-game dynamics or fun, rather than attachment to the cause, were main motivators, *Project Discovery* demonstrated that gamers could produce high-quality data anyway (Sullivan et al., 2018). On the opposite, the task to be pursued in *Foldit* requires creativity and strategy, so the right kind of users would find it fun per se, and the choice of structure it as a game was due to increase its accessibility. Presenting the task as scientists perform it would be intimidating, but simplifying it and extracting the meaningful constraints of the system and translating them in games rules and mechanics make the task more understandable for a wider crowd.

The crowdsourcing aim, the users involved or to be reached and the case by case issues have to match stakeholders' resources and needs. Choosing the right strategy between gamification, serious games development or taskification is the first step.

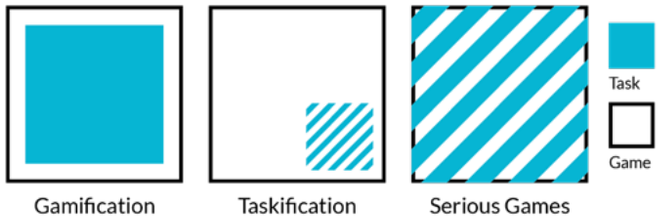


Fig. 31 A visual representation of the three approaches presented to combine crowdsourcing and games

2.2 Fun is all? Investigating underlying motivations

Par. 1.2.1 discussed the relevance of fun enjoyment in crowdsourcing. Games were proposed as tools to arouse these feelings and hence sustain engagement in the system. The case studies presented in the last paragraphs have shown how games can effectively support and enhance fun and enjoyment in crowdsourcing. Those are great examples of game-based crowdsourcing systems supporting the idea that fun derived from gaming is highly effective in gathering people around crowdsourcing activities and keeping them engaged. However, researchers have questioned if this was actually true and have investigated more deeply how gaming experiences could affect participation in crowdsourcing projects.

Bowser and colleagues (2013) investigated the effectiveness of gamification on technology enthusiasts from the millennial generation. In particular, they wondered if it could be possible to improve the experience by turning the citizen science application into a gamified mobile one. To test their hypothesis, they gamified *Biotracker*, an application that crowdsources plant phenology data. In *Biotracker* users can create or check in floracaches, virtual representations of real plants mapped in the application. A user has to visit, photograph, and identify a plant to create a new floracache, which clearly requires some plant expertise from the user.

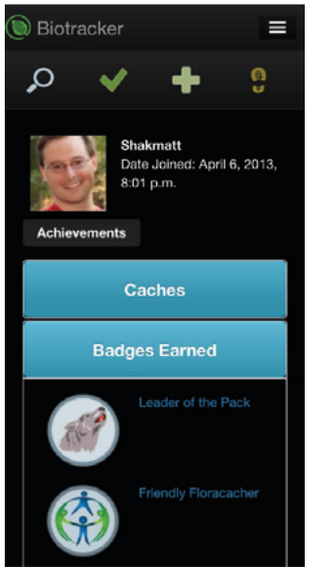


Fig. 32 User profile in Biotracker. *Badges show user achievements (Bowser et al., 2013).*

A user has to declare the plant's current phenological state ("all leaves withered", "full flowering", ...), comment on an aspect and photograph the plant. *Biotracker* inputs users with activities that rely on these two kinds of interaction, rearranged in various ways. Those data then are collected in the Project Budburst[22] database.

These activities are actually similar to video game quests, asking users to "create a cache of an invasive species" or "check in with another person". After completing an activity, the user earns a badge that is placed on their virtual profile pages. A leaderboard lists the top ten players who have checked into the most caches at any given time. Users who achieve the first spot receive a special badge to credit their commitment.

22 Project Budburst (http://www.budburst.org/) is a platform aimed at tackling specific, timely, and critical ecological research questions. Citizen scientists help by reporting careful observations of phenophases to the platform.

Bowser and colleagues' study involved 71 Millennial college students between 18 and 24 years of age. They had to attend a five-week unit on citizen science of 1.5 hours per week. They completed a questionnaire in the first week about their experience with citizen science. Most of them were unfamiliar with the topic and not even one had participated in citizen science. After four-week that introduced students to citizen science and gave them practice as participants, they finally tested the *Biotracker* application and evaluated it through a 15-question survey. Bowser and colleagues discovered that Millennials recognize wanting to have fun and some gamified elements as motivational elements to get involved in the project. Among the four items related to gamification into the survey that the students filled, "competing with my peers is motivating" and "earning badges is motivating" were significant items, while "doing my best is motivating" and "completing activities is motivating" were not (Fig. 33).

However, those gamified artefacts seem not to be enough to catch the interest of the most. Only 10 participants reported being somewhat or very likely to use the *Biotracker* in the future, while 37 of the users stated that they were very unlikely or somewhat unlikely to use in the future and 24 were indifferent. Without considering usability issues, users stated that they would not use it if doing so was not convenient. Students affirmed: «I would be pretty unmotivated to participate because the subject matter does not interest me to any real degree», «If I could use an app in a way that did not interfere with my daily activities I would absolutely participate in citizen science projects. I don't want to have to go out of my way to use the app», «When choosing a citizen science activity to complete, I would most likely choose one that is interesting, but that is not too time consuming». Broadly, users stated that they would be motivated to use the app if it was related to activities that matched their interests and did not diverge from their usual routines. However, participants who would use the *Biotracker* app were not more motivated by an interest in plants than those who would not use it.

Bowser and colleagues state that their hypothesis that gamifying a mobile application for citizen science could engage an additional user group result to be true. They affirm that a portion of Millenials may engage with a gamified citizen science application precisely because it is gamified.

Motivation	Total (n=71)	Would use Biotracker	Would contribute to other projects
Is fun	36 (51%)	U= 146.5 p< 0.01	U= 391.0 p< 0.01
Supports my interest in plants	26 (27%)	U= 287.5 p= 0.76	U= 502.5 p= 0.12
Helps me learn about plants and their environment	41 (58%)	U=134.0 p< 0.01	U= 501.5 p= 0.11
Contributes valuable scientific data	42 (59%)	U= 201.0 p= 0.07	U= 460.0 p< 0.04
Contributes to the public good	45 (63%)	U= 234.0 p= 0.21	U= 456.5 p< 0.04
Can connect me to a community of similar people	31 (44%)	U= 152.5 p< 0.01	U= 536.0 p= 0.26
Could be a fun social activity	31 (44%)	U= 144.5 p< 0.01	U= 472.5 p= 0.06
Doing my best is motivating	34 (48%)	U= 230.0 p= 0.19	U= 556.5 p= 0.37
Competing with my peers is motivating	34 (48%)	U= 180.0 p< 0.04	U= 462.5 p= 0.07
Earning badges is motivating	33 (36%)	U= 154.5 P< 0.01	U= 542.0 p= 0.28
Completing activities is motivating.	44 (62%)	U= 202.0 p= 0.06	U= 507.0 p= 0.12

Fig. 33 Independent t-test results of the research. In yellow, relevant results (Bowser et al., 2013).

Gamification appears to be an approach useful to motivate users. However, Bowser and colleagues record only 10 users out of 71 that stated they would use *Biotracker* in the future. Except for gamification features, these users were influenced by the idea that the application could help them learn about plants and their environments. Users' motivation may be more complex than what it seems.

Before this study, a previous one by Rotman and colleagues (2012) investigated motivations in using a previous version of

Biotracker which was not gamified. They collected data through an online survey and qualitative interviews. From these data, they identified many motivators. However, as a fine-grained analysis of the interviews revealed, these motivations were not equally salient nor concurrent. They stressed that the initial and more substantial volunteers' motivations were different from later and long-term motivations. In particular, the firsts were egoism-related reasons, while other kinds of motivations (collectivism, altruism, ...) show to effect engagement after some time, but to determine sustained participation. The motivators identified to initially drive users were:

1. familiarity with, or personal curiosity about a specific species, a landscape, or environment, or a process;
2. previous engagement in scientific projects, in schools or as a hobby;
3. an existing hobby that is closely related to citizen science (e.g., photography);
4. a career-building step for aspiring scientists.

To sum up, Rotman and colleagues (2012) record that volunteers' initial curiosity towards the project originated from personal interest and perceived individual gain. Recognition and attribution, feedback, community involvement and advocacy instead were motivations that affected only ongoing participation. Rotman and colleagues resumed their discoveries in a model that illustrates a dynamic model for the engagement cycle of volunteers and the related motivational pivotal points (Fig. 34) to stress the relevance temporal nature of motivation.

As shown, personal interest prompts initial involvement, but it needs care to develop into active collaboration and continued involvement. This interest collides with scientists' need for data for their studies, which pushes them to cooperate with volunteers. The synergy between the user's interest and scientists' need starts their initial collaboration. Then, the other motivational factors start affecting ongoing participation. The model that Rotman and colleagues draw stresses how crucial it is to explicitly recognize volunteers' motivations to reinforce the initial personal interest and keep them engaged with the system. Feedback, community

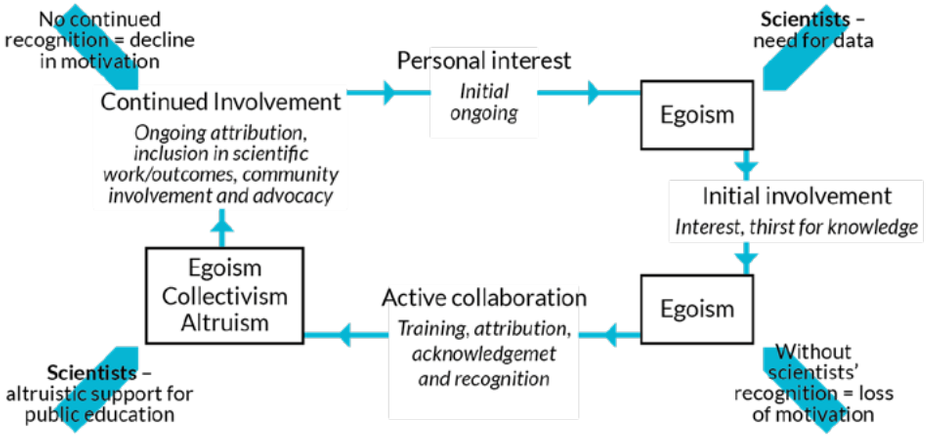


Fig. 34 A process model of volunteers and scientists involved in a citizen science project. Once addressed a personal interest, the correct motivational triggers can ensure long-term participation (Rotman et al., 2012).

involvement and advocacy show to influence participation only late in the cycle. At the end of a task or a project, citizen scientists reconsider their participation in the project relying mostly on previous experiences and less on future expectations. They reassess how the activity satisfied their various motivational factors, even secondary, and recognition and how they felt.In sum, personal interest plays a central role in sustaining the whole experience, since it inputs all the other motivations which lead to sustained participation.

Citizen science projects that employ games or gamification as well show this pattern. Several studies on the matter report that volunteers usually feel motivated to participate mostly because of a personal interest in scientific topics (Curtis, 2015; Iacovides et al., 2013; Raddick et al., 2013). Even the results from Bowser and colleagues' gamification of *Biotracker* reinforce this concept, by contrast. They wonder if gamification could engage technology enthusiasts from the millennial generation. Although their work suggests that Millennials enjoyed gamified features, it was insufficient to keep them engaged. When students were set free to decide whether or not to use *Biotracker*, gamification lost its power. Students argued they were lacking personal interest in the first

place, so they wouldn't feel motivated to use *Biotracker* in the future. Hence, engagement was not a valuable substitute for personal interest as an initial motivator.

In essence, the concept that the main driver is one's interest on a specific topic applies even to game-based or gamified crowdsourcing. It may appear that the idea derived from Brabham (2008) that fun is a valuable bargaining chip for crowdsourcees (par. 1.2.1) is wrong. Actually, these studies do not deny this possibility, but they set a need for reframing the terms upon which employing games might be effective to attract and retain users.

Firstly, so far the focus was on initial motivation. This is a highly important step in the engagement cycle and can support the attachment to the cause, but it is not the only one. Some motivations appear in later stages of engagement but are crucial to guarantee sustained participation. Rotman and colleagues have emphasized the urge for designing the whole experience to provide the correct motivations through the engagement cycle outlined in Fig. 34. These later motivations are as crucial as initial motivation in their model. The extent to which games can influence motivation should be analyzed distinguishing between initial and later motivations. Even if games do not impact on first motivation, they might affect ongoing long-term motivation. Moreover, games and gamified artefacts satisfy personal interest in different ways. This will be further discussed in par. 2.2.1.

Secondly, games might actually affect initial motivation. Indeed, if volunteers were keen on games, a well-designed game-based system could be a sufficient motivator to participate. *Project Discovery* (par. 2.1.3) is an example in this sense. Even if they were not interested in the kind of tasks to be accomplished or the general purpose of the project, they had a great personal interest in *EVE Online*. Therefore, the assumption that games can drive more people into crowdsourcing systems should be verified with the right user segment. Par. 2.2.2 digs deeper into these considerations and reflects on them.

2.2.1 How games affect ongoing motivation in crowdsourcing activities

The previous paragraph discusses how volunteers' in citizen science projects tend to be driven by personal interest more than anything else.. Many studies in the citizen science field indicate that gamification might not be so motivational in driving someone to contribute as other factors (Skarlatidou et al., 2019), hence it should not be considered as a main means to reach more contributors. However, it is necessary to keep into consideration that long-term motivations differ from initial ones. Sun and colleagues (2012) emphasise that initial motivation in participating in crowdsourcing activities are not even bound to predict sustained participation. Game employment might be not an initial motivation but a long-term motivation that affects the engagement cycle only later and guarantees users' retention. The use of games seems actually suited to improve user experience (Skarlatidou et al., 2019), which results in better retention rates as it will be explained through the following case studies of Iacovides and colleagues (2013) and Wang and colleagues (2020).

Iacovides and colleagues (2013) report that game employment seems to be a great technique to lengthen participants' involvement in citizen science activities. They analyze the influence of games – namely *Foldit*, a game to design protein folding (par. 2.1.2), and *Eyewire*, a game to map the brain – on motivation in citizen science projects through interviews. In particular, Iacovides and colleagues focus on two hypotheses (H1, H2): whether games might be effective (H1) in attracting users or/and (H2) sustain engagement. They discover that primal motivation stems from a previous personal interest in science, compliant to Rotman and colleagues' findings (par. 2.2). Although, Iacovides and colleagues noticed that games or games features seem powerful in maintaining volunteers engaged. In particular, they stressed that points and leaderboards push citizen scientists to do their best and keep participating day by day («the points don't motivate me but they do drive me further»). They identified team-play as a strong motive among the interviewees as well («if there were no group I

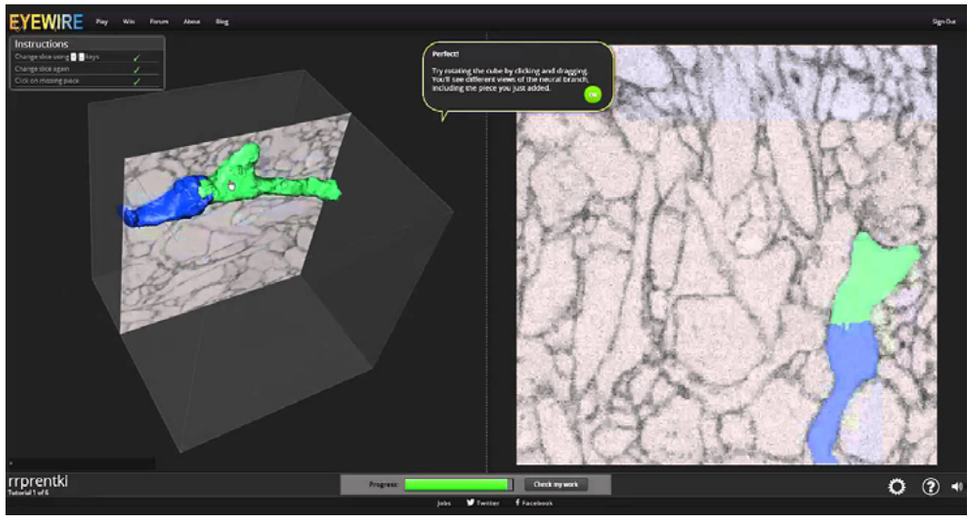


Fig. 35 The gameplay of Eyewire. On the left a three-e-dimensional representation of a neuron, on the right the two-dimensional image that the player has to map.

wouldn't be involved»). Therefore, games appear to be useless to attract volunteers but efficient to sustain the experience through their features.

Even later research on the acceptance of crowdsourcing games over time have registered this positive impact on sustaining motivation. Wang and colleagues (2020) traced users' experience of *KpopRally*, a music video-tagging game (Fig. 36), over a three-week period. In the game, people can both play solo or competitively on 5 round sessions. Players chose the best word to describe the video and get points according to percentage of users who agree on the chosen answer. The aim is to get the highest score possible and sub-challenges unlock achievements.

Wang and colleagues surveyed users after the first use, one week later and two weeks after the first use. They report the occurrence of time-based variations in users' experience. In particular, game employment significantly influenced acceptance at all tested stages. Enjoyable experiences appear critical in the continuous use of the system. Wang and colleagues affirmed willingness to be entertained is the main motivation in crowdsourcing games usage over time.

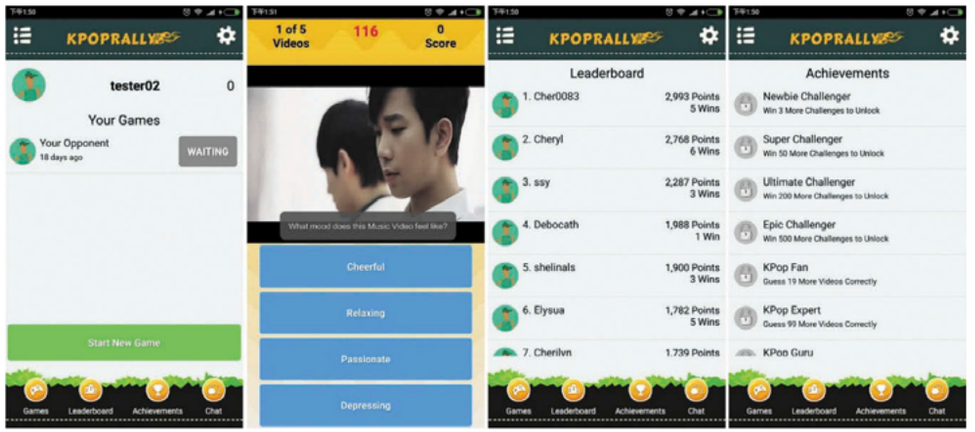


Fig. 36 Screenshots from KpopRally *application (Wang et al., 2020).*

Iacovides and colleagues' research hypothesis (H1, H2) derived one from a Rotman and colleagues' suggestion and the other from Paharia (2012). After their study (par. 2.2), Rotman and colleagues' (2012) proposed exploiting the intrinsically rewarding systems of games to (H1) attract people towards less appealing topics or (H2) increase their engagement in a topic that interests them. Paharia instead accounts that gamification is just a process that amplifies engagement (H2), and cannot create it. Based on his experience in design gamification for over 100 companies, he affirms that «the entity being gamified needs to have some intrinsic value already» (2012). Otherwise, gamification is pointless. The findings from Iacovides and colleagues appear to sustain Paharia's view even in the contest of game-based crowdsourcing, denying H1. Even later gamification literature reasserts this concept. Huotari and Hamari (2016) consider gamification as an aid to pleasing user experience as well. They stress that gamification is a process that does not ensure success but can only strive to enhance the experience. Games effectiveness in crowdsourcing systems appears to depend on the intrinsic value of the system.

However, the whole discussion relies on misused terms. Iacovides and colleagues (2013) consider *Foldit* and *Eyewire* both gamified artefacts and GWAPs. However, the two terms are conflicting. Gamification (par. 2.1.1) is a process that adds game features to a

system whose aim is different from pure entertainment in order to engage users more. GWAPs (par. 1.2.2) are systems designed for entertainment and at the same time for collecting valuable data from the gameplay. Those systems are transposals of tasks simulated by computers into games played by humans. Therefore GWAPs definition fits them well. They do not employ just some parts of games but they are complete games. Therefore, they are not gamified artefacts.

Stated that gamification acts as support, GWAPs define a distinct category that can follow different patterns. Drawing from one literature to support the other lead confusion in the field. Even though Iacovides and colleagues' findings (2013) are not invalid due to this misuse of terms, it is crucial to have this distinction clear. As explained by Paharia (2012), gamification does rely on the intrinsic value of the system and enhances it, acting like an additive layer on the product or service and improving the user experience. On the other hand, GWAPs, as SGs, are the system themselves. Users might perceive differences in the two kinds of products and act differently when dealing with a GWAP or a gamified artefact. As already said, successful case studies as *Project Discovery* suggest these differences exist. It might be even possible that they affect initial motivation, differently from gamification. Users keen on games might value GWAPs and SGs for their being games, and not for their second purpose. Therefore, they would engage with the crowdsourcing system for entertainment as their principal personal interest. It is crucial to acknowledge the limits and conditions in which GWAPs and SGs can achieve this objective.

2.2.2 Target matters. A discussion on user groups in game-based crowdsourcing

Many projects have been developed as games aimed at attracting a lot of people through their promise of being entertaining. Quite a few have been able to actually pursue this intention. Disregarding those which fail to be actually entertaining, it is proposed that the cause could be in a mismatch between the expected target audience and the reached one. Different researches have noted that even if some game-based crowdsourcing systems were designed to attract gamers, they failed and actually reached other audiences, such as science enthusiasts.

This is particularly true in the field of citizen science games. They may set their target to a large and inclusive audience but actually result in engaging mostly science enthusiasts. Iacovides and colleagues (2013) affirm that the sample of *Foldit* and *Eyewire* players that they analyzed (par. 2.2.1) represented only science-interested people, but not gamers. Users joined the games after they discovered them on science-related magazines, websites, TV shows and Twitter feeds. Out of 8 interviewees, only one claim to play other games, namely *Minesweeper*[23].

Curtis (2015) registered similar findings in her study on *Foldit* players. The sample that she investigated through an online questionnaire could be not well representative of the huge population of *Foldit* players as only 37 people replied. However, Curtis reports that the active playing population of *Foldit* is around 200/300 users. Therefore, the response rate is between 12% and 18.5%, which may be higher than in other studies on citizen science, hence are still relevant for the research in the field. However, the sample was self-selected and might contain individuals with strong opinions about the game. In any case, the results of Curtis' study are still interesting to report.

Fig. 37 shows the impressively low ratings that she registered for motivations as "I like puzzles" (11%), "I like computer games" (8%) and "It's fun" (5%). 59% of respondents stated that *Foldit* was the only computer game that they played. On the other hand, approximately 50% of them had already participated in other citizen

23 *Minesweeper* is a single-player puzzle video game developed originally in the 1960s. The player aims to remove hidden "mines" without detonating them. The shown fields display a number which indicates the number of mines in the nearby fields, so that the player can attempt to avoid the mines.

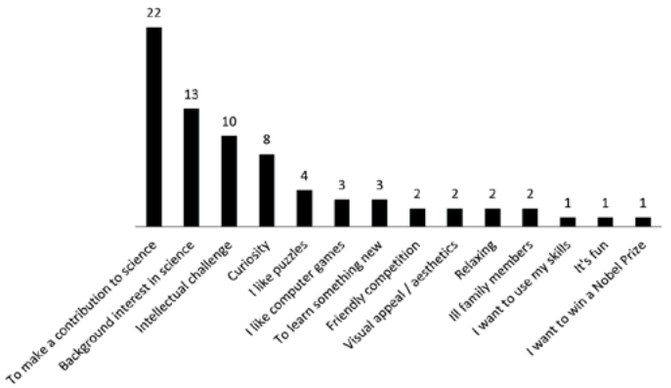


Fig. 37 Reasons why respondents participate in Foldit on a sample of 37 respondents. The number at the top of each bar represents the number of respondents providing that response (Curtis, 2015).

science projects. These data might account for a population segment highly bewitched by citizen science and quite disengaged with games. Their personal interest in participating was contributing to science (rated around 40%) and not fun.

Both Curtis' (2015) study and Iacovides and colleagues' one (2013) seem to support that game-based crowdsourcing is not intrinsically bound to attract all kinds of users. They identify the players population as a consistent user base of individuals keen on science. However, games can still play a central role in improving user experiences. Iacovides and colleagues' research have proved this concept, as discussed in par. 2.2.1, and Curtis' results sustain it too. She argues that the enjoyment of playing can be inferred from other comments, particularly those that refer to the community aspect of *Foldit*. The item "intellectual challenge", which reached 27%, might hide the enjoyment in tackling complex problems as well. Nevertheless, it is evident that there is an issue in the intention of reaching as many people possible and the actual user base. To better understand this phenomenon, researchers have studied users as distinguished groups. Their objective was to isolate needs and preferences for each group to target them effectively.

Prestopnik and Crowston (2012) investigate a set of citizen science projects that shared the same goal of classifying insects, plants or animals and were shaped more in a "tool-like" or "game-like" manner. They want to understand users' motivation in relation to these different designs and collect data through several focus groups.

They distinguish three user groups in game-based citizen science projects: 1) experts, 2) enthusiasts and 3) gamers. For the sake of simplicity, I group experts and enthusiasts as science-interested people, what Prestopnik and Crowston named "nature" participants. They suppose gamers to represent a larger pool of possible participants compared to science-interested groups. On the other hand, they may lack scientific knowledge, which increases the challenge of designing the system to be understandable for them, other than enjoyable. Prestopnik and Crowston suggest that, by designing to motivate gamers, it may occur to demotivate science-interested people and vice-versa. The two user groups demonstrate different drivers to engage with the system. Science-interested people want to explore their passions and interests, while gamers seek for entertainment and enjoyment.

Bowser and colleagues (2013) mention convergent and divergent motivations among the two user groups (Fig. 38). However, it is remarkable to notice they do overlap only in the domains of "community involvement" and "general socialization". Indeed, although "intrinsic rewarding" and "personal interest" appear to drive both user groups, they are umbrella terms that are not bound to

Motivation	Gamers	Cit. Sci. Volunteers
Fun (intrinsically rewarding)	X	X
Personal interest	X	X
Learning or education		X
Contributing to science		X
Contributing to public good		X
Community involvement	X	X
General socialization	X	X
Personal performance	X	
Competition with peers	X	

Fig. 38 A comparison of the motivation between gamers and citizen science volunteers (Bowser et al., 2013).

represent the same motivations. Personal interest sets the users' aim, i.e. entertainment for gamers and engaging with science for citizen scientists. Intrinsic motivation stems from the satisfaction of those different objectives. These different objectives define needs that diverge in the two user groups. It is therefore evident that designing choices might fit a user group and not the other.

Bowser and colleagues (2014) investigate those differences even further. They tested how the two user groups experience using *Floracaching* (Fig. 39), another gamified application supporting Project Budburst (par. 2.2). Like *Biotracker*, *Floracaching* requires users to report phenology data, but only about specific plants, which serve as virtual caches. This means volunteers need to visit the specific location of the plant to add information about it.

They surveyed 58 individuals and held optional focus groups. Among the participants, 22 self-identified as nature participants, volunteers interested in the science behind the project, and 36 as gamer participants, users interested in the gaming experience.

Fig. 39 "Check-in" to a Willow oak in Floracaching. Users may submit different types of data, such as the state of a plant (i.e., first flower, full leaf) or a photograph (Bowser et al., 2014).

They found gamer and nature participants were motivated by a sense of discovery, appreciation of outdoor experiences, education, social interaction and desire to contribute to science. However, they also evaluated the system differently on some core points. Gamer participants pushed to a more game-like design and enjoyed competition in any form. They call for additional guidance and feedback as well as values, e.g. teaching how to research and collect data. On the opposite, nature participants considered game feature accessory and see an inherent value in using the application to engage with nature. They appreciate competition when related to their scientific skills ("range of sites checked into" or "accurately keyed specimens") and not gaming ones ("getting the most the fastest"). Some nature participants even preferred to play alone and did not value social interactions. They preferred to discover autonomously the experience and use the application as support to scientific activities, and not as a guide. Bowser and colleagues conclude that design to capitalize on common motivators (as social interaction) and diversify the experiences for the two kinds of users could engage effectively both groups. In particular, the idea for multiple fruitions might support task decomposition, assigning tasks to different users accordingly to their skills.

Although gamers and citizen scientists can converge on some motivators, their perception of the application was basically opposite. Indeed nature participants were quite satisfied with *Floracatching* design, while gamers express a need for a more game-like experience and additive values. The usage of gamification can explain this mismatch. It has already been shown how gamified artefacts can only engage people interested in the core activity (par. 2.2.1). Since citizen nature participants were interested in the activity per sè, the application was appreciated. Gamers participants instead were not interested and prompt the developers with adding values to address their interests (i.e. learning a new skill, like how to do scientific research).

Although this does not mean it is impossible to design for many groups, it is crucial to know their desires and behaviours to design consciously. In particular, the collected complaints from gamers seem to stress that errors in targeting them are quite frequent.

The main of these mistakes is obviously not addressing their interest at many levels. One issue might be recruitment. As aforementioned, recruitment for citizen science games is often fostered through scientific magazines or platforms (Curtis, 2015; Iacovides et al., 2013). These channels are likely to fail reaching gamers, so it is no surprise they are few or even absent in the user base of science initiatives such as *Foldit* or *Eyewire*.

Another issue is the already discussed misuse of gamification to attract gamers (par. 2.2.1). Indeed, it may appear as an empty shell with no value to them and simply fail to address gamers' motivations. Developing a game for citizen scientists, for general crowdsourcees, for gamers, are all different scenarios that need different considerations. A proper study of the target is crucial to design correctly for the desired user groups. The previous studies have not investigated their target previously to the design of the system, but they have only registered their reaction to the final product. It is clear that to reach gamer users, a deeper investigation of who are gamers and which are their motivations is essential and would be undertaken in the next paragraph. This reasoning on user groups would be discussed again in par. 2.3 to propose taskification as a useful method to target gamers, making clear the value of this approach compared to gamification and SG design.

2.2.3 Understanding players motivation through Self-Determination Theory

A great way to explain players' motivation is through SDT, short for Self-Determination Theory, a macro psychology theory of human motivation and personality. It focuses on «the investigation of people's inherent growth tendencies and innate psychological needs that are the basis for their self-motivation and personality integration, as well as for the conditions that foster those positive processes» (Ryan and Deci, 2000a, p. 68). It addresses self-motivated and self-determined individual behaviour and the motivation behind human choices which do not suffer external influence.

SDT distinguishes between different kinds of motivation, namely intrinsic and extrinsic motives. Intrinsic motivation can occur only when one is inherently, personally interested in a specific activity, which they value for its novelty, challenge, or aesthetic value (Ryan and Deci, 2000b, pp. 59, 60). This is the case with sports activities. Extrinsic motivation instead relates to doing an activity to achieve some separable outcome (Ryan and Deci, 2000b, p. 60). This may be money or other kinds of material rewards.

Other crucial elements analyzed by SDT are basic psychological needs. They are energizing states that lead to health and well-being when satisfied or contribute to pathology and ill-being when not (Ryan and Deci, 2000a, p. 74). SDT identifies 1) competence, 2) relatedness and 3) autonomy as basic psychological needs that must be satisfied for people to experience an ongoing sense of integrity and well-being (Ryan and Deci, 2000a, pp. 74, 75).

The studies on video games Ryan and colleagues (2006) have employed SDT to understand players' motives and enjoyment. Excluding when it comes to playing video games as a work activity, video gaming does not produce extra-game rewards or approval for participants, therefore it is usually an activity driven mainly by intrinsic motivation. Ryan and colleagues proposed that people play video games because they are intrinsically satisfying for their psychological needs. They developed the Player Experience of Need Satisfaction (PENS) based on SDT. PENS includes:

- in-game competence, which measures the participants' perception that the game provided a challenging but not overwhelmingly difficult experience and enhanced efficacy;
- in-game autonomy, which measures perceived freedom and opportunities to do activities that interest the players;
- presence, which measures the sense of physical, emotional and narrative immersion in the gaming environment;
- intuitiveness of controls, which measures the experience of the interface that controls the character's actions in the virtual environment.

Ryan and colleagues (2006) ran four studies and gathered data by administering pre- and post-play questionnaires to the participants in the various studies. While the first three studies were

laboratory designs and the recruited participants could or could not be gamers, the fourth study focused on examining the motives of regular game players. It aimed at assessing persons actively involved in MMOs.

In this study, they even apply Yee's (2006) measures of player motivation, which are overarching, non-exclusive and focused precisely on MMOs. Yee's assessments include:

- achievement, namely interest in game mastery, competition and gaining power within the game;
- social, namely interacting with other players (helping and chatting, forming long-term meaningful relationships, being part of a group effort);
- immersion, namely interacting with the virtual environment (discovering, role-playing, customizing and escaping from real life problems).

Ryan and colleagues chose an online community active in discussing games and other Internet-based activities and invited their members to fill a survey intended for persons with past experience in MMO environments. 730 people responded to the survey. Ryan and colleagues recorded strong relations between the SDT-derived relatedness construct and Yee's social motive. They found a link between presence and Yee's immersion constructs as well. Autonomy satisfaction and Yee's achievement motive instead were not correlated.

Ryan and colleagues unveiled a relation between game enjoyment and intentions for future play and the SDT-derived measures of autonomy, competence, and relatedness need satisfactions, while none of Yee's variables related significantly. Hours per week was positively related to competence, relatedness and achievement. Intuitive controls support players' experiences of in-game competence and, in some game contexts, in-game autonomy, hence enhancing game enjoyment and preferences. Presence was related to need satisfaction. Greater perceived autonomy and competence enhance feelings of presence. In the same way, intuitive controls were positively associated with presence. In other words, presence steams from satisfying psychological needs and the possibility to focus on the game while not worrying about controls.

Ryan and colleagues demonstrated the relations between the three basic needs (competence, relatedness and autonomy) and games. In particular, single-player games (tested in the first three studies) satisfied autonomy and competence needs, while multiplayer ones satisfied all of them. Psychological needs' satisfactions, enhanced by intuitive controls, account for game enjoyment, preference for future play and partly (only competence and relatedness) hours per week. To sum up, it can be simplified observing that people's motivations may appear complex but in the end they play to satisfy their psychological needs.

Though psychology helps understanding the underlying motivations of people to participate, this knowledge is hard to translate into actual design. For this reason, the next paragraph digs into the tools which support the design of game-based crowdsourcing.

2.3 Converging players to new activities: research aim

Chapter one and two have reported a wide literature review on game-based crowdsourcing. It has been discussed what is crowdsourcing and why fun is one of the crucial motivators to participate. Hence, it is introduced the idea that games, as "fun-machines", can be a great support and catalyst for crowdsourcing activities.

Gamification, SGs development and taskification are all sensible ways to achieve this mixing of crowdsourcing and game design. However, motivations are complex and fun alone is not always enough. The difference between initial and ongoing motivation, introduced in par. 2.2 and developed in par. 2.2.1, has further uncovered the extent to which fun and games can affect the experience in game-based crowdsourcing. In particular, reasoning on target (par. 2.2.2) came naturally to understand how games can impact on initial motivation. Thus, par. 2.2.3 digs into the psychological drivers of gamers as users attracted by gaming experience.

Prestopnik and colleagues noted a difference in preference between science enthusiasts and gamers. They claim that there is «work to do in this area, especially building and studying citizen science (or other purposeful) games targeted explicitly to gamers, especially gamers of different types» (Prestopnik et al., 2017, p. 266). They envision that purposeful game design intersects with the creation of meaningful play experiences, the economics of the game industry and the data requirements of scientists. By doing so, it would reach great numbers of players that contribute to crowdsourcing activities. The idea of targeting defined user groups among gamers and directing their enthusiasm and energy to achieve real-world goals is extremely promising. However, it is little developed.

Taskification appears to be a trend that is slowly catching on. It was first theorized by Prestopnik and Crowston (2012). Gamification took hold around the same period but got great media coverage and success, contrary to taskification. The concept of taskification was not openly discussed and studied like gamification. Nevertheless, it expanded in practice. In 2014, Cancer Research UK published *Reverse the Odds*. The game required that players completed some scientific tasks to progress. The task is a marginal element of the gameplay, but it restricts players from continuing. As explained in par. 2.1.4, taskification differs from SGs development and gamification because it does not involve the whole gameplay or experience, but a small part of it. This is the reason why taskification can occur even after the game has been designed. In 2015, CPP and collaborators released *Project Discovery* on *EVE Online*, pushing their players classifying human proteins to collect rewards (par. 2.1.3). In 2020, Gearbox and collaborators released *Borderlands Science* for *Borderlands*' players. Participants to *Borderlands Science* could gain more rewards from the minigame accessible from an arcade on the starship Sanctuary III and use them in the actual game. The game requires that players complete block puzzles based on strands of DNA to correct errors made by computers organizing these data.

The idea of augmenting an entertainment game with purposeful gameplay is particularly interesting. Taskification is a method that enables this enhancement. It targets directly gamers to redirect them to a crowdsourcing activity. The experimentations in *EVE Online* and *Borderlands* show that such taskifications can be meaningful for players, beyond that beneficial for both the game industry and scientists.

In their agenda on gamified crowdsourcing, Morschheuser and colleagues claim that future research should «consider novel

trends in games design and crowdsourcing» (Morschheuser et al., 2017, p. 38). Taskification can be included in this kind of research. It is clearly a technique that combines game design and crowdsourcing. Moreover, it could become a novel trend in the near future: only a couple of gaming companies (CPP and Gearbox) have implemented tasks in their games. Moreover, CPP is iterating on *Project Discovery*, hence demonstrating that it is valuable for the company and it wants to keep it part of the game experience.

It emerged from the literature review and it has been explained that taskification is far less established than gamification and SGs development. Although, it is a valid way to exploit game design to engage users in crowdsourcing activities as well. This research hence focuses on taskification as a matter of enquiry. In particular, it investigates taskification as a design methodology to reshape games.

3 Research methodology

This research focuses on taskification as a novel and unexplored methodology in game and crowdsourcing design. It is a concept theorized by few (Prestopnik and Crowston, 2012) and still not widely shared. However, the use of it is recognizable in some practical examples. There are many aspects of it that could be investigated.

The research question (RQ henceforth) came naturally by observing the state of the art and the current literature: no one reported the process of design of taskification. Although it is constituted of and implies clearly different activities, taskification is not recognized as a different process from game design or gamification. Until now, designers and researchers still do not share the concept of taskification as, for example, gamification. They do not describe it because they treat it as a common game design process. However, turning an entertainment gameplay into a purposeful one is arguably the same process of designing a game. Therefore, the question which arises is:

How to design a game taskification?

3.1 Research question and hypothesis

The question presented above appears extensive, broad and limitless. However, it is interesting to dig into it, and further unpack it. The literature did not provide many insights on how to taskify a game in practice and, besides, on how practitioners reached their designs. There is a lack of theoretical understanding of the structure and components of a taskified game and consequently it is hard to approach the design of such a system. It is hard to imagine attempting to design a taskified game with so low guidance. Therefore, the actual RQ of this research is:

> RQ: How to **guide** a game taskification design process for crowdsourcing?

In practice, design processes are usually supported with tools. There are many, from facilitators like the IDEO Method Card[24] that enable the exploration of new design approaches to understand users' needs, to frameworks like Schell Games' Transformational Framework[25] that supports the design of games for social change. The hypothesis (H) was that a design tool could support a game taskification.

> H: A **tool** could guide a game taskification design process for crowdsourcing.

The aim of this research is therefore to test this hypothesis. The following paragraphs describe the methodology implied to address this H and understand if it can be an answer to the RQ.

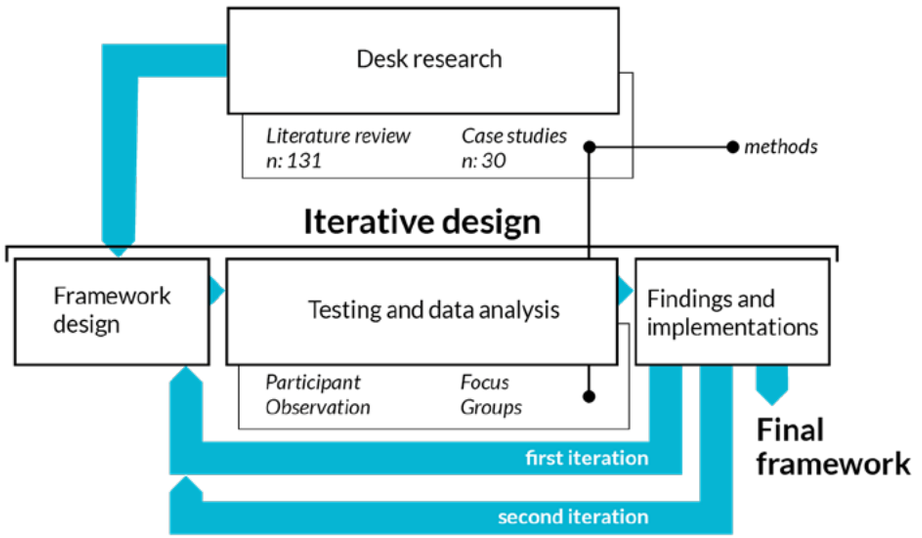
24 https://www.ideo.com/post/method-cards

25 https://www.schellgames.com/blog/building-a-transformational-game/

3.2 Iterative process

The research began with a desk research (Travis, 2016) which then nourished the design of the framework, refined through iterative design (Laurel, 2003). The research process consists of four steps, shown in Fig. 40 and described through this paragraph. Steps 2 to 4 were iterated three times to achieve the final result.

Fig. 40 The sum of the research process.



1. Desk research: literature review and analysis of the state of the art

First of all, an exhaustive literature review (n: 116 papers and n: 30 case studies) was carried out to understand better the field of intervention. The research started by analyzing crowdsourcing games, a topic that lies at the intersection of many different fields. The main sources spread from crowdsourcing studies and game studies. Games for Impact emerged as a recurring theme along with the matter of research and was analyzed as well.

The literature review helped to identify various approaches employed to combine games and crowdsourcing. These approaches — gamification, SGs design, taskification — are explored in par. 2.1.1, 2.1.2 and 2.1.3. Three case studies referred to these approaches were collected to understand the state of the art and were deeply analyzed. Other case studies were collected through the analysis of literature (n: 4) and the analysis of citizen science projects (n: 23).

The literature review revealed that a broader concept was better suited to describe all these phenomena and the term game-based crowdsourcing seemed to serve the purpose properly, which was hence always used during this thesis. Moreover, taskification appeared as an understudied phenomenon, hence it was further investigated.

A lack of methodologies to design taskification compared to SGs design and gamification was noticeable. This lack prompted this thesis' RQ, namely: "How to **guide** a game taskification design process for crowdsourcing?". It arises the hypothesis that a tool could provide the support needed in the design process.

The research hence focused on collecting information from liminal fields of research to fill the gap and develop a design framework for taskification.

2. Design of the framework

Once gathered enough information through the literature review and the state of the art, a draft of a conceptual framework was designed. The framework aims at providing enough reference points to support in taskifying a game without previous trai-

ning. It was developed to be a guiding tool which could contain all the essential factors of a taskified game. Its design builds on the knowledge gathered from the literature review, including the findings obtained from previous crowdsourcing and game design frameworks. Moreover, this knowledge is augmented through the analysis of guidelines related to game-based crowdsourcing systems. These guidelines were collected and clustered to identify recurrent design themes that were hence integrated into the framework structure.

3. Testing and analysis of the results

At this point, the research moved to the testing phase. The test aimed at evaluating the framework ability to guide the taskification process. Hence, it was employed in a series of pilots (n: 3) within a workshop in which participants (n: 9) had to taskify a game. During the workshop, participants had to choose a game and a citizen science project from a limited set of options and then combine them by designing the taskification. Participants analyzed the game structure and the crowdsourcing activity and designed the integration of the taski into the game following the framework. The workshop structure, its activity and the demographics of participants are deeply detailed in par. 4.2.

During each workshop, qualitative data was collected to understand if and how the framework supported participants' design process. The qualitative methods implemented to produce these data were participant observation (Jorgensen, 1989; Spradley, 1980) and focus groups (Creswell, 2009; Laurel, 2003). I conducted a moderate participant observation (Spradley, 1980) which balances between "outsider" and "insider" roles, allowing me to engage in the process as well as detach myself from the activities and observe objectively. I participated actively as moderator to explain the framework and to provide time limits for the activities, so that participants could go through the whole process in about eight hours. I did not interfere in participants' design choices but I answered about the framework and the workshop sturcture questions when requested. At the end of the workshop, I conducted a semi-structured focus group with participants to let them the possibility to express their opinion on the strong and weak points of the framework.

Then, the data collected in each pilot has been carefully analysed. I reported observed behaviours and comments from the focus group. I cross-checked the notes that I took during the process with the audio registration of the focus group to avoid information loss. The analysis aimed at turning data into organized information that was used to notice patterns and relevant factors which could support the research hypothesis or provide other insights.

4. Findings and implementations

After the analysis of the data, I collected the findings about the strong and weak points of the framework. Some issues emerged from the data analysis which needed a redesign to support the design process. Although the framework was never reshaped, I harness the knowledge acquired from each pilot to refine the way it was proposed and explained to participants. After the new implementation (going through step 2 again), the framework was tested again in a new pilot (step 3), to end again in step 4. There were three pilots and two iterations, then I collected the final results of the experimentations.

4 Testing and results

This chapter presents the design of the framework for taskification design based on the analysis of the literature review and the state of the art, its testing and subsequent implementations. Par. 4.1 describes the definition of the framework, explaining the connections between the different elements in the schema and the theoretical background that supports them.

The following paragraphs describe the results coming from the application of the framework in three pilots, and as many workshops, where the data was collected through participant observation and focus groups. The pilot and workshop structure are detailed in par. 4.2.

The three workshops occurred on the 9th (team A), the 11th (team B) and the 18th (team C) of July. The workshop lasted around eight hours, including a lunch break and the focus group. Team A and team B participated in the workshop online, while team C met in person.

4.1 Defining a framework for game taskification for crowdsourcing

Frameworks tailor the design process to the specific issues of the matter at hand. They present the core elements to design and their relationships in a clearly structured way. It is possible to design good game-based crowdsourcing systems with well-known design processes, like the agile one (e.g. see Ridge, 2011). However, frameworks sum knowledge related to the specific field of design or bridge various fields to form interdisciplinary tools to sustain the process.

Frameworks are an underdeveloped segment of tools in the field of game-based crowdsourcing, where the attention is mainly drawn on gaining practice-base knowledge. Indeed, ractitioners and researchers usually rely on experimental approaches that are ad-hoc for the matter of specific studies (Brito et al., 2015). The specific literature on crowdsourcing and game design provides several frameworks which can be translated in the field of game-based crowdsourcing systems and support their design.

Based on the wide literature review conducted, I designed a framework to taskify games – where the aim is to provide designers with clear guidance for building purposeful activities that exploit crowdsourcing logics within games. As there is little knowledge of

taskification, I relied on established knowledge from close fields, namely crowdsourcing and game design. Taskification combines an entertainment game with purposeful activity. In this sense, it is similar to the development of SGs, which are hence also considered in the design of the framework.

For this reason, the framework grounds its fundamentals and builds on two established frameworks, Simperl's (2015) framework for crowdsourcing design and the MDA (Hunicke et al., 2004), a well-known tool for game design. Additionally, the framework relies on the SG-related concept of *diegetic connectivity* (Lane and Prestopnik, 2017 – the concept is detailed in par. 4.1.3) to balance entertainment and tasks. Finally, the framework also benefits from a systematic review of guidelines from various literature on game-based crowdsourcing that highlighted other fundamental aspects of these systems; aspects which have been implemented in the framework design as well.

4.1.1 Simperl's framework for crowdsourcing design

To the matter of crowdsourcing, Simperl's framework (2015) appears brief and clear. She reflects on the basic mechanisms of the crowdsourcing process and draws her framework upon four design dimensions:

- what to crowdsource, a mapping of the high-level goal addressed through the completion of the tasks;
- who is the crowd, that is reached through open call but practically is restricted by the crowdsourcing platforms, advertising channels, knowledge and skills prerequisites;
- how to crowdsource, considering points 1-2 the requester has to design and execute the tasks, define assessment criteria and how to consolidate individual contributions into a result;
- how to incentivize, choosing the right mechanisms and refining them to affect the number of contributors and their level performance and engagement.

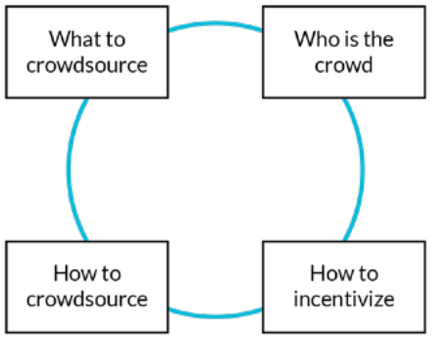


Fig. 41 The fundamental elements of crowdsourcing design according to Simperl.

Simperl explains these points further.

"What to crowdsource" may seem obvious when one decides to develop a crowdsourcing system. Poor resources are often the reason to deploy crowdsourcing. The crowdsourcer accounts that crowdsourcing is more effective in terms of time, costs, or quality of results than other options. However, the aim of the crowdsourcing activity is often too general to be turned one to one into an open call targeting a wide unknown group of people. Shaping the high-level goal of the project into specific tasks is crucial. In practice, this handles two issues: 1) what contributions to expect and 2) what to present to the crowd.

Contributions can respond to either an open or a closed task. Open tasks may suffer its subjective nature. Since it is not clear what answers are correct, results may be hard to use. Contrary, closed tasks require that the crowdsourcer have high control over the matter to crowdsource. The crowdsourcer has to know many elements in advance to set strict boundaries to the task so that results are homogeneous. However, if the initial constraints are somehow wrong or incomplete, the results of the crowdsourcing activity will reflect that. In general, open and close tasks have pros and cons that have to be carefully considered.

The next step is to care for the presentation of the task to the crowd. The system should display relevant items and tools to the crowd through its interface. When designing the interface the aim

should be to balance the time necessary to complete the task with the knowledge and context needed to do it well.

While dealing with contribution and presentation, the crowdsourcer should weigh details and wonder about the consequences of the choices that they make. They can even consider setting a number of tasks for different crowds with different levels of expertise or including the crowdsourcing activity a greater ecosystem that mixes it with in-house expertise and automatic tools.

"Who is the crowd" is the dimension which wonders about the target. Anyone can ideally be part of the crowd. However, many elements influence who actually participates in crowdsourcing. Certain types of crowds can be relevant if not even necessary to achieve a specific goal. The crowdsourcer will initially have very little knowledge about the crowd's skills, availability, and willingness to contribute.

It is important to know the elements that can affect participation from the desired crowdsourcees in a crowdsourcing project. Specific tasks that require skills and know-how which may cut down the number of possible contributors. Alongside, the way the task is crowdsourced, which is explained later in the "how to crowdsource", may influence participation as well. Apart from task-related issues, mediums employed to reach to the crowd play a significant role. The platforms chosen to support the project and the channels used to promote work as a filter to possible contributors. If a crowdsourcer wants to gather a highly educated crowd to tackle a hard task, he can set the project on InnoCentive or advertise its initiative through academic channels.

Crowdsourcers have a wide variety of methods to direct the desired users to join their crowdsourcing project. Using a particular platform, learning to predict the crowd members performance from their previous interactions and aligning incentives and motivation are all great ways to counterbalance the lack of control that the crowdsourcer can have on the arrangement of the crowd.

"How to crowdsource" is the dimension which accounts for the shaping of the task. After defining what task ("what to crowdsource"), one has to determine its granularity, transparency and validation.

The level of granularity differentiates macrotask and microtask. Macrotasks are meant to be outsourced as they are to one or more contributors. Open innovation, challenges, participatory government, or virtual labour marketplaces usually employ this schema. The assessment of the macrotask can occur manually (done by the task owner or by a panel of experts), openly or privately, through an algorithm or community votes. Microtasks are smaller chunks of the task that can be executed independently. process. While microtask might be quicker to solve, to reach high efficiency it may be necessary to translate the initial task into complex workflows. Developing workflows may require additional design effort, plus they need high crowd coordination.

The level of transparency distinguishes between explicit and implicit crowdsourcing. Explicit crowdsourcing relies on professional crowdsourcing platforms, social media and other PR channels to solicit contributions. Implicit crowdsourcing implies that the crowd does not explicitly participate to solve the tasks. GWAPs or for participatory sensing are examples of implicit crowdsourcing. In these systems, users play games or collect and share information via mobile devices or other sensors and the crowdsourcer uses the data that the users produce for their own purposes. Implicit crowdsourcing often induces a change in the incentives schemes. Indeed, participants may not need motivations to join as long as they "do not mind" the additional, potentially intrusive crowdsourcing tasks while they go about their own activities.

Finally, validation is a core element of all crowdsourcing projects. Indeed, contributions may not be usable as produced by the crowd and the crowdsourcer will have to assess them. Microtasks outcomes are usually meant to be easy to assess automatically. Redundancy is a widely employed approach. It aids to identify the answer which is most likely to be correct by using a weight metric (e.g., majority voting, previous performance of crowdsourcees).

"How to incentivize" is the dimension which relates to participants' motivations. The people contributing have their own motivation and behave according to it and to the incentive mechanism employed by the crowdsourcer. This mechanism can be classified as love, glory, and money.

Love refers to scenarios in which the crowd considers the tasks intrinsically enjoyable or rewarding. *Glory* is the status that comes with their involvement. Both love and glory imply that participation is voluntary and there are no profits. They are valuable approaches since they save costs and can attract a large number of contributors, at least for a while. However, they may be hard to replicate for all kinds of tasks. In particular, repetitive, unpleasant tasks or unrelevant for the audience are difficult to satisfy through love or glory. Contrary, reward models are often easier to control and study, although they do have their own issues. The crowdsourcer has to choose what to pay for and how much, which both affect the behaviour of the crowd. However, a rich set of motivations drives crowdsourcees and rely on financial incentives only might waste great chances to increase engagement and performance.

Simperl affirms that GWAPs and gamification are an interesting case in studying incentives. They rely on the assumption that adding game elements to the task raises crowdsourcees' motivation (par. 2.1). She argues yet that developing a good GWAP is often more critical than what may seem. On the one hand, some tasks will keep being more appealing and accessible to gamers than others. On the other hand, the game functioning is equally important. The crowdsourcer needs some knowledge of the task domain to implement game elements like feedback and levels. To build levels, the crowdsourcer has to be able to distinguish between easier and more challenging tasks among the ones they submit to the crowd. Developing a feedback system instead implies that they can rely on a gold standard, which takes time and resources to build, to confirm to users what they did correctly or not. In any case, the players constantly demand additional features to satisfy their desire for challenge. A one-off crowdsourcing project, for example, would not benefit from such a complex setting to run.

Simperl's framework provides precise dimensions for crowdsourcing development. Simperl digs deep into the components of these dimensions and explains clearly the crucial elements to consider and design. It even provides some insights on game embodiment for incentivisation reasons. However, it does not focus on it, so it is worthwhile to discuss even a framework for game design.

4.1.2 MDA, a game design framework

As Brito and colleagues (2015) affirm, it is useful to consider game design principles to develop game-based crowdsourcing systems. A great framework in the field of game design is the MDA framework (Hunicke et al., 2004). MDA stands for Mechanics, Dynamics, Aesthetics. This well-known framework was thought to bridge the gap between game design and development, game criticism, and technical game research. Consequently, the MDA is handy for understanding games by decomposing, studying and designing both game designs and game artefacts.

The framework conceives games as artefacts rather than media. The idea is that the content of a game is its behaviour and not the media that streams out of it towards the player. This concept supports clearer design choices and analysis by framing games as systems that build behaviour through interaction. At the same time, it relies on the games' consumption process. Games are products created by designers and developers and consumed by the players. The MDA formalizes this consumption process and divides it into 1) rules, 2) system and 3) "fun" (Fig. 42).

These elements are then translated into their design counterparts (Fig. 42), namely:

- Mechanics, «the particular components of the game, at the level of data representation and algorithms»;
- Dynamics, «the run-time behavior of the mechanics acting on player inputs and each others' outputs over time»;
- Aesthetics, «the desirable emotional responses evoked in the player, when she interacts with the game system».

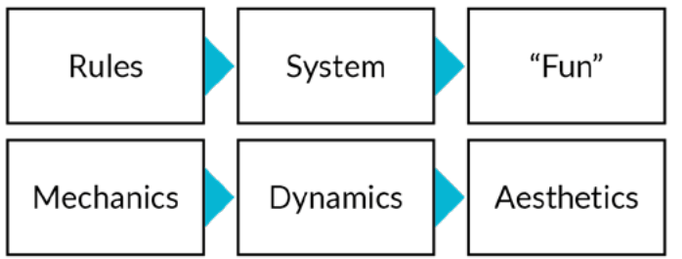


Fig. 42 Elements of the games' consumption process according to the MDA and their design counterparts (Hunicke et al., 2004).

These design counterparts provide different "views" (Fig. 43) of the game. From the designer's perspective, the mechanics define the dynamic system behaviour, which in turn produces particular aesthetic experiences. From the player's perspective, aesthetics establishes the tone, which arises from observable dynamics and eventually, operable mechanics. It is worthwhile to reflect on both perspectives while designing a game. Even small variations in one layer can cascade into the others. Moreover, focusing on one or the other perspective fosters either experience-driven or feature-driven design. Hunicke, LeBlanc and Zubek explain further how to use MDA to design experience-driven games.

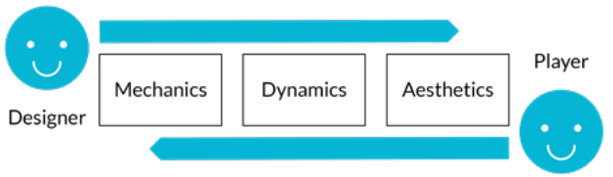


Fig. 43 Designer and player have different perspectives over the game (Hunicke et al., 2004).

Models for gameplay are chosen upon the aesthetics. Aesthetics represent the "fun", which can have manifold manifestations (the authors rely on LeBlanc's taxonomy of fun, explained in par. 2.1). Defining an aesthetic model helps describe gameplay dynamics and mechanics. For example, a "competitive" feature may be an aesthetic element of the game one is developing. Competition requires both adversaries and the desire to win in order to occur, which are elements of dynamics. Therefore, it essential to support adversarial play and provide clear feedback about who is winning and which are mechanics elements. Aesthetics helps articulate design goals and discuss game flaws, hence measure the development progress.

Hunicke, LeBlanc and Zubek even stress the importance of developing dynamic models that can predict and describe gameplay dynamics. They affirm that it is possible to avoid some common design pitfalls by doing so. For example, the model of 2 six-sided dice provides the probabilistic distribution of the random variable 2 D6 and aids discovering the average time it will take a player to progress in a game like Monopoly. Dynamic models may reveal losses of dramatic tension and agency at some stages in the gameplay so that designers can address and fix them.

Mechanics drive the dynamics. E.g., «the mechanics of card games include shuffling, trick-taking and betting – from which dynamics like bluffing can emerge» (Hunicke et al., 2004, p. 4). By adjusting the game mechanics, designers can fine-tune the games overall dynamics.

Mechanics, Dynamics and Aesthetics have to be tested and iteratively refined through playtesting. MDA implies iterations in the process to reach the best solution possible, which aligns all the elements. «Using the MDA framework, we can reason explicitly about aesthetic goals, draw out dynamics that support those goals, and then scope the range of our mechanics accordingly» (Hunicke et al., 2004). Hence, the MDA framework is a valuable tool to imply to reflect on games structure and integrate into them a new purposeful activity, which is the aim of taskification. However, the MDA does not focus on the design of purposeful games, hence it is crucial to analyze another framework for this specific aim.

4.1.3 Diegetic connectivity

Diegetic connectivity is «a mindset more than a model» (Lane and Prestopnik, 2017, p. 230). Its scope is balancing paramount elements that characterise entertainment systems, and in particular sustaining handling relationships between task, mechanics, and story. The task is the real-world aim and outcome of the system. The mechanics are the rules that limit players in their reach for game/task objectives. The story is a tool for serious game design to stimulate curiosity, produce meaningful connections, satisfy emotional needs and intrinsic motivation.

Lane and Prestopnik (2017) consider *diegesis* the element that connects task, mechanics, and story. Hence, it is useful to tackle several issues in SG design. Diegetic elements are elements that are integrated into the story world, while non-diegetic elements

refer to real-world elements. Diegesis enables to contextualize the task in the game world, provide meaningful rewards with in-game value production and stimulate task outcomes with mechanics (Fig. 44).

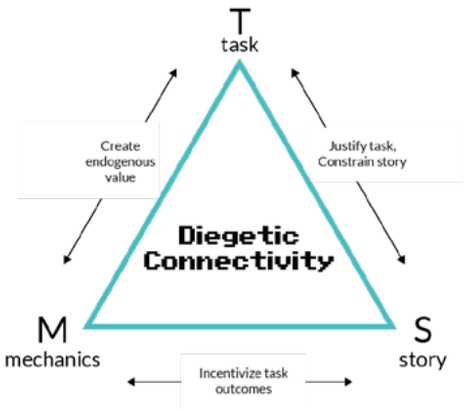


Fig. 44 Diegetic connectivity use story to connect various aspects of the game to improve player motivation, engagement and task outcomes (Lane and Prestopnik, 2017, p. 231).

The task can merge seamlessly into the game thanks to the story. Diegetic elements – like narrative structure, characters, the game world itself – can help designers justify the presence of the task into the game. Alongside, the task affects the story. The kind of task – classification, recognition, transcription and others – requires that players do an activity that has to make sense into the game setting. For example, in an investigative game, it would be reasonable to search valuable information into documents and ask players to take notes of them (in Pilot 2, the team explored this option, see par. 4.2.2). Lane and Prestopnik stress that aesthetics, visuals and sounds are particularly important to connect the task diegetically, convincing players that it is an integral, consistent part of the story (Wolf, 2013).

At the same time, the story can support the creation of endogenous value. Lane and Prestopnik (2017) explain that there are many rewards which are highly meaningful to the player because of the story. Some of them are new chapters of the story itself, player avatar upgrades, in-game economy that enables to purchase or trade items, action-oriented mechanics like exploration and combat that rewards with story-justified moments of excitement and intrigue. Lane and Prestopnik comment that links between story and mechanics abound in well-designed entertainment games. They suggest inheriting this approach from commercial-off-the-shelves games and applying it to SGs.

The story can bridge mechanics and tasks. «For example, story-driven quests could encourage players to exercise second language reading comprehension skills in conversations with other game characters» (Lane and Prestopnik, 2017). However, mechanics can enhance player uncertainty, curiosity, motivation, and engagement even with respect to the task when justified with diegetic connectivity. According to Lane and Prestopnik, diegetic connectivity creates chances to nourish task engagement by raising and then satisfying player curiosity.

Moreover, Lane and Prestopnik point out that diegetic complexity and task complexity should be carefully balanced. They propose to map them on two axes (Fig. 45) to understand the relationships between them and various design trade-offs resulting for each combination.

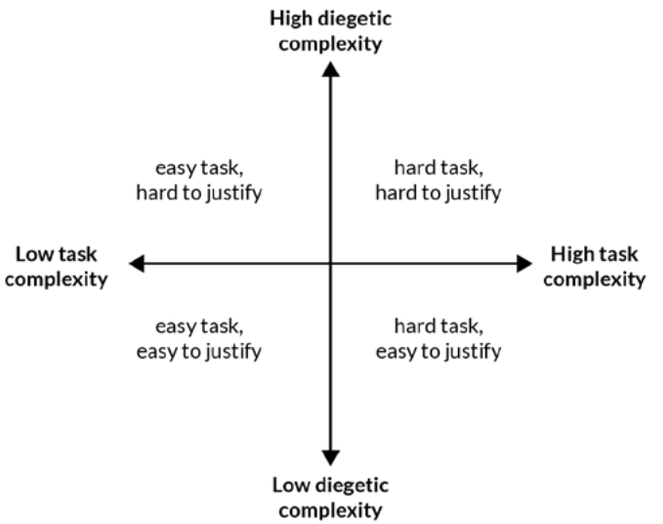


Fig. 45 Relationship task-game mapped on two axes: diegetic complexity and task complexity (Lane and Prestopnik, 2017, p. 237).

Lane and Prestopnik make some considerations on the combination of task and diegetic complexity based on their experience. Those considerations are the following:

- low task and diegetic complexity: it can be easy to justify the task and built strong relationships with diegesis, but this option seems to provide too little active engagement to be successful;
- low diegetic complexity but high task complexity: it may have greater potential for success, but it depends on its specific implementation, so it requires further study;
- low task complexity but high diegetic complexity: like the previous one, there may be high potential to be successful, but more studies are necessary to confirm this;
- high task and diegetic complexity: this option is highly risky because it may overwhelm players and lose their interest.

Lane and Prestopnik propose this approach for SG design, opposing it to other "gamification approaches". They stress that SG designers should care about the task in their designs more, claiming that the tasks can be something unique and special in their games and even motivators. In the light of this reasoning, diegetic connectivity helps to acknowledge both the motivating power and the limits of the task integrated into the design. It is not a tool for tricking players to do an activity or disguise the task. It supports the design of strong, story-oriented connections between task and mechanics. Lane and Prestopnik claim that this approach is «more akin to commercial entertainment than typical "gamified" experiences» (Lane and Prestopnik, 2017, p. 238). It immerses the player in the game world and maintains a cohesive and fun feeling.

4.1.4 Guidelines Review

Guidelines stem from all kinds of studies. This paragraph sums many formalized recommendations found in the literature to create a comprehensive, structured system of guidelines. Although, comprehensive may not mean completely exhaustive. The aim is to classify the existing guidelines to outline the relevant areas of design in the field of game-based crowdsourcing. For this reason, there may be some uncovered topics which need proper guidelines that research has still to formalize and miss in this review.

Moreover, it is important to explain some of the recommendations from Schrier (2016). In her book *Knowledge Games*, the author describes games whose objective is to knowledge production. Many of these games rely on crowdsourcing to achieve this objective. Schrier defines some guidelines which were meaningful for the scope of this research, so they were integrated. Some other of her guidelines instead were dismissed because they were too much related to their specific context of use.

Building on the analysis of the literature, I clustered the guidelines in macro-categories shown in Fig. 46. Each macro- and micro-category makes reference to a specific literature and in particular to specific guidelines, that are detailed below. Among the main ones: GWAP, gamified crowdsourcing, SGs or taskified crowdsourcing. This comprehensive view shows common points across the different game-based crowdsourcing systems.

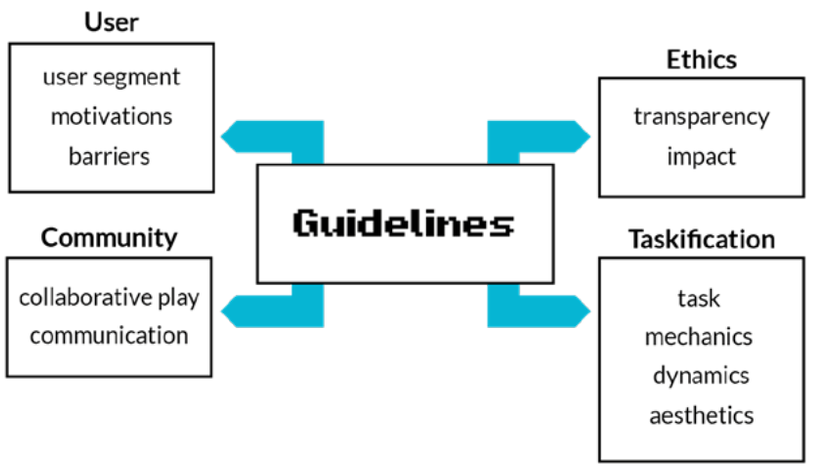


Fig. 46 Overview of the system of guidelines presented below.

The first group of guidelines regard the **user** of game-based crowdsourcing. Users are the engine of crowdsourcing systems, so they are essential for the sustainability of a project. It is crucial to define the kind of user to target (user segment), how to stimulate and motivate them (motivation) and what could retain them to join (barriers).

Guidelines related to **user segment** in game-based crowdsourcing systems include:

- «Consider how to target the right players for your game» (Schrier, 2016 on games for knowledge production);
- «Decide whom you want to participate—and whether your game is the best way to reach them or motivate them to continue playing» (Schrier, 2016 on games for knowledge production);
- «Sustainable gamification designs should also consider personal factors as well as orientation to work and games» (Morschheuser et al., 2017 on gamified crowdsourcing).

Guidelines related to **motivations** in game-based crowdsourcing systems include:

- Consider «the results of extant empirical studies [...] and theoretical frameworks on the design of game mechanics for crowdsourcing work (von Ahn and Dabbish, 2008), in order to incentivize right activities in the right form» (Morschheuser et al., 2017 on gamified crowdsourcing);
- «Help participants feel like they matter and that their participation is meaningful» (Schrier, 2016 on games for knowledge production);
- «Monetary incentives should be implemented cautiously in combination with gamification» (Morschheuser et al., 2017 on gamified crowdsourcing);
- «Support a variety of motivations, often simultaneously, such as a player's desire to help and contribute to the collective good and conduct prosocial activities, while also engaging in enjoyable gameplay» (Schrier, 2016 on games for knowledge production);
- «Appeal to different types of play styles and play motivations, as well as personal passions and interests» (Schrier, 2016 on games for knowledge production);
- «Consider how to support players' needs for practicing skills and gaining experience in a field, and enable people to show off their new skills or competencies in realms outside the game (e.g., cite players in research articles; provide certifications of skills)» (Schrier, 2016 on games for knowledge production);
- «Contemplate different levels of participation and how to balance quality of participation with quantity, and the differing motivating factors for each» (Schrier, 2016 on games for knowledge production);
- «Provide an appropriate number and type of choices, and enable the right level of control, so that people feel both needed and autonomous but not overwhelmed» (Schrier, 2016 on games for knowledge production).

Guidelines related to **barriers** in game-based crowdsourcing systems include:

- «Reflect on how cultural conceptions of play and leisure influence how people might interact with your game and its results» (Schrier, 2016 on games for knowledge production);
- «Weigh obstacles to access and participation in the game, whether related to resources (e.g., having the correct device or time to play) or to literacy and expertise (e.g., being game literate)» (Schrier, 2016 on games for knowledge production);
- «Many potential players do not read the instructions; it is therefore necessary that the game be intuitive and that learning be as short as possible» (Lafourcade et al., 2015 on GWAPs);
- «Players play all the longer as the game can be immediately accessed and instantaneously restarted. It results in a quantitatively enhanced acquisition of resources. Minimizing the delay and the number of clicks needed to launch a new game is a strategy to favor as much as possible» (Lafourcade et al., 2015 on GWAPs).

The second group of guidelines regard **community**. Community sustain participation, foster engagement and relatedness to the project. Both playing together (collaborative play) or more simply connecting and discussing with each other (communication) help build a strong community.

Guidelines related to **collaborative play** in game-based crowdsourcing systems include:

- «Ponder the broader cultural context of games and how this influences and is influenced by your game and its design» (Schrier, 2016 on games for knowledge production);
- Consider «implementing cooperative gamification approaches and affordances such as virtual teams and shared goals that might promote cooperative behaviors» (Morschheuser et al., 2017 on gamified crowdsourcing);
- «Teams contribute to involvement by allowing for a greater range of interaction between participants e.g. collaboration and competition between players. Not all participants may want to engage in group play however so it is important to ensure there is also a single player mode. Foldit for instance, allows for both» (Iacovides et al., 2013 on SGs);
- Consider «implementing cooperative gamification approaches and affordances with social factors such as rankings or public visualizations of individual achievements, should be preferred if the context allows the use of such motivational affordances» (Morschheuser et al., 2017 on gamified crowdsourcing);
- «By playing alone in his/her corner, the player gets quickly bored; the ability to visualize with the interface that other people play simultaneously causes a certain emulation[26], even if the player does not actually know them» (Lafourcade et al., 2015 on GWAPs).
- «Consider the ways people might (or might not) work together on a problem, activity, or task» (Schrier, 2016 on games for knowledge production);

Guidelines related to **communication** in game-based crowdsourcing systems include:

- «Support collaboration and other social interactions inside and outside the game, whether through sanctioned platforms (e.g., forums) or not designer-led but designer-allowed platforms (e.g., fan sites)» (Schrier, 2016 on games for knowledge production);
- «Consider how to use social interactions to support different aspects of a knowledge game, such as recruitment to participate, in-game training, or continuing to engage in the game» (Schrier, 2016 on games for knowledge production);
- «Reflect on how the game connects with other platforms or media and whether it should be part of a transmedia experience» (Schrier, 2016 on games for knowledge production);
- «Forums and chat tools help to facilitate a sense of community and continued engagement by supporting further interactions between participants. In addition, scientists could use these tools as a way to sustain involvement by providing information about progress and recognizing contributions, e.g. through regular project updates. While chat tools provide immediate contact with other people logged in, forums are a good way of

26 This is actually trickier than it seems. A Titfactor's study on emphasis frames in crowdsourcing tagging games (Kaufman et al., 2016) report that stressing the participation of an increasing number of fellow players may trigger social loafing. Social loafing is the tendency to exert less effort when activities are pooled (like in crowdsourcing) compared to when one is acting alone. This can even worsen if players feel their contributions are redundant with those previously offered, which is detrimental since redundancy can be a factor to prove output validity and performance. It may not occur when a participant feels deeply involved and, therefore, inclined to invest energy, but others may contribute less thinking that other participants compensate for their low level of contribution. There are some ways to avoid this effect, like reducing the perceived number of participants or reinforcing the value of users' unique perspectives. It is crucial to consider these effects and try to identify them during playtesting to fix the issue.

curating content/discussions» (Iacovides et al., 2013 on SGs); «not only consolidates the group by promoting conviviality, but also makes the players increasingly more expert. Furthermore, any means of communication at their disposal gives them the opportunity to provide feedback about the game and its modalities, which is important for designers» (Lafourcade et al., 2015 on GWAPs);
- «Design social interactions to encourage player-to-player learning, teaching, and sharing of expertise to spur knowledge production and problem solving»; «Encourage discussion, deliberation, and constructive argumentation among participants» (Schrier, 2016 on games for knowledge production).

The third group of guidelines regard **taskification** design. The design focuses on the basic structure of the crowdsourcing system (task) and of the game (mechanics), considering its actual functioning and interactions (dynamics) and emotional elements (aesthetics).

Guidelines related to the **task** in game-based crowdsourcing systems include:
- «Deliberate which specific problems your knowledge game seeks to solve and how to best design for and scaffold them» (Schrier, 2016 on games for knowledge production);
- «Consider how to best train and integrate novices into the game and to the field, and cultivate the necessary expertise to engage in the game» (Schrier, 2016 on games for knowledge production);
- «Allow players to test and retest hypotheses or replay different tasks to ensure reliability and validity» (Schrier, 2016 on games for knowledge production);
- «Be aware of how the complexities of a problem or issue are not perfectly able to be simulated and which aspects are lost or gained through a game»; «Deliberate what is lost and gained by simulating systems in your game»; «Reflect on how to best simulate real-world processes through games, particularly in terms of balancing accuracy, playability, and comprehension» (Schrier, 2016 on games for knowledge production);
- Consider «task characteristics[27] and especially the task complexity when designing gamification approaches for crowdsourcing systems» (Morschheuser et al., 2017 on gamified crowdsourcing);
- «Think about how to best represent problems in a game environment» (Schrier, 2016 on games for knowledge production);
- «Design for an integrated relationship between human being and technology so that the abilities of each are optimized»; «Consider how people and technology can most advantageously work together» (Schrier, 2016 on games for knowledge production);
- «Select the methodologies you are using to produce knowledge (and which ones you are not using), and consider how that choice affects your results and gameplay»; «Revisit how knowledge is defined and developed through your game as well as the assumptions underlying its design» (Schrier, 2016 on games for knowledge production);
- «Reflect on any biases and preconceptions and how they may affect your game»; «Consider how any data generated is subjective, and subject to biases»; «Consider how to express, question, and validate the trustworthiness of any knowledge produced in the game» (Schrier, 2016 on games for knowledge production);
- «Consider how you are framing the serious and fun aspects of the game and how this may influence the gameplay and the quality of the knowledge production» (Schrier, 2016 on games for knowledge production);

27 Task characteristics can even drive to innovative game designs. Tiltfactor (Flanagan et al., 2013) has attempted to incentivize players to avoid exclusively entering obvious tags in a tagging game to gather more specific ones. They decided to let players free to type that kind of tag but undesirable to do so as players would lose points if their opponent picked the same word. Punishment is an unusual mechanic for crowdsourcing games but it increases the challenging aspects of the game, especially in late-game. The design results to be proper to achieve the desired tags and engage players.

Guidelines related to **mechanics** in game-based crowdsourcing systems include:

- «Include appropriately challenging problems and obstacles» (Schrier, 2016 on games for knowledge production);
- «A system of points and ranking between the players generally makes the game more attractive: establishing and stimulating the competition between gamers drives some of them to spend more time playing, thus acquiring practice and experience. This makes them not only more productive, but overall increases the quality of the contributions. It is very rare for a player to be put off by the presence of a ranking» (Lafourcade et al., 2015 on GWAPs);
- «Players aquire experience by playing; they answer in an increasingly more relevant manner to increasingly more specific and more specialized questions. [...] It is necessary that the level of difficulty grows together with the player's experience» (Lafourcade et al., 2015 on GWAPs);
- «Team-play will encourage participation if participants view it as a meaningful activity in terms of contributing to the project goals. Similarly, points and badges should be used as way to support primary motivators, e.g. as a method of recognising contributions and allowing players to establish their expertise» (Iacovides et al., 2013 on SGs);
- «Consider how to fully integrate game mechanics with game goals and project goals so that the mechanics are not just plopped on top of the experience» (Schrier, 2016 on games for knowledge production);
- «Consider the extent to which the game needs to be fun and enjoyable—or if there are other types of meaningful game experiences» (Schrier, 2016 on games for knowledge production).

Guidelines related to **dynamics** in game-based crowdsourcing systems include:

- «Decide when to encourage competition, collaboration, or cooperation, with regard both to the game and project's goals, as well as to individual player characteristics» (Schrier, 2016 on games for knowledge production);
- «Support people's need to win and complete activities, even when a full solution may be unreachable or partial» (Schrier, 2016 on games for knowledge production);
- Be aware of cheating and vandalism, try «to uncover the sensible areas using pilot players, preferably skilled in computing» (Lafourcade et al., 2015 on GWAPs);
- «To encourage collaboration, jigsaw problems so that participants have slightly different tasks or responsibilities and so that players are more interconnected» (Schrier, 2016 on games for knowledge production);
- «There may be a form of appropriation of the game by the players. For example, a player may decide to focus on a theme that is of particular interest to him/her ("I am the expert in this field") and make a point to provide comprehensive and carefully controlled data. This form of specialization can set an example through the interactions within the group and bring other players to also feel they are specialists in a domain» (Lafourcade et al., 2015 on GWAPs);
- «Enable your game to have multiple and flexible paths to possible solutions»; «Enable players to provide alternate and diverse perspectives and viewpoints on problems, even ones that already seem solved» (Schrier, 2016 on games for knowledge production).

Guidelines related to **aesthetics** in game-based crowdsourcing systems include:

- «The game must present a ludic interest (is it not obvious?) at the interface level to attract gamers, but even more at the content level in order to keep them. The content must be able to renew itself with a small dose of repetition offering the player the possibility of amending himself/herself in the event of failure» (Lafourcade et al., 2015 on GWAPs);
- «Consider how to use and evoke emotion properly in gameplay, as well as its potentials and implications» (Schrier, 2016 on games for knowledge production);
- «Balance a need for fun with other design goals» (Schrier, 2016 on games for knowledge production);

- «Pique curiosity through incongruity of expectations, uncertainty, and lack of consistency (in a way that is relevant to the project goals and design needs of a game)» (Schrier, 2016 on games for knowledge production);
- «Use story and analogies to motivate players and help support in-game tasks, but also consider the implications of using stories» (Schrier, 2016 on games for knowledge production).

The last group of guidelines regard **ethics** within game-based crowdsourcing. It is crucial to remind designers that such systems have major ethical issues both regarding how they communicate with the crowd or the public (transparency) and they affect people's lives (impact).

Guidelines related to **transparency** in game-based crowdsourcing systems include:

- «Decide how you are using the players and their gameplay and to what extent any aspects of the game's design, data collected, or results need to be explained or made more transparent» (Schrier, 2016 on games for knowledge production);
- «Clarify any benefits or drawbacks for the researcher/designer and the player» (Schrier, 2016 on games for knowledge production);
- «Decide to what extent you will share any findings, data, or designs with your players and the general public» (Schrier, 2016 on games for knowledge production);
- «Deliberate the relationship between your game and the "real world." Demarcate boundaries (if any) between your game and beyond» (Schrier, 2016 on games for knowledge production);
- «Consider how games can help illustrate and make transparent, and also replicate and reinforce, institutions and systems that are often opaque» (Schrier, 2016 on games for knowledge production).

Guidelines related to **impact** in game-based crowdsourcing systems include:

- «Reflect on how your game cultivates critical questions, values, empathy-related skills, and ethical thinking about its use and the meaning of the knowledge produced through it» (Schrier, 2016 on games for knowledge production);
- «Consider how your design might be coercive and whether aspects of the design, or emergent play, are exploitative or alienating» (Schrier, 2016 on games for knowledge production);
- «Consider how you perceive and treat amateurs, and how this is reflected in the design of your game» (Schrier, 2016 on games for knowledge production);
- «Consider how you will protect, communicate, and/or manage the privacy of any contributions and actions made by, in, and through your game» (Schrier, 2016 on games for knowledge production).

Finally, I identified a **rule of thumb**, which can be the perfect guide when other guidelines fail to aid design choices:

- «Ponder how the game, tools, platforms, people, and context work together to distribute knowledge and support problem solving» (Schrier, 2016 on games for knowledge production).

4.1.5 The framework

Relying on the knowledge collected to draft the framework (described from parr. 4.1.1 to 4.1.4), this paragraph presents the final schema tested in the pilots. The framework (Fig. 47) consists of two macro-areas, representing respectively 1) the crowdsourcing design (top) and 2) the taskification design (bottom).

The area of crowdsourcing design contains Simperl's (2015) framework: "What", "Who", "How to crowdsource" and "How to incentivize". Since those elements are equally important, they are linked through a circle to represent their connection without defining hierarchies. The "Who" element is supported by the "User" element derived from the guidelines review.

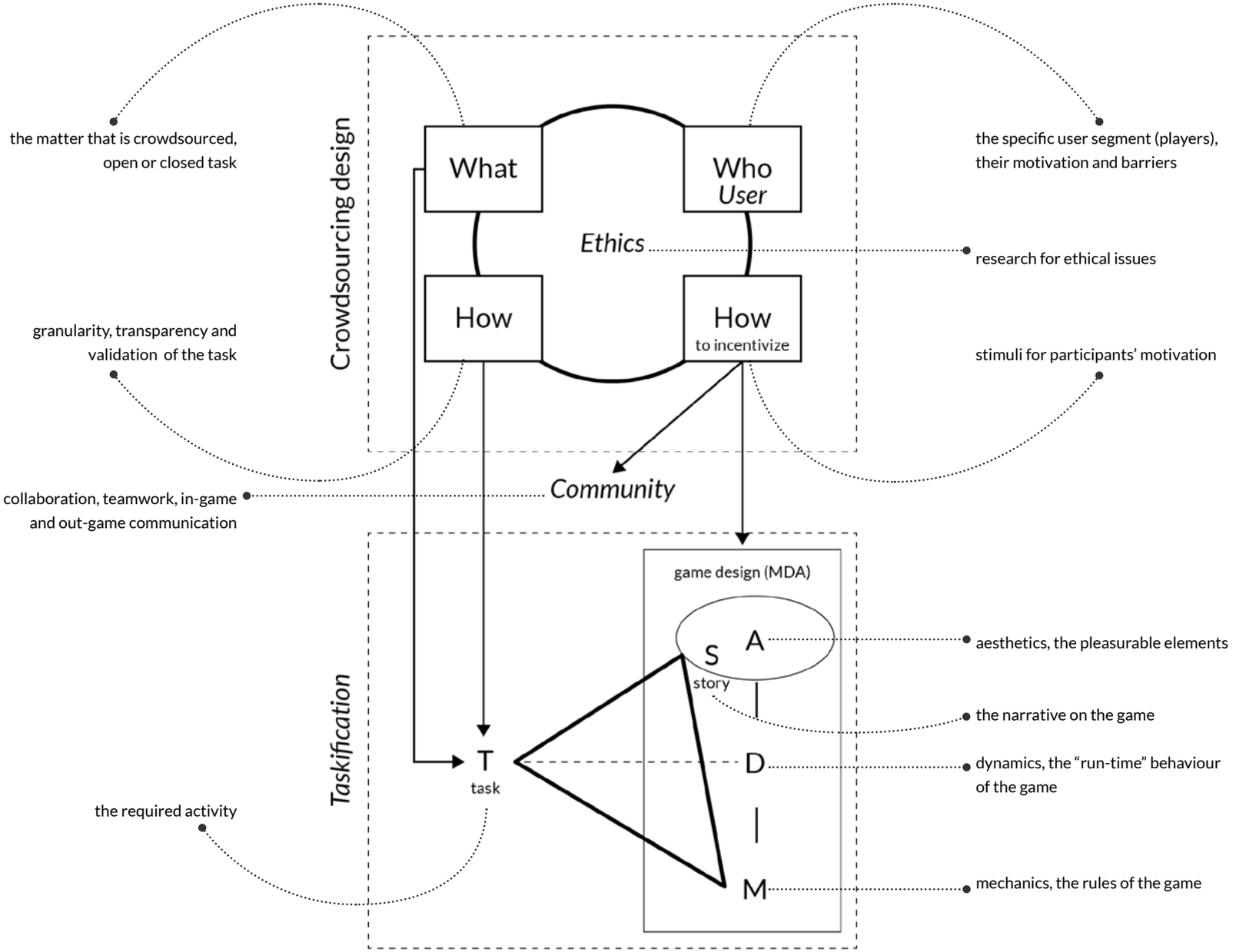


Fig. 47 Taskification design framework for crowdsourcing activities. In italics, the elements infferred from the guidelines review.

"What" refers to the matter that is crowdsourced, the high-level goal that the system attempts to address. The "What" guides the definition of the crowdsourcing task in terms of what contributions to expect and what to present to the crowd. The designers should define either an open or a closed task, providing relevant items and tools to the crowd through the interface, so that they can perform the task, and deciding to either split the activity in various tasks at different levels of expertise.

Along with the "What", the "How to crowdsource" support the definition of the task, in particular granularity, transparency and validation. The level of granularity sets if a task is a macrotask or a microtask. The first presents the activity as a whole, while the second can be only little pieces of a task. Transparency determines whether the task is explicit or implicit. In the case of taskified games, the task is always implicit, because the main activity for participants is playing. Finally, there is the need for a validation system in every crowdsourcing project. Contributions could be assessed by manual control or automatic tools, like algorithms who record redundant answers and confirm them true because of their quantity.

"Who", endorsed by the "User" element, aims at identifying the desired target to be participating in the crowdsourcing activity. While whoever can engage in a crowdsourcing project, not everybody does and for various reasons. It is important to reflect on the elements that can affect participations, both positively (motivations) and negatively (barriers) and to use the correct platforms and communication to reach the audience. In the case of taskified games, it would be proper to design a purposeful activity for a game whose players are the ideal target for the task to be crowdsourced. Then, it is crucial to understand how to communicate the value of participating in the crowdsourcing project.

Finally, "How to incentivize" is the element that focuses on the motivation of participants and how to encourage them. In the case of taskified games for crowdsourcing purposes, the stimuli to engage with the crowdsourcing activity is the game itself. This means that players participate because they consider the tasks intrinsically enjoyable or rewarding.

The "Ethics" element, another cluster of the guidelines, stands in the middle of the framework of crowdsourcing design. Ethical matters are essential throughout all the process, so it is important that they are always considered in all steps. Par. 1.4.2 has explored the complex topic of ethics in game-based crowdsourcing, showing how hard it is to handle. Designers need to carefully consider ethical issues that can arise from the employment of both crowdsourcing and games. The major aspects that arose in the guidelines review regarding ethics were the need of both transparent relationship between the player and the crowdsourcer and of careful analysis of possible impacts on society and individuals. This knowledge can be further supported by the themes deemed in par. 1.4.2, like unethical persuasion (Berdichevsky and Neuenschwander, 1999), exploitation (Kim and Werbach, 2016; Standing and Standing, 2018), manipulation (Kim and Werbach, 2016; Sandovar et al., 2016; Standing and Standing, 2018), power imbalances (Standing and Standing, 2018), deception (Zagal et al., 2013), physical and psychological damage (Sandovar et al., 2016; Kim and Werbach, 2016). The following are some examples of ethical issues to tackle considering each aspect of Simperl's framework:

- **What and how to crowdsource** should not be unethical, e.g. harmful for someone – training an algorithm to steal private data for example;
- **Participants** should not be vulnerable categories, like children, who hardly could understand if they have been manipulated, and should decide to participate to the activity consensually, knowing the terms of use, which should be transparent;
- The system should not **incentivize** to contribute by manipulating users and tricking them into exploitative cycles or using dark game design pattern (Zagal et al., 2013).

Moreover, taskification may be ethical as long as it aims at social innovation, like helping research against cancer or other diseases. If taskification was implied for commercial purposes and profit, it would be probably exploitative. Participants could spend their energies and efforts on behalf of industries which would not share any benefits with them. It is important to reflect on the use of taskification for crowdsourcing considering these aspects.

The area of taskification design combines the MDA framework (Hunicke et al., 2004) and the diegetic connectivity (Lane and Prestopnik, 2017). It links the task (T) to the MDA, creating a triangle in which MDA is a side of the shape and T is a vertex connected to each element of the MDA. This area details the elements from the upper part: "What" and "How to crowdsource" defines the task, while "How to incentivize" focus on how to engage users. In this case, engagement is stimulated by the game, in particular by its mechanics, dynamics and aesthetics – the elements of the MDA framework. Games produce fun experiences that attract people through their aesthetics elements: narrative, challenges, discovery, and so on (see par. 2.1 for a wider discussion). These experiences are supported by the game dynamics, namely its functioning system. This system in turn works thanks to the rule that composes it, i.e. the mechanics of the game. By combining and tuning these elements, games can engage players perfectly.

Alongside, even the community aspects of the game are important. "Community" is a topic that emerged from the guidelines review and is an element which sustains engagement and participation, hence it is a way to incentivize users to contribute. Community can either be accessible in-game, through chat or other communication systems as well as through collaborative play, or outside the game, in dedicated forums or similar.

Diegetic connectivity, an approach that connects purposeful activity with games narrative and mechanics, enables the connection between the task and the MDA. Diegetic connectivity enables the connection between the task and the MDA. Diegetic connectivity links the task (T), the story (S) and the mechanics (M). The story, intended as fantasy and narrative, is an element of the aesthetics according to the MDA framework (Hunicke et al. 2004). The mechanics instead appear in both the diegetic connectivity approach and the MDA. Hence, the task is linked to the MDA through aesthetics and mechanics. When designing a taskification, it is crucial to reflect on each of these relations. The challenge is to create new bonds between the elements of the MDA and the task without compromising the balance of mechanics, dynamics and aesthetics that the original game has already achieved.

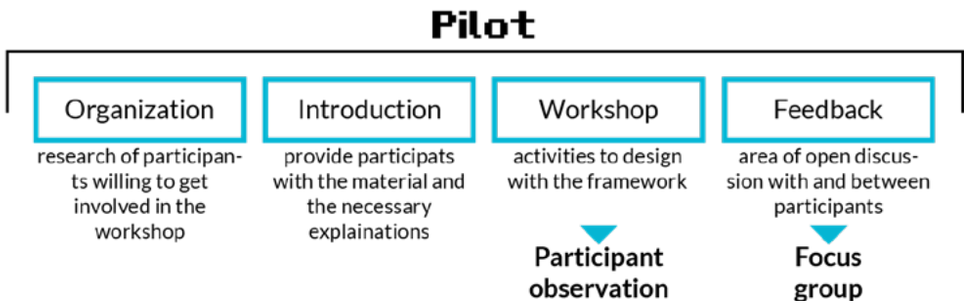
The framework aims at providing knowledge and structure to design taskified games for crowdsourcing applications. To understand if the tool was actually able to guide the design of such systems, it was implied in an experimentation composed of three pilots. Those pilots followed the design process of three teams of three people each who were asked to taskify a game with a citizen science project. The following paragraph explains the workshop structure and participants demographics.

4.2 Testing through pilots

The framework aims at guiding the design of a game taskification. To experiment with its capability to do so, it was employed in a series of three pilots. Participants in the pilots were asked to design a game taskification following the proposed framework. Fig. 48 show the pilot protocol, composed of four steps, namely organization, introduction, workshop and feedback.

The organization phase focused on finding participants for the workshop. The defined target for the testing were game designers. A taskification would require many different professionals along with game designers, like developers, researchers and crowdsourcing experts. However, due to lack of assets, the research focused on testing game designers' understanding of the framework. The majority of games on the Citizen Science Games website are products of universities or science-related organizations like NASA or Cancer Research UK. There are some cross-disciplinary teams, like ScienceAtHome[28], Balanced Media|Technology (discussed in par. 1.4.3), MalariaSpot[29], QuestaGame[30] and Citizen Sort[31]. Only a few gaming companies, namely Glitchers, CCP Games and Gearbox Software, have collaborated to design games to sustain scientific research.

Fig. 48 The pilot protocol, composed of introduction, workshop and feedback.



This analysis shows that researchers are far more often in charge of producing game-based citizen science systems than gaming companies. Gaming companies and their employees could reasonably lack the knowledge and expertise to design such systems. Therefore, the framework might be especially useful for them. The research focuses on game designers among gaming companies' employees because the framework is tailored for guiding game design. It does not consider programming, visual, sound or production issues. These are fundamental aspects of game development and game design should coordinate with them. However, for the scope of the research, the focus was game design and hence game designers.

A call for participants was launched among students and graduates from Politecnico di Milano and Università Statale di Milano. Nine participants between 25 and 30 y.o. responded to the call and participated in the workshop. Three were females and six were males. Participants were students (n: 2), employees (n: 3) or fresh graduates (n: 4). They all had previous experience in game design, either by attending game design classes during their studies (n: 7) or developing a thesis on the topic (n: 2). They were students or graduates in computer science (n: 2), automation engineering (n: 1), interaction design (n: 3), game design (n: 2) and communication design (n: 1).

28 https://www.scienceathome.org/
29 https://malariaspot.org/en/
30 https://questagame.com/
31 https://www.citizensort.org/

Team A	Team B	Team C
male computer science student	female automation engineering employee	male computer science employee
female game design graduate	male game design student	male interaction design employee
male interaction design graduate	female communication design graduate	male interaction design graduate

Fig. 49 Teams and their composition.

Three pilots were set based on participants' availability and considering some time to implement potential necessary improvements. Participants were arranged in three teams (A, B, C) of three people each. Fig. 49 show the composition of each group. Pilot 1 was on the 9th of July and involved team A. After a couple of days, on the 11th of July, team B participated in Pilot 2. Finally, Pilot 3 occured on the 18th of July with the attendance of team C.

The introduction phase established the grounding elements of the workshop, describing the concept of taskification and providing examples of taskified games. To clearly explain what is taskification and how it works, two projects were selected: *Borderlands Science* (Gearbox, 2020) and *Project Discovery* (CPP, 2015).

Borderlands Science[32] is a taskification in *Borderlands*[33], an open-world action role-playing first-person shooter video game. It is a mini-arcade puzzle game (Fig. 50, Fig. 51) that helps to map the human gut microbiome. Mapping this microbiome would support research on diabetes, depression, autism, anxiety, obesity and other diseases. The minigame rewards participants with in-game currency.

32 https://borderlands.com/en-US/news/2020-04-07-borderlands-science/

33 https://borderlands.com/en-US/

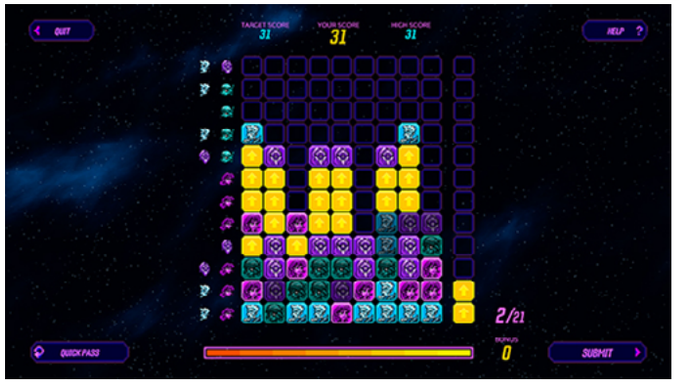


Fig. 50 The arcade that enables access to Borderland Science.

Fig. 51 The minigame interface. The player has to match similar tiles moving them up (the yellow tiles mark where tiles have been moved).

Project Discovery (explained deeply in par. 2.1.3) is a taskification inside *EVE Online*, a sci-fi MMORPG. Its current aim is to classify exoplanets through a minigame and provides in-game currency upon participation.

Then the introduction phase exposed an overview of the framework and explained its elements. In this way, participants could better understand the stages of the workshop and their relation to the framework. Finally, an icebreaking activity took place to help team members start interacting with each other.

Fig. 52 The structure of the workshop, defined by its main activity.

The workshop in its turn had four stages: definition, brainstorming, integration and synthesis.

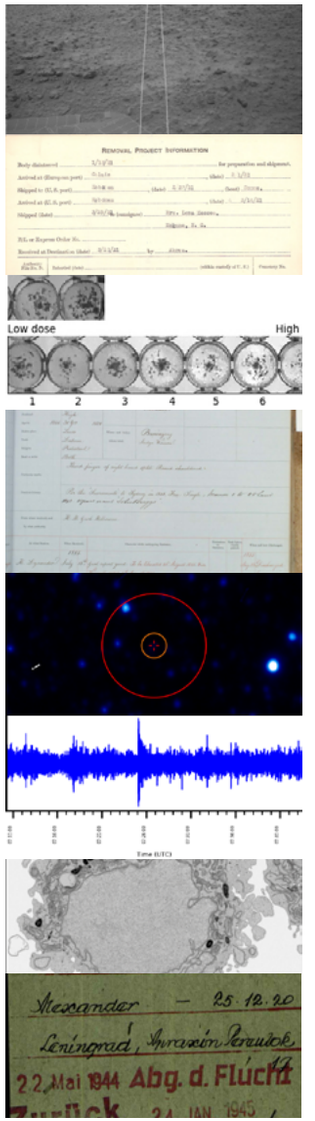
In the definition phase, teams had to choose the citizen science project and game to combine together. Participants could decide which to pick from a list of citizen science projects and a list of games.

Citizen science projects were selected from Zooniverse, «the world's largest and most popular platform for people-powered research»[34]. Zooniverse has supported some of the largest, most popular and most successful citizen science projects on the internet. It started from the original Galaxy Zoo project and it counts around 2 million registered volunteers. The projects on the platform demand the active participation of human volunteers to complete tasks, which was ideal for the objective of the workshop. Only active projects were chosen for the workshop. Projects were chosen to suggest various connections to the games. They were divided into four clusters for easing readability and facilitating understandability.

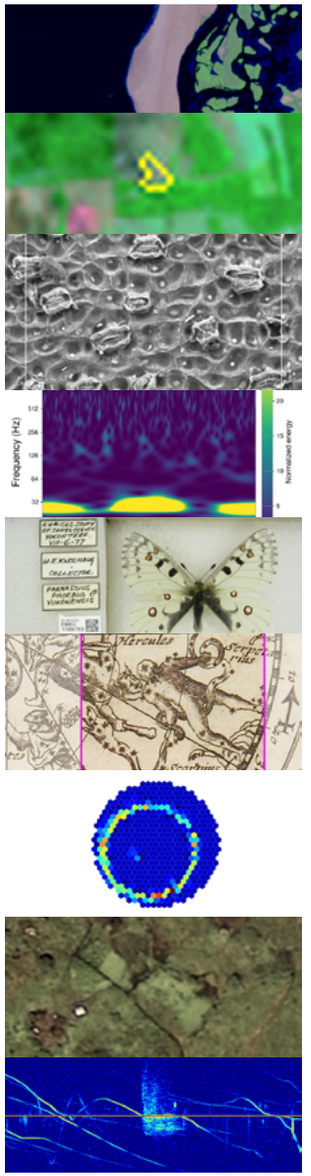
The following list contains the citizen science projects proposed during the workshop in alphabetical order, providing a short description for each:

34 www.zooniverse.org

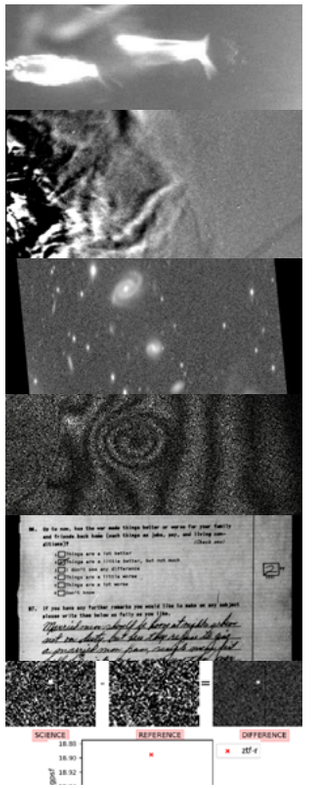
1. AI4Mars: help improve the rovers' ability to identify different, sometimes dangerous terrain, by mapping the terrain;
2. American WWI Burial Cards: transcribe the card register of burials of deceased American soldiers of World War I;
3. Bash the Bug: classify samples from patients with Tuberculosis treated with different antibiotics;
4. Criminal Characters: transcribe the records of persons imprisoned from the 1850s to the 1940s to discover who committed offences and why;
5. Disk Detective: search for dusty debris disks similar to our asteroid belt, and gas-rich primordial disks, the birthplaces of planets;
6. Earthquake Detective: classify types of seismic events such as earthquakes or tremor just by listening to them;
7. Etch a Cell: identify mitochondria in images of a cell at a very high level of magnification;
8. Every Name Counts: transfer names from documents on the victims and survivors of National Socialism to the internet;



9. Floating Forest: trace patches of kelp, submerged floating forest, to track this ecosystem state;
10. ForestEyes: track rainforest deforestation and build a database to improve automatic algorithms;
11. Fossil Atmospheres: analyze the cells of ginkgo trees, which evolved before the dinosaurs, to learn about the ancient atmosphere of the Earth;
12. Gravity Spy: classify sources of noise to improve gravitational waves detection;
13. Notes from Nature: transpose herbarium specimen images and metadata to the internet;
14. Mapping Historic Skies: analyze the Adler Planetarium's celestial cartography collection and map its constellations;
15. Muon Hunter Classic: identify muon rings and filter out those pesky muons that are masquerading as gamma rays;
16. Power to the People: find rural buildings in satellite imagery of sub-Saharan Africa to provide them with electricity;
17. Radio Meteor Zoo: identify the various and complex shapes of meteor echoes during meteor showers;



18. Run, Herring, Run!: count the herring in each frame by marking all the ones you can see;
19. Solar Stormwatch II: improve tracking techniques by tracing the outline of storms in images of the solar atmosphere;
20. Spiral Graph: draw galaxy shapes to produce an approximation to measure the winding of spiral arms in galaxies;
21. Steelpan Vibrations: map of the vibration of steelpans taken with a high-speed camera;
22. The American Soldier: transcribe pages of commentary that touch upon myriad facets of soldiers' wartime experiences;
23. Zwicky's Quirky Transients: classify celestial objects detected by ZTF telescope;



Taskification is a process that aims at directing a player from playing a game to partaking in a purposeful activity, like contributing to crowdsourcing project. Hence, it is more successful when applied on games with a massive user base. For this reason, only games that reached at least 10 million players in their lifetime were considered. Moving a game with this amount of players would mean reaching five times the number of people on Zooniverse.

The following are the games included in the workshop, reported in alphabetical order and with the reference which demonstrates their great user base:

1. *Animal Crossing* (https://www.theguardian.com/games/2020/may/13/animal-crossing-new-horizons-nintendo-game-coronavirus);
2. *Apex Legends* (https://www.pushsquare.com/news/2019/03/apex_legends_reaches_50_million_players_in_its_first_month);
3. *Ark: Survival Evolved* (https://www.thexboxhub.com/xbox-one-x-update-ark-survival-evolved-brings-visual-enhancements-galore/);
4. *Crossfire* (https://variety.com/2020/film/news/crossfire-movie-sony-1203502166/);
5. *Call of Duty WWII* (https://www.polygon.com/2020/4/10/21216734/call-of-duty-warzone-50-million-player-count);
6. *Candy Crush Saga* (https://www.theguardian.com/technology/2014/jun/23/candy-crush-saga-freemium-games);
7. *Dota 2* (https://www.riftherald.com/2016/9/13/12865314/monthly-lol-players-2016-active-worldwide);
8. *Dungeon Fighter Online* (https://www.gamesindustry.biz/articles/2018-06-21-dungeon-and-fighter-gross-lifetime-revenue-exceeds-usd10bn);
9. *Farmville* (https://venturebeat.com/2011/01/03/zyngas-cityville-becomes-the-biggest-ever-app-on-facebook/);
10. *FIFA 20* (https://www.ea.com/games/fifa/fifa-20/news/fifa-20-10-million-players-infographic?es_p=10359447);
11. *Fortnite* (https://www.theverge.com/2020/5/6/21249497/fortnite-350-million-registered-players-hours-played-april);
12. *Forza Horizon 3* (https://www.theverge.com/2020/5/6/21249497/fortnite-350-million-registered-players-hours-played-april);
13. *Grand Theft Auto V* (https://www.pcgamesn.com/ps4-player-population);
14. *Heartstone* (https://www.pcgamer.com/blizzard-celebrates-100-million-hearthstone-players-with-free-card-packs-for-everyone/);
15. *League of Legends* (https://www.gamesindustry.biz/articles/2018-08-10-culture-clash-why-arena-of-valor-is-struggling-in-america);
16. *Left 4 Dead 2* (https://arstechnica.com/gaming/2018/07/steam-data-leak-reveals-precise-player-count-for-thousands-of-games/);
17. *Maple Story* (Deterding, Sebastian; Zagal, José (April 17, 2018). Role-Playing Game Studies: Transmedia Foundations. Routledge. p. 318. ISBN 9781317268314.);
18. *Metal Slug Defence* (https://games.app-liv.jp/archives/66256);
19. *Microsoft Solitaire* (Marcus, Aaron (July 20, 2015). Design, User Experience, and Usability: Interactive Experience Design: 4th International Conference, DUXU 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings. Springer. p. 284. ISBN 9783319208893.);
20. *Minecraft* (https://www.pcgamesn.com/minecraft/minecraft-player-count);
21. *Overwatch* (https://www.pcgamer.com/overwatch-has-more-than-40-million-players/);
22. *Pac-Man* ("Men's wear, Volume 185". Men's wear. Fairchild Publications. 185. 1982.);
23. *Players Unknown's Battlegrounds* (https://in.ign.com/playerunknowns-battlegrounds-mobile/142341/news/pubg-mobile-gets-600-million-downloads);
24. *Rocket League* (https://www.rocketleague.com/news/rocket-league-s-2020-infographic--five-years-and-counting/);
25. *Star Wars: The Old Republic* (Deterding, Sebastian; Zagal, José (April 17, 2018). Role-Playing Game Studies: Transmedia Foundations. Routledge. p. 319. ISBN 9781317268314.);
26. *Street Fighter II* ("Business Week". Business Week. Bloomberg (3392–3405): 58. 1994.);
27. *Roblox* (https://www.nytimes.com/2020/08/16/technology/roblox-tweens-videogame-coronavirus.html);
28. *Robocraft* (https://arstechnica.com/gaming/2018/07/steam-data-leak-reveals-precise-player-count-for-thousands-of-games/);
29. *Teamfight Tactics* (https://venturebeat.com/2019/09/25/teamfight-tactics-hits-33-million-monthly-players-making-riot-games-happy/);
30. *Tetris* (McGonigal, Jane (2016). SuperBetter: The Power of Living Gamefully. Penguin Books. ISBN 0143109774.);
31. *The Last of Us* (https://www.ign.com/articles/2018/12/17/sony-may-have-revealed-how-many-people-have-played-ps4s-biggest-games);

32. *The Sims 4* (https://venturebeat.com/2020/01/30/the-sims-4-has-reached-20-million-players/);
33. *Unturned* (https://arstechnica.com/gaming/2018/07/steam-data-leak-reveals-precise-player-count-for-thousands-of-games/);
34. *World of Tanks* (https://worldoftanks.eu/en/game/);
35. *World of Warcraft* (https://www.polygon.com/2014/1/28/5354856/world-of-warcraft-100m-accounts-lifetime);
36. *Yu-Gi-Oh! Duel Links* (https://www.konami.com/games/corporate/en/news/release/20170126/).

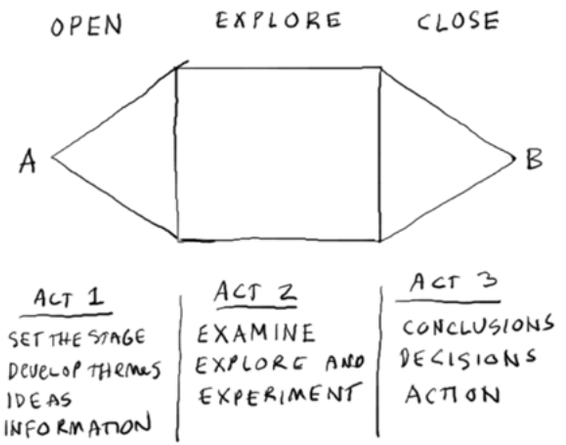


Fig. 53 A model for brainstorming (Gray et al., 2010).

The brainstorming phase was based upon the Gray and colleagues' model (2010) in Fig. 53. Three phases constitute the model: open, explore and close. The "open" phase is for divergent thinking. Participants began to propose ideas and let them flow. The goal is to reach the widest possible spread of perspectives. The more ideas, the more material there is to work in the next stage. The "explore" phase advances emergent thinking. Participants analyze and experiment with the ideas collected in the previous phase. They look for patterns, analogies and new perspectives by sorting ideas. The aim is to encourage unexpected, surprising and delightful things and let them emerge. The "close" phase is for convergent thinking. Participants assess ideas, look at them with a critical or realistic eye. They attempt to understand which are promising and worthy of their time and energy. The objective is to narrow down ideas to reach a conclusion. At the end of the brainstorming, participants had clear what they wanted to design.

In the integration phase, participants focused again on the framework. They deepened the idea resulting from the brainstorming. They broke it down following the framework structure to understand if it worked. They attempted to care for the various aspects of the combining process. Participants finally formalize their concept with images, sketches, descriptions, storyboards, all the tools they find more useful for their purpose.

After the workshop, the pilot ended with a moment dedicated for feedback. Participants had the possibility to leave their feedback regarding the pilot, the framework and the result that they achieved. I conducted a focus group to encourage participants to express their opinion, thoughts and doubts. The focus groups were semi-structured and relied on three general question:

- How was the workshop?
- What do you think of the framework?
- What do you think of your final design?

The following question stemmed either by the answer to these questions, to dig into the thoughts of the participants to the workshop, or by observed behaviours, to understand the reason behind them. I collected the data both through notes and recordings, in order to cross-check later participants' comments.

The activities relied on digital tools, hence participants had the possibility to access it without place restrictions. This choice was guided by the presence of the pandemic of Covid-19. Although no formal restriction prevented physical meetings, it was highly preferable to avoid gatherings. Moreover, many people were not reachable as they moved to other regions due to the pandemic. Participants were free to express preference between meet in person or virtually.

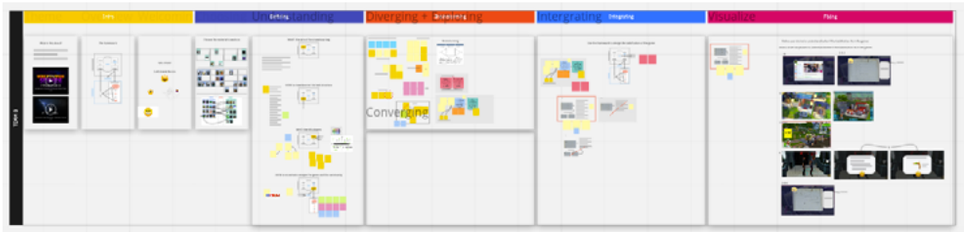


Fig. 54 The virtual board designed for the workshop, completed by team B.

A virtual board (Fig. 54, Fig. 55) was available to all the team members to share ideas and tools, e.g. the framework. The board was designed to display the structure of the workshop (the four phases aforementioned) and to support both the explanation of the activities and the activities themselves. The aim was to maximize participants' understanding of the requests and give them all the tools to accomplish them. Labels clarify the aim of each workshop section. The lists of games and citizen science projects were designed to be easy to quickly analyze them. Both were divided into clusters to improve search and identification. Each element in the lists contained hyperlinks that returned extensive descriptions of the element itself. In particular, citizen science projects linked to the actual project, so that participants to the workshop could attempt to accomplish the task requested by their own and understand better the functioning of the project. Dot-voting and virtual post-it were already set to be quickly accessible for novice users of the virtual board. For the ones attending the workshop online, a video call service allowed them to speak to each other to comment and explain the content on the virtual board.

The framework and the overall structure in four steps of the workshop were always the same. However, the explanation of activities and their relation to the framework was improved each time according to observed issues and previous participants' feedback. Each iteration enabled a reframing of the activity to better explain them to participants. Issues emerged in Pilot 1 were addressed with a redesign of the virtual board and rephrasing of explanations, while after Pilot 2 only the explanations were enhanced and some suggestions added. In this way, participants

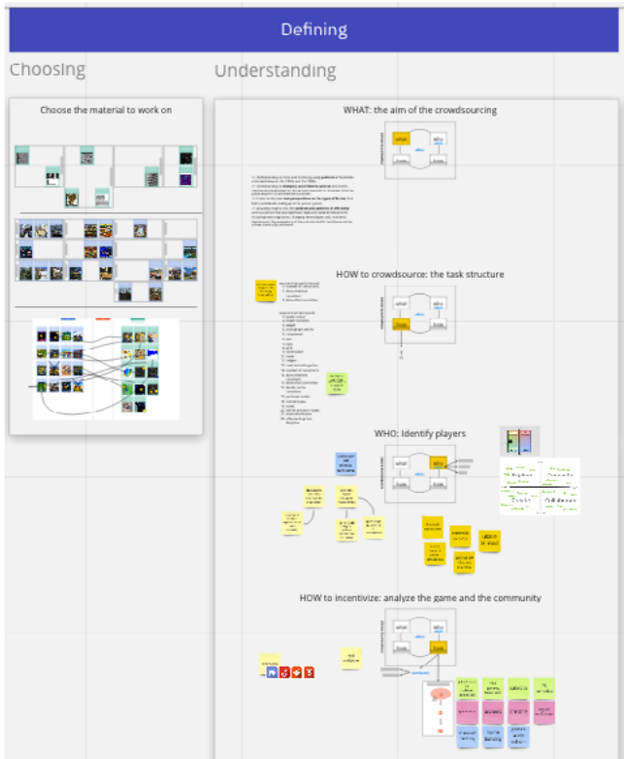


Fig. 55 A close-up of the defining phase from team B.

improved in using the framework to taskify the game with the citizen science project. The issues, the consequent reflections and redesigns are detailed in the description of the pilots (par. 4.2.1, 4.2.2 and 4.2.3).

The paragraphs in the following describes the activities of the three pilots, posing attention to the dynamics that took place in the workshops, during which participants were involved in designing a game taskification for a citizen science project. What follows is therefore the result of the observation conducted, as well as of the focus groups held with the participants right after the workshop.

The data produced through the participant observation report participants' behaviours, teams dynamics, personal and collective reflections. I paid much attention on the reasoning occurring regarding each element of the framework to understand when and how the tool was able to boost the design process proposing relevant topics and issues. Similarly, the data stemmed from the focus group focus on these aspects.

At the end of each of the next paragraphs there is a discussion about the main issues emerged from the observation and focus group. Alongside, it is explained how specific redesigns attempted to face these issues and provide a better experience for the following participants in the workshop.

4.2.1 Pilot 1

Pilot 1 was on the 9th of July with team A, composed of a male computer science student, a female game designer graduate and a male interaction designer graduate. The workshop was held online, started around 10 AM and ended around 17 PM. First there was an introduction to the workshop of 20 minutes, where the theme of taskification, the aim of the workshop and the framework were explained to participants. At the end of the introduction, there was an icebreaking activity (5 minutes) so that participants could introduce themselves and get comfortable. Then team A was guided through the various activities of the workshop following roughly this timetable: definition (1 hour), brainstorming (1.30 hours), lunch break (1.30 hours), integration (1.30 hours) and synthesis (1 hour). At the end, I held a focus group with participants of 30 minutes.

Participants in this pilot reasoned deeply on the idea of taskification itself. They attempted to find theme-based connections between citizen science projects and games, e.g. they reflected to combine *Ark: Survival Evolved* (Studio Wildcard, 2017) with *Fossil Atmospheres*, which are both are set in prehistoric ages. One participant wondered what was the better choice between explaining in depth the citizen science project or disguising it in the game as much as possible, which was a remarkable consideration to define the approach to the design.

Team A agreed to work on *The Last of Us* (Naughty Dog, 2013) and *Power to the People*[35]. *The Last of Us* is an action-adventure game set in a post-apocalyptic world. The *Cordyceps* fungus has muted so it can infect humans and turns them into aggressive creatures. The protagonist Joel has to guide a teenage girl, Ellie, to a rebel militia, the Fireflies. Ellie is immune to the fungus and the Fireflies hope to produce a vaccine based on her immunity. *Power to the People* aims at finding rural homes to provide them with electricity. The lack of census data in poor rural areas, where most poor people live, makes grid design difficult and time-consuming. Participants in the project identify buildings in satellite imagery of sub-Saharan Africa to train algorithms to do the same. Locating people's houses will help designing smart electrical systems to meet their needs.

They analyzed the two systems – the game and the crowdsourcing project – and discussed the relevance of the narrative in *The Last of Us*. They worked on fitting *Power to the People* to the storyline of the game to make it seamless to the original one. They discussed that the game narrative was more important than the reality of the project, so they focused on justifying the images of the citizen science project to fit into the fictional context. They did not spend much time on the mechanics: notably, they considered that the task already shaped the mechanic by itself and no other contribution was needed from them apart to integrate it into the story. They felt from the beginning that the task could not be changed so they did not even attempt to elaborate it further. They connected the task to the game reward system by providing additional resources to face the game challenges to players completing the task. In general, their focus was strongly on the aesthetics components of the game.

35 https://www.zooniverse.org/projects/alycialeonard/power-to-the-people

They discussed even technical integration. They envisioned both the development of a new DLC for the game or addition of a minigame through a patch: they then decided to focus on the second option.

During the analysis of the game, members of team B were asked to reason on the community of *The Last of Us*, since it is an element of the framework. The team imagined to harness the large community and planned a first check of *The Lats of Us* players' interest via social networks, to understand if the taskification would have been appreciated by this target. Instead, since it is a single-player game, Team A did not experiment with collaborative play. They deemed it problematic to add cooperative missions or mechanics, because it could have twisted the atmosphere of the game and its dynamics. Hence, they decided to avoid a risky design choice and maintained a single-player structure. However, they thought about displaying individual performance in completing the task compared to the community ones. Players would see the community effort, without interacting with it but knowing others are contributing.

From an ethical point of view, they did not recognize possible issues. They designed the minigame in a way that meant to not affect the main storyline of the game, so people would choose freely to participate. Moreover, they decided that information about the cause of the project would be explained to players ahead to ensure aware participation. The team expressed doubt to the extent the player could lose perception of the real-world impact of their action. They wondered how to emotionally engage the player in the citizen science project and send a message through the game.

Although the workshop activities relied on the framework, participants did not understood clearly what part of it they were using in each stage. The framework was presented as a whole and participants followed more the mentor (me) than the tool. They went through all the elements while designing but they were not able to notice it because they spent more attention on my suggestion than on the framework. This issue triggered a specific re-design, being implemented for the field testing in the second iteration of the workshop design: Pilot 2. Each stage started highlighting the part of the framework which was at the core of the given activity, as shown in Fig. 56. When at hand, subcategories of the element inferred from guidelines were displayed to provide further guide in the design process. In this way, visuals helped participants in understanding what was the focus of the design activity at a given point.

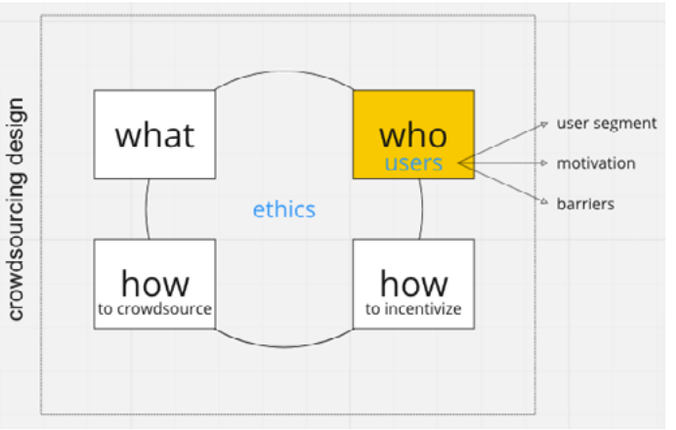


Fig. 56 An highlighted element redisegned for Pilot 2 in the virtual board. Relative subcategories of the element inferred from guidelines were displayed at this point.

Going back to the description of the activities occurred in Pilot 1, the brainstorming phase led Team A to finally define a concept. They introduced the task as a secondary quest. In the first part of the game, the player meets Tess, who instructs the protagonist Joel about its main quest. Team A decided that Tess explains even this secondary task to the player. Tess would give Joel a device with which he can enter some servers during his journey. Those servers have access to satellites which are still working. Team A exploited the post-apocalyptic setting and pretended the satellite images represented isolated survivors' dwellings. Tess would ask Joel to look for those survivors and report their presence on the device.

Team A designed the task to provide feedback to the player, in order to visualize their contribution in advancing knowledge. After completing the task, a graph would show player accuracy compared to other players. The player can understand its perfor-

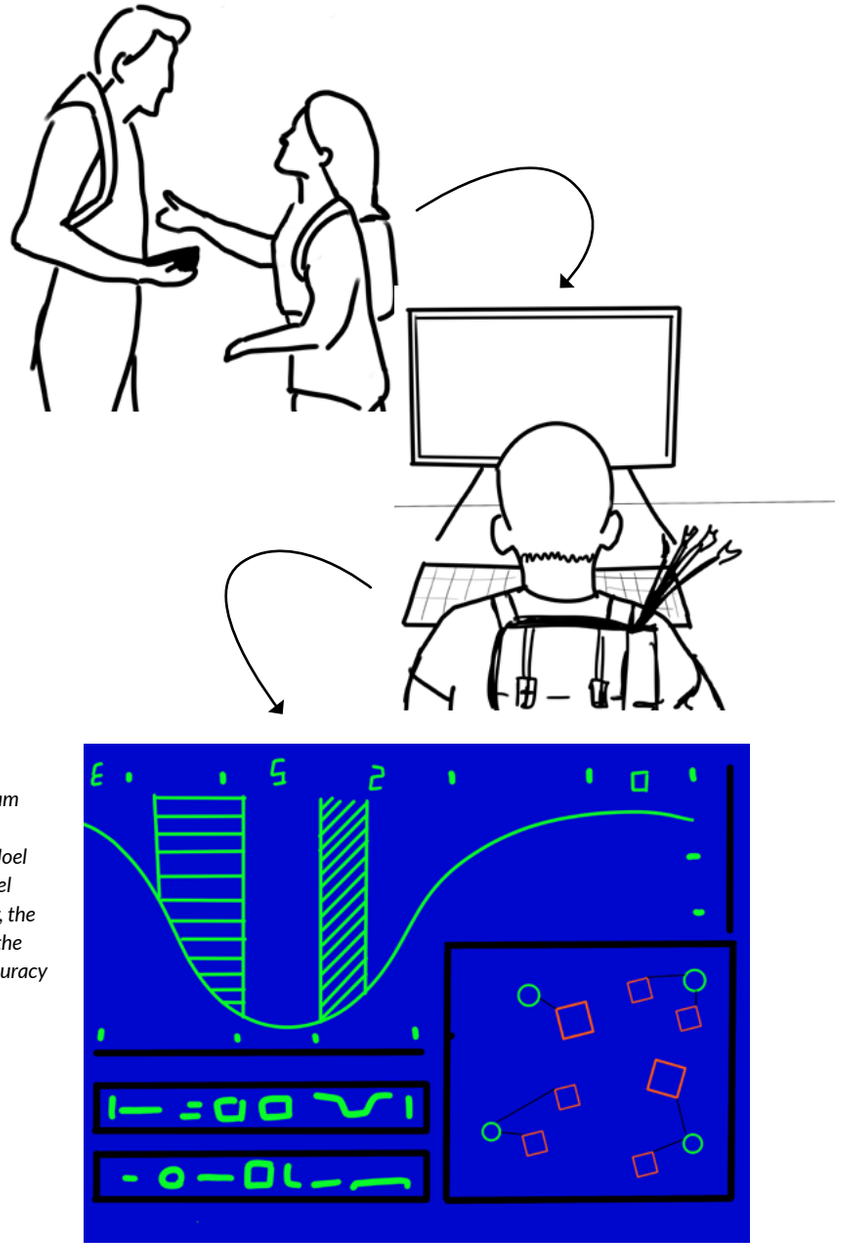


Fig. 57 Sketches from Team A workshop. From top to bottom: Tess explains to Joel how to use the device, Joel finds a functioning server, the in-game visualization of the task with the player's accuracy graph.

mance and feel part of a community by seeing collective results. Data would be anonymous, but their accumulation would display players' effort toward a common goal. Team A imagined that Joel is not the only looking for survivors and Tess would explain this.

The player would earn additional resources by completing the task. When a new level starts, Joel and Ellie would discuss their encounter with the people living in the villages found through satellite images. During the discussion, they unveil some elements of the lore and explain the resources gained.

Team A decided that the quest would end once the player reaches the hydroelectric dam existing in the original story. At the dam, the player would deliver the device to Tommy, Joel's brother, who would promise to provide electricity to the buildings found by the player. In this way, Team C kept the theme of helping in bringing electricity to people in need.

At the end of the workshop activity, participants were involved in a moment of collective discussion, aimed at encouraging them to express their opinion on the experience and the framework. One member of team A affirmed the framework was very useful. She felt that it helped them reflect on many factors and the framework supported well the design process. She particularly appreciated the integration phase, stressing that thinking about the community aspects of the project was interesting. She claimed that thinking about the community helped her consider not only the game but also users' reactions. She considered useful basing the design on the MDA, pointing out that it made it possible to design a coherent taskification following a logic of completeness and general consistency. Overall, she claimed the framework reached its aim and that without it they would have been still figuring out how to act after eight workshop hours.

Another participant said that there was little choice among citizen science projects, noticing that they were mainly image and sound recognition tasks. He argued that it is not possible to shape these tasks freely because the mechanic was already the task itself and it was not possible to develop it further. Hence, aesthetics and dynamics become more relevant in the design process. He felt that they did just make an excuse to insert the task in the game and

its storyline. He suggested that citizen science projects in the field of physics may foster more mechanics design. He realised that the game was unchangeable in its opinion. He dismissed options which could deeply modify the game structure. Probably, this participant was misled by the relevance of the narrative in *The Last of Us* and fearful of damaging the game dynamics, he focused on the story. The other members agreed on the idea that aesthetics and dynamics were more important and that mechanics was not a core element.

The last member generally agreed with the others. He thought the output that they developed was interesting. It was clear that the aim of the framework was guiding a process to fit the citizen science project in an existing game. He wondered if the framework could support taskification design even if the people doing it have no education in game design. Indeed, the framework is rooted in the MDA, which requires some expertise.

The team agreed that story-driven games appear to be the best or even the only option that can fit with citizen science projects. They thought that the narrative was highly important to make sense of the citizen science into the game; because the story provides excuse to integrate the task. They claimed that they found it not possible to taskify a sandbox or a strategy game. They thought that citizen science projects can combine well with games that contain similar themes, as *The Last of Us* and *Power to the People*.

After analyzing the data from the observation and focus group of Pilot 1, as aforementioned the major issue that emerged was the lack of independence in using the framework. Participants were not able to understand which part of the framework they were focusing on and highly rely on my guidance. While it was predictable that mentoring would have been important since it was not scheduled for participants to study the tool aforehead, it made it hard for participants to follow the framework and understand how it worked. This was highly problematic because users had difficulties in determining if the framework was useful to the design process as they did not understand how they were using it. Hence, I redesigned the visuals of the framework on the collaborative tool, highlighting the part that was used in each phase. In this way, future participants could better understand the relation between the framework and each designing phase.

Another issue was the use of the framework even in the brainstorming phase. Indeed, this misled team C, who did not explore freely various possibilities, but stuck to the elements in the framework. The framework is highly useful in guiding the analysis of the game and the citizen science project in the “defining” phase, but turned to be damaging for the creativity requested in the brainstorming. Therefore, in the following pilots participants were asked to brainstorm freely and only later confront their ideas with the framework. Actually, as the brainstorming phase follows the defining one, particiapnts are already influenced by the knowledge gathered through the analysis of the systems that they are trying to combine. Hence, the framework would be still guiding the design process but in a implicit way, which enables more freedom in brainstorming.

4.2.2 Pilot 2

Pilot 2 was on the 11th of July with team A, composed of a female automation engineer employee, a male game designer student and a female communication designer graduate. The workshop was held online, started around 10 AM and ended around 18 PM. First there was an introduction to the workshop of 20 minutes, where the theme of taskification, the aim of the workshop and the framework were explained to participants. At the end of the introduction, there was an icebreaking activity (5 minutes) so that participants could introduce themselves and get comfortable. Then team B was guided through the various activities of the workshop following roughly this timetable: definition (1 hour), brainstorming (1.30 hours), lunch break (1.30 hours), integration (1.30 hours) and synthesis (1 hour). At the end, I held a focus group with participants of 40 minutes.

Team B discussed theme-based connection as well, also exploring other combinations. They proposed mixtures that would

completely twist the original game. An example was their idea of combining *FIFA 20* (Electronic Arts, 2019) and *Earthquake Detector*, adding earthquakes to the football game and asking players to avoid earthquakes by doing the task. Other proposals relied on dynamics-based connection, e.g. *Candy Crush* (King, 2012) and *Run, Herring, Run!*. In this case, the task would appear as a fast-paced event during the gameplay. Team B imagined these two could be combined by having a quick time event in which herrings would "run" through the screen at some point during the gameplay. They commented that even combining citizen science projects and games with no evident connection was not intuitive; however they considered it interesting to discuss. In particular, they reflected on games which have toxic community issues and the impact that implementing a citizen science project could have on such games.

After thoughtful discussion, Team B agreed to work on *The Sims 4* (Electronic Arts, 2014) and *Criminal Characters*[36]. *The Sims 4* is a life simulation game. The player controls sims' lives from a god-like perspective and customizes every element of the game. *Criminal Character* aims at understanding crimes and their history by capturing the first large-scale data on the life histories and offending patterns of Australian criminals. The database covers a period extending from the end of the convict era (1850s) to the beginning of the Second World War (1940s). Participants in the project transcribe the documents to make them digitally available.

The "What" in the framework regarding crowdsourcing design was not clear and a participant proposed starting from the "How" element. After explaining again the aim of the "What", they completed it and proceeded. The observation showed that Team B better understood the meaning of each element in the framework thanks to the highlighting employed after observation of team A in Pilot 1. At the same time, the terms are not self-explanatory. It is necessary to provide a proper introduction for participants to use the framework.

36 https://www.zooniverse.org/projects/ajpiper/criminal-characters

Team B was careful to not impact on the overall game and its meaning, by altering its mechanics and dynamics as well as the genre of the game itself. A member recognized that the analysis of "barriers", a subcategory of "Who" which aims at understanding what the user could perceive as a motivation to not participate in the activity, already made clear the need to preserve the game. Looking for "barriers", they had identified the need for creating a seamless experience to the original game. The task was inserted in an investigative dynamic, being careful to not turn the game into a role-playing game (RPG). Indeed, the first ideas of participants were to make the sim find clues around the neighboorhood and attempt to identify the killer. However, this activity could change too much the dynamics of the game. Hence,they introduced the task as a new job opportunity for the sim. However, the interaction between the player and the sim remained unchanged. Team B put a lot of attention in preserving the player in its god-like status.

An interesting discussion regarding ethical issues took place during Pilot 2. While this discussion was almost completely absent in Pilot 1, team B explored the matter of ethics of implementing a citizen science project with historic references into a light-hearted game like *The Sims 4*. They debate that *The Sims 4* does not attempt to be realistic, and elements coming from the real world, as documentation of historic criminals, could be problematic. They decided to insert disclaimers to remember the historical setting of the task and did not ask players to transcribe the accusation to avoid it to conflict with the tone of voice of *The Sims 4*.

The final concept by Team B relies on the job mechanic in *The Sims 4*. With a new expansion pack, the player would get a new job opportunity, the hell employee. By applying to this job, the sim would travel in time where they could find shred documents describing criminals of that period. The sim would have to collect all the parts of the document and take note of its content. Team C disguised the transcription task as the need to take notes. They deconstructed the task to better combine it with the dynamics of the game. Once the player has collected all the part of the document, they have a complete profile of the criminal. Hence, the game would produce a sim in the neighbourhood with the characteristi-

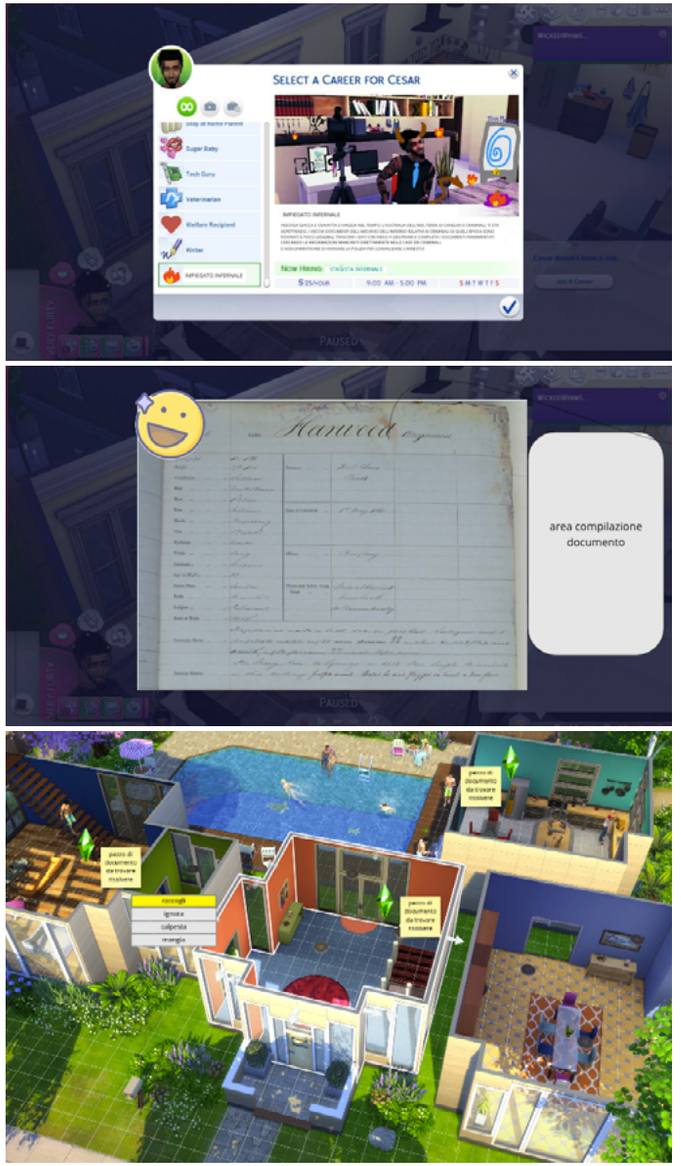


Fig. 58 Composite images from Team B workshop. From top to bottom: the "hell employee" job description, the document transcription interface and the investigation section.

Fig. 59 Composite images from Team B workshop. From top to bottom: the criminal identification section and the "mission complete" alert showing the reward.

cs of the criminal (height, hair colour, etc...). The player would have to recognize the criminal sim and then expose them to the police. If the criminal sim is captured, the players get a reward, namely a themed item (a hat from 1850, a historical boomerang...).

Team B was generally impressed by the result that they were able to achieve. Discussing with them during the focus group, they sound satisfied with the concept that they developed and the process that they went through. A member affirmed that the workshop phases and activities were clear. She explained that the framework helped them set what they were discussing during the workshop. In particular, she stressed that the analysis phase was well described and ordered. She claimed that the analysis made them think about factors that they would have been missed

otherwise. She felt that she knew what she had to do all the time demonstrating the effectiveness of the framework in leading the taskification activities. She explained that they did not get stuck and find good hints using the framework. She thought that this could be caused by the chosen game, since the team had good knowledge of it. For this reason, it was easy for them to harness *The Sims 4* aesthetics, dynamics and mechanics and taskify it.

Another member claimed that they had been guided properly in the process. He argued that they slow down the process in the last phase of the brainstorming, the converging phase. Nevertheless, he expected such activities to be time-consuming. He said that they reached a good detail level in just eight hours and with only three participants. He considered that the result demonstrated the validity of the framework. He stated that the integration was quite good. He affirmed that *The Sims 4* was an easy pick, likewise to other sandbox games. These two members agreed that the framework helped them deal with an unknown topic, namely citizen science projects.

The other member did not argue back nor concur. She focused on the crowdsourcing design part of the framework. She felt that it was little overseen at a certain point, hence the objective was less clear from that time. She suggested to either keep an eye on it during the later phases or check it again at the end of the process. However, she explained that she was uncomfortable working on such a famous game. She argued that the workshop appeared out-of-scale, meaning that the game was not manageable for students or fresh graduates. She seemed to imply that the workshop was tailored for the actual developers of the games proposed in the lists. She claimed that the framework and the workshop worked quite well, but at the same time she was unsure since she felt not to be on target.

Considering the feedback of participants and the observed behaviours, the major issue was the fact that there was no check on elements that were defined in the first stages of the workshop, in particular the ones set in the "defining" phase. Although there were no changes in the structure or the visual presentation of the tool, this issue was borne in mind in the mentoring. Indeed, when facing the "integrating" phase in the ending phases participants to Pilot 3 were stimulated to reconsider the elements set in the "defining" phase to check if the whole design was consistent and solid.

4.2.3 Pilot 3

Pilot 3 was on the 18th of July with team A, composed of a male computer science graduate, a male interaction designer employee and a male interaction designer graduate. The workshop was held in presence, started around 10 AM and ended around 18 PM. First there was an introduction to the workshop of 20 minutes, where the theme of taskification, the aim of the workshop and the framework were explained to participants. In this case, there was no icebreaking activity because participants already know each other and were already comfortable. Then team C was guided through the various activities of the workshop following this timetable: definition (1.45 hours), brainstorming (1.45 hours), lunch break (1 hour), integration (1 hour) and synthesis (1 hour). At the end, I held a focus group with participants of 50 minutes.

Team C focused more on turning the task into a pleasant quest in the game. For example, they proposed exploiting the driving mechanic in *Grand Theft Auto V* (Rockstar Games, 2013) to map the steelpan vibrations. The idea was to ask players to drift along the edge of the ellipses that one could recognize in the image, which would appear over the in-game map.

They discussed that open-world games gave more freedom to implement citizen science projects. At the same time, they recognized that it is more complicated to deal with open-world games since their structure is much more elaborated. Open-world games enable players to move freely in the virtual environment and choose to engage with missions, quests or other elements. It becomes hard to encourage players to choose one activity over the other because the game structure is non-linear and people decide freely. They got interested in the idea of exploiting birds-eye images in open-world games. Such games usually include mini-maps or maps

to support player orientation. Team C thought these maps could combine well with aerial views like the ones from *Power to the People*. However, they were unsure how to merge the new mechanic required by the task. Indeed, maps are usually used in games to display the environment and players expect to reach whatever is on the map. Implementing in the game environment a replica of the content from images of the citizen science project is not feasible. The content is unknown and it is exactly what the task attempts to identify. Hence, team C found it impossible to pursue this idea.

They reasoned on the possibilities to combine the 2D elements of the citizen science projects with the available games, that were both 2D and 3D. They agreed that whatever the image, it was possible to disguise it as something else and not disrupt the task. They discussed whether to implement the task as a minigame or part of the game. Team C recognized that inserting a citizen science project and making it outstanding could be risky. The publisher should be into contaminating the entirety of its game for the sake of the citizen science project, which is unlikely.

Team C agreed to work on *Grand Theft Auto V* and *Moun Hunter Classic*[37]. *Grand Theft Auto V* is an open-world adventure game. The player takes the role of a gangster in a fictional city in the USA. It has both a story mode and a multiplayer one. *Muon Hunter Classic* is a citizen science project that requires participants to help astronomers to find muons disguised as gamma rays, recognising which images in a database actually represent mouns. The task is relatively easy as mouns leave a remarkable circular mark in the images. Hence, images showing a circle are deemed mouns.

They chose *Moun Hunter Classic* because the images of mouns are abstract and easy to apply to different contexts. Team C discussed even using them as normal maps for environmental elements, e.g. mountains. They imagined a quest where users had to walk around the edge of certain mountains. If the system registered that the player walked around a circle, then the normal map would be verified to represent a moun. The team spent a lot of time looking for solutions to reduce errors in crowdsourcing. They wonder how to deal with the fact that the task can not return feedback about the correctness of the classification. They discussed possible issues in doing the task (e.g. finding noisy images hard to classify) or to avoid cheating behaviours. They decided to implement control groups to check players' reliability and distinguish accurate players from malicious ones. They spent time understanding how to justify the control group in the story. They explored the possibility of using collaborative play to support participation in the task.

Then, they focused on how to engage players in the task. They reflected on the reward systems to ensure players' interest in participating. They analyzed various options that make sense with the setting of the game to maintain consistency with the finctional world. However, not all options worked with the dynamics of the game. They discussed them and picked the one which provided a more seamless experience. Following the suggestion from one of the members of team B, it was proposed to team C to look back to the analysis during the definition phase, but they did not consider again that analysis despite the various suggestions to do so.

Their final concept exploited a theme already existing in *Grand Theft Auto V*. Team C thought the task as a secondary quest. The player would unlock this quest by interacting with a group of NPCs (non-playing characters) believing in government conspiracies. This group already exists in the original setting of the game and provides other quests. In this case, they would ask the player to help them get rid of some infected cows in their herd. They claim that the aliens have kidnapped those cows and get some of them pregnant. The conspiracy theorists argue that pregnant cows are marked with a circle, while other cows are sane, but the mark is visible only during the night. The player's mission is to find and kill the infected cows. The circles are actually moun images coming from the *Moun Hunter Classic* project. By killing the cows, the player confirms that the image related to that cow represents a moun.

37 https://www.zooniverse.org/projects/zooniverse/muon-hunter-classic

To avoid that players kill the whole herd to get the mission done quickly, team C inserted a control group with known images. A group of cows would be marked with images known for not containing mouns. If the player kills these cows, the conspiracy theorists get mad at them for ruining their herd. Otherwise, the player would unlock a unique vehicle. By keeping playing the quest, the player receives new upgrades for that vehicle (machine guns, wings, rockets, and so on) that can be later used to personalize the vehicle.

Fig. 60 Composite images from team C workshop. Top: a conspiracy theorist with an alien. Right: a cow with the moun mark on the back. Bottom: an example of the reward gained by the player through the mission.

Team C appeared thrilled by the experiment. A member claimed that he expected the taskification to produce forced designs. He doubted that games and citizen science project could combine. He thought that integrating a citizen science project could disrupt games' rules, and he was surprised to find out that it was not the case. He was satisfied by the end result of the workshop.

He agreed with another member that claimed that he expected to end up developing a game in the game like in the case of *Borderlands Science*. This member explained that *Borderlands Science* did not catch him because it only rewards money. Another participant stressed that *Borderlands Science* is particularly efficient: for a short playing time, players perform many classifications. Their concept instead would produce fewer classifications in the same playing time. The other participant stressed that the minigame was useful only to farm currency, hence provided an unbalanced experience. He claimed that *Borderlands Science* was more focused on the citizen science project and less on the game. For this reason, he felt that a player would participate for a short period to gain in-game currency and then would spend time on other activities, forgetting about *Borderlands Science*. He found that it was not a good integration.

The last member agreed and discussed that their concept weighed stakeholders' interests in a more balanced way. Their concept provides a more appealing experience for the player and the publisher would gain a DLC to sell and boost engagement in their game. Team C noted that they discussed even business and technical issues during the workshop. He explained that he enjoyed the challenge of combining a citizen science project with a game. He stated that the requirements from both parties sustained each other instead of clashing. He claimed that the intersection of these constraints produces a small spot to reflect on. He thought that it was not related to the chosen citizen science project and game. He explained that the game provides opportunities as tools and materials to work on and the citizen science project provides opportunities as hints to develop a new mechanic. Hence, they combine well together, they balance perfectly and do not cover each other.

He explained that open-world games appear to him as an easy solution to design a taskification as the one in *Borderlands Science*, i.e. including an arcade machine somewhere in the game. However, as the other participants, he was pleased to discover that it was possible to exploit the in-game mechanics. The other members agree and they were also satisfied to have integrated the task to the storyline. In particular, they were glad they could take advantage of an existing storyline, so that they did not have to invent much more, risking to disrupt the game. They felt the taskification felt exactly as a *Grand Theft Auto V* mission.

A member said that although he found it hard to follow the set ofactivities of the workshop because he is inclined to design in an unstructured way, the framework helped set staples. The

team diverged a lot with the brainstorming and he claimed that the framework helped to set some decisions. Another participant claimed that the framework showed well the relation between the two systems. He claimed that the citizen science project and the game were both represented clearly in the schemata. Ruminating about the framework structure, he suggested dividing the framework by a vertical line: the crowdsourcing on the left side while the game on the right side. This suggestion is further explored in par. 5.1.1. He affirmed that the analysis during the definition phase was useful to generate ideas. He also said that the upper part (crowdsourcing design) containing the "Who" element reminded them that they were working on *Grand Theft Auto V*, which has its specific contexts and users. This prevented them from speculating on all the possible options and pushed them choosing the best solutions. However, the participant noticed that it was thanks to my mentoring that they employed the framework that way. He admitted that they did not focus on the framework much without my suggestions. He expressed some concerns about how they would have designed the taskification without my mentoring. However, he did not find it problematic in a workshop since there is always someone guiding.

Another member commented that without the framework, they would have probably got stuck at the brainstorming phase. A participant observed that the steps set by the framework were mandatory to taskify the game, and recognized that they foster idea creation. At the same time, he admitted that they did not resume the work done in the late stages to see if it was coherent when compared with the analysis done during the definition phase.

A participant said that they focused more on the game aspects rather than the crowdsourcing ones. Another member argued that it happened because the citizen science project was easy. He claimed that the upper part of the framework (crowdsourcing design) was useful to understand the project. A participant agreed and observed that it helped them deconstruct and understand the task. They commented that it would have helped even more with complex projects like *Power to the People*. A member said that knowing the citizen science project was composed of microtasks enabled him to understand how to integrate it. However, he also claimed that the lower part of the framework (taskification design) is enough to design for a citizen science project with an easy structure like *Moun Hunter Classic*.

A member reflected that the "What" element was not particularly stimulating but it helped pair games and citizen science projects from an aesthetics point of view. He argued that if the task is reshaped to fit in the game, the "What" is not relevant. He considered the "Who" more useful for the game than for the citizen science project in a taskificiation design. He claimed that the "How" was the core element of the framework. He complained that the framework provided a schema to analyse the game, i.d. the MDA, while it did not for the task analysis. In his opinion, the framework was unbalanced.

Team C thought that their concept works because they combined a simple task with a huge game. They argue that doing the opposite would be really hard. A participant said that justifying the taskification in the storyline is not possible with all games. Another one was impressed that the taskification process did not seem more suitable for a certain category of games.

It was extremely interesting to record that the last group was the only to relate the success of the taskification not with the genre of the game that they chose to taskify, but to a larger concept: the extension and complexity of the game. Indeed, team A claimed narrative games give plenty of opportunities for taskification, and team B argued that simulation games are the perfect match instead. This and other considerations of the results of Pilot 1, Pilot 2 and Pilot 3 are further discussed and analyzed in the next chapter.

5 Conclusion

Taskification is a method to combine games and purposeful activities by inserting a task in a part of an entertainment game, like a quest or a minigame. Taskification is particularly useful for game-based crowdsourcing systems. Crowdsourcing requires high numbers of participants and it is possible to find so many users in games. Players can be direct to the crowdsourcing activity by taskifying a game, namely integrating seamlessly the task into the game experience.

This thesis focuses on taskification as an emergent method in game-based crowdsourcing. In particular, it investigates the design of taskified games for crowdsourcing purposes. The leading RQ of the research was: “How to **guide** the taskification design process?”. To answer this question, I designed a framework to guide the design of such integration and tested it with game designers through three pilots. Each pilot was structured in three main moments: preparation, workshop, discussion. Data was collected throughout the entire process conducting participant observation and running focus groups aimed at understanding if these users consider the framework useful to taskify a game, as well as their rumination on the experience.

5.1 Result discussion

Looking at the three pilots, specifically at the activities of the three workshops and the comments raised during the focus groups, in sum the testing gave good feedback to initial hypothesis that a tool could guide the design of taskification. Testers claimed the framework was useful, except for two members of Team A and one of Team B who expressed uncertainty. Therefore, six out of nine participants declared their appreciation for the framework. The majority of them (4) linked the framework value to the low resources employed to taskify the game. They explained that it enabled them to achieve a good result in a short time (eight hours) and with few people (each team was composed of three persons). Every team had at least a member who mentioned this point. Team C even stressed that the result of the taskification process exceeded their expectation, stating that the framework supported their creative process and led them to surprising outcomes. This demonstrates that the framework can effectively provide knowledge to game designers to face taskification challenges. In particular, the framework can boost the game design process for taskification and stimulate diverse solutions.

In parallel to these positive considerations, three participants claimed to be uncertain about the validity of the framework, advancing diverse motivations. Those motivations are described in the following:

1. Framework target

A participant (team A, Pilot 1) wondered if people not proficient in game design would be able to use this framework. However, it is necessary to point out his being an expert in interactive storytelling but not in game design. Therefore, his concern was probably referred to how his background affected his design activity, making his task harder than for the rest of the team, composed of game designers. As a matter of fact, what emerged is that he had troubles in the workshop because of his lack of knowledge in game design. In particular, dynamics and mechanics are difficult topics for inexperienced designers, as I could clearly notice by observing the design activity of this participant, as well as his interactions during the workshop that made evident a certain lack of familiarity with the game design terminology and some of its fundamentals. It becomes hard for such people to participate in the process and propose modes to integrate the task into the game. In this case, team A had two other experienced members who supported the design process well. Hence, the team was able to end the workshop with a good result. This suggests that the tool is proper for mixed groups, but there must be at least one game designer.

The MDA framework is a core element of the tool and it is an expertise of game designers. Hence, it is necessary that mixed teams include at least a game designer to properly use this taskification framework. A game designer with good proficiency in mechanics, aesthetics and dynamics would take a leading role in a mixed group that design a taskification through the framework that has been used in the pilots. Although it does not imply that other professionals cannot use it, it is important to know that the skills of game designers are required to build a proper team to design a game taskification.

2. Experimental setting

Another participant (team B, Pilot 2) was unsure about the framework to be useful because she did not test it in a familiar context, i.e. on a game that she designed. She wondered if the taskification framework could fit in a design process different from the one presented in the workshop. Her worry was that the actual

design context of the huge companies which develop the kind of games proposed in the workshop might present some specific requirements that did not emerge from the restricted experimental environment of the workshop. Hence, she did not deny that the framework is useful. She rather cast some doubts on its employment in the work pipeline of the big companies that produce the games proposed for the workshop. This could be obviously a next step for the research to improve the tool and adapt it to users' needs.

3. Unbalanced process

The last participant who advanced concerns (team A, Pilot 1) felt that the process was all about finding excuses to justify the task into the game. He and his team focused on aesthetics, in particular on the story. They had chosen *The Last of Us*, a game with a great narrative component, and *Power to the People*, a prosocial citizen science project that has a complex task structure. The framework did not provide much support to their design process because little reasoning was done on the dynamics and mechanics, which was driven by the highly narrative game that the team chose. The other groups instead valued all MDA elements equally as their games as well relied on all those elements equally. The upper part (crowdsourcing design) was useful to set the ground, but they little employed the lower one (taskification design). They did not modify the task but they inserted it as it was in the storyline of *The Last of Us*. The opinion of this participant was correct regarding their use of the tool in the workshop: they focused on the task integration in the story, overseeing the rest (mechanics and dynamics). This participant claimed that his team could not develop much more these elements. He explained that the task already assessed mechanics and dynamics. Hence, team A's only perceived need was to integrate the task with the aesthetics, considering that the game chosen is characterized for being highly narrative, and therefore for having an important storyline to consider and be coherent and consistent with. The framework appeared to lack the ability to guide the game design process.

Among these comments, however, only one denies the framework's validity, namely the last about "unbalanced process".. The other two set interesting questions about the target and usability of the framework, which can be further investigated by future research. The one that denies the validity of the framework may have encountered peculiar circumstances that led to such feedback. Analysing once again the design activity looking at the overall study, it is necessary to pose the attention on the fact that the game chosen by team A is extremely narrative. *The Last of Us* relies on a strong story that engages the player deeply. On the other hand, the citizen science project chosen by team A, hence the task that they attempted to integrate, is highly complex. *Power to the People* requires participants to complete a series of steps to fulfil a task and submit it. This combination may have led team A to focus on the story, really important for the game at hand, and leaving the task as it was, because too hard to deconstruct. Team A may have struggled using the framework no because it was useless but because the circumstances led them to give attention to certain elements more than others. The framework structure may be still valid, but in some cases like this, the relevance of the elements is not equal. Further research could unveil if it is actually true that not all elements should be weighted equally to reach a good taskification.

It is remarkable that the most of the issues emerged in Pilot 1 ("framework target" and "unbalance process"), only one in Pilot 2 ("experimental setting") and none in Pilot 3. It seems the iterative design of the pilots solved the issues or at least improved participants' understanding of the workshop activities so that the issues were not perceived as problems anymore. Next paragraph discusses other meaningful insights gathered by the data analysis, going deeper in the strengths and weaknesses of the framework, reporting as well the opportunities and threats noted through the observation and focus groups.

5.1.1 An in-depth analysis of feedback

The observation and focus groups revealed many strong points of the framework as a tool to guide the design of taskifcation for crowdsourcing. This paragraph digs more into the data collected in the three pilot to critically discuss remarkable insights. What emerges in general is the ability of the framework to systematize and operationalize the relationship between the two systems of game and crowdsourcing.

A member of Pilot 3 (team C) claimed that the framework clearly displays the relation between the two systems and their constitutive elements. Even a member of Pilot 1 (team A) argued that the tool supported the analysis well. Various elements were recognized by participants as particularly useful to analyze the game and the citizen science project. In particular, a member of Pilot 1 praised the "Community" element while two members, respectively from Pilot 2 and 3 recognized the "Who" element to be useful. Then, a participant from Pilot 2 noted that the subelement "Barriers" was helpful.

The observation and the following discussion with those who took part in the experimentation revealed the effective contribution of the framework. Especially by highlighting core elements and their relations, the framework succeeded in providing a quick understanding of how to taskify. At the same time, it shows stakeholders' interests and sphere of influence. A member of Pilot 3 claimed that the framework helped them design a taskifcation which balanced stakeholders' interests. He explained that the citizen science project and the game were both preserved and did not overcome each other, so the process was successful.

Moreover, the framework appears to be able to support the decision making processes. A member of Pilot 3 specifically mentioned that the framework helped the whole team to set decisions.

At the same time, however, the framework resulted to be still in need of improvement. Some weak points emerged indeed from the observation and focus group. In particular, the upper part of the framework (crowdsourcing design) presented some issues. A member of Pilot 2 felt that, after the analysis phase, her team did not reason on that part. This was probably due to the workshop structure. In any case, the analysis phase of the workshop was not pointless. I observed that another member of the same team (team B, Pilot 2) during the integration phase noted that the team was basing their design choices on the knowledge fixed in the analysis phase. Although not formally taken into account, the upper part of the framework (crowdsourcing design) guided the game design process even in later stages. It still presented other minor issues. The "What" element was confused with the "How to crowdsource" element by team B, in Pilot 2. A member of team C, in Pilot 3, deemed the same element as not really stimulating. He argued it was useful only to support aesthetics connections between the game and the citizen science task. Hence, the upper part of the framework could need some refinement for game designers to use it easily. Extending the reasoning, a possible solution could be merging the "What" and "How to crowdsource" elements. This change should be checked with a crowdsourcing expert to well understand if it could cause any issue. The lower part (taskification design) also presented minor issues. In Pilot 3, a member stressed that the framework is useful yet "unbalanced". He meant that the MDA provides a good way to analyze the game while the task is unexplained, however he also suggested that it could be further developed for also analyzing the task, so that it would support designing its integration in the game.

Finally, the "Ethics'' element was little investigated by all teams, transversally to the pilots. The major reason could be that, from a certain point of view, the proposed taskifications were intrinsically ethical considering that they involved citizen science projects. Players would have a return from the improvement of scientific research in terms of new medical treatment or better data to confront climate change. Otherwise, they could play "for charity", to help the people affected by a certain disease or with a lack of resources. If well informed, this kind of play would be ethical as well. Indeed, all participants in the workshop focused on privacy and information issues. "Ethics" are still fundamental in this framework:

if the crowdsourced matter benefited a company which would profit from the players' work, it would be hardly ethical. It is crucial to search for manipulative, harmful or exploitative applications of taskification. Likewise SGs and gamified artefacts, taskified games can have great negative side effects that should be recognized beforehand, hence avoided.

Although a good overall efficiency, the data gathered reveal that there are several opportunities to further refine the tool. "Community" was an underdeveloped element in all concepts from the workshop. Participants did not develop games for collaborative play. Only team A (Pilot 1) decided to include a feedback graph to show the single player's score related to the other players. However, they planned no interactions. Participants did not experiment with known stimuli for gamers like epic feelings or *naches* (McGonigal, 2011). The context of "Community" can probably be expanded to provide further knowledge that can support the design of game-based crowdsourcing systems. Contrary, an element which was particularly discussed by all teams was the reward. Indeed, the actual driver for players to participate in an in-game task is the reward. The player has to perceive it as valuable to engage in the task. The reward system is part of the MDA, so it is indirectly addressed by the framework. However, its relevance may justify further investigation and greater representation in the tool. This could push game designers to develop more refined taskifications that exploit games' reward systems in a better way to engage players.

Another interesting opportunity to improve the framework comes from the reasoning sprung in Pilot 3. Team C used control groups in their design to ensure good data retrieval. Various papers in the literature review investigated how to reduce errors in data gathered from crowdsourcing projects as well. However, it appeared to be more a programming problem rather than a design one, hence the framework did not contain a reference to it. Team C instead demonstrated that it has consequences on the game design. Their reasoning on a control group led them to change the narration many times. Their struggle was to make sense out of the high number of known wrong options that they inserted as a control group. Although it was not a request of the workshop to consider how to retrieve good data, an actual taskification would require a system to ensure data quality. Acknowledging that such a system impacts also the game design, it is important to include it in the framework. Game designers should design how to avoid cheating or other damaging phenomena together with developers. They need to know the system requirements also in terms of data quality to understand the influence that it can have on the design.

The framework was tested in a linear design process during the pilots, as the workshop was structured in a linear way. However, an iterative design process enables polishing the original idea by improving the design on each iteration. It usually contains these phases when applied to game design (Fig. 61): ideate, prototype and playtest (Bertolo and Mariani, 2014). Using an iterative design process to taskify a game would address considerations like the one from a member of team B. As aforementioned, this participant expressed some concern about the fact that her team worked on some elements of the framework and then did not consider them too much. The iterative design process is meant to help focus on all elements in the framework, as it leads to analyse and confront its design with the framework at each iteration.

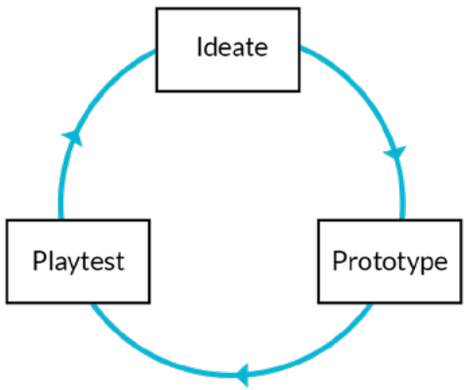
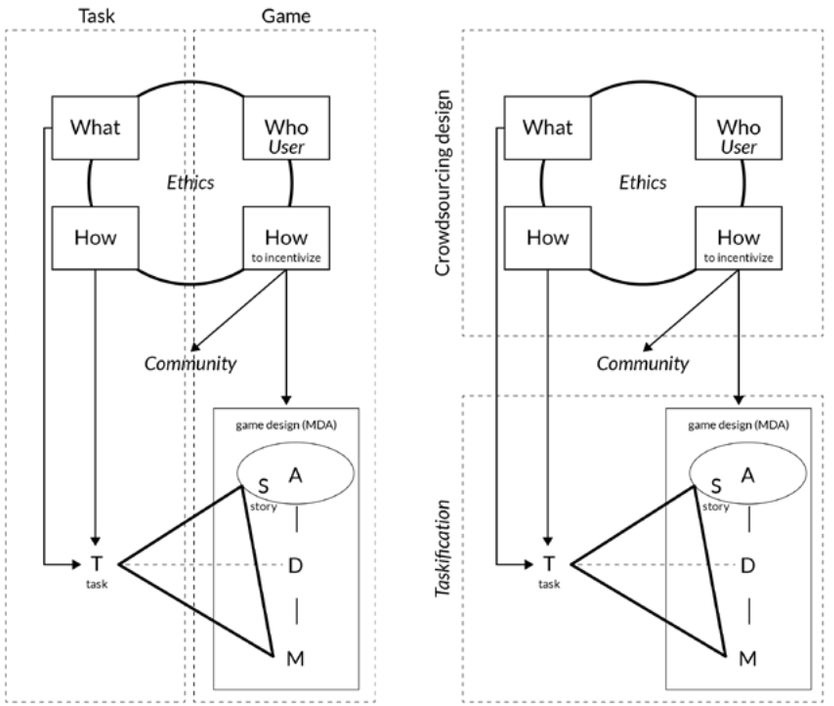


Fig. 61 A schema of the iterative design process.

Another opportunity is the suggestion from a member of team C of dividing the framework vertical (left: the task, right: the game) and providing a framework for the task analysis. The partition left-right seems to be an interesting option to the actual partition up-down. Indeed, during the pilots game designers reasoned more in terms of crowdsourcing as task than in term of crowdsourcing as a system. Hence, the partition left-right could improve the usability of the framework for game designers by following their mental models. At the same time, it could be harder for collaborators to join in the discussion, because the design of the framework would facilitate game designers' way of thinking. It might not be a major issue since the primary target is game designers. Careful testing may demonstrate which is the better solution.

Fig. 62 A comparison between vertical (left) and horizontal (right) framework partition.



There are some possible threats for good employment of the framework as well. First of all, the team involved in Pilot 3 (team C) raised the matter of efficiency. They discussed that this framework guides the taskifications design which may not be efficient in terms of data produced per playing time. Actually it could be an important requirement for the crowdsourcer. An easy way to address this issue could be to stress this factor in the framework, that, although is contained in the "How to crowdsource" element, may need to be more explicit. Game designers would hence consider also this element when designing for taskification. Anyway, it is a hypothetical issue not yet tested. Therefore, it is not possible to declare if the framework as it actually is guides the design of games that are inefficient for the crowdsourcer's needs. Even if the task was integrated into a long game activity, this does not assure fewer data would be produced than if it was integrated into a short game activity.

For example, a short but repetitive secondary mission could not encounter the interest of players. Therefore, less players would accept the mission and the overall time spent by players on that activity would be low. On the contrary, a long but captivating mission could engage many players, and the high number of players would raise the overall time spent by players in that activity. Time-efficient tasks and related missions may not be the best solutions to gain high participation rates. It is surely an interesting matter that could be further developed to understand its implications and consequences.

Another potential threat is the need for mentoring. The framework requires a little introduction and guidance in the first use. A member of team C, in Pilot 3, made clear that mentoring is actually crucial. This could be a barrier to usage for game designers and it would be necessary to test whether a fully-explained version of the framework, like the one in par. 4.1.5 could prevent this issue. Indeed, I did not ask to the participant of the workshop to learn in advance and autonomously the tool, and it would be an interesting scenario to investigate. It would give space to unveil and experiment the usability of the tool in terms of clarity and accessibility.

At the same time, it is hard to say if they would be willing to spend a little time to understand the tool to save time later. Many participants in the workshops mentioned that the framework was useful exactly because it helped them design the taskification in a short time. It would be important to investigate the extent to which the target would be interested in the tool and the level of effort that they would be willing to take to benefit from it.

Last, a threat to the correct use of the framework could be the fear of disrupting or deviating the original game. It was common among participants. While it is helpful to avoid taking damaging design choices, it is important to avoid restraining creativity. It could be possible to introduce "disrupting" activities to counter this fear. These activities would suggest game designers do what they would not or brainstorming unlikely possibilities to integrate the task into the game. Therefore, they may realize that they can push the game boundaries and enrich the gaming experience without compromising the game itself. New opportunities may arise if game designers let them more freedom of exploring ideas.

5.2 Contributions to knowledge

This thesis has investigated a promising and understudied concept in the field of game-based crowdsourcing systems, namely taskification. The major contributions of this study are the following:

1. the positioning of taskification in relation to gamification and SG development;
2. the definition of an ordered collection of guidelines for the development of game-based crowdsourcing systems;
3. a theoretical framework to design game taskification.

The topic of this thesis emerged from a wide literature review on game-based crowdsourcing, which made it clear the peculiarities of taskification with respect to SGs design and gamification. The three approaches have been introduced to then expose the uniqueness of taskification in comparison to gamification and SGs design. By analyzing the relation between the game and the task in all three cases, it is clarified what distinguishes taskification from the other two. In particular, taskification is a less intrusive approach than gamification and SGs design because it operates only on a small portion of the experience. Hence, the first contribution of this thesis is the positioning taskification and the discussion on why it should be studied separately.

A second contribution is the wide set guidelines on game-based crowdsourcing systems that have been collected and cluste-

red. This collection gathers information from the field of gamified crowdsourcing systems, GWAPs, serious games for crowdsourcing. It provides useful insights to design better game-based crowdsourcing systems. Clusters improve their readability and show recurrent and relevant themes. Moreover, the collection enables designers to roughly understand which may be the preferable solution between a taskified game, a gamified task or a SG for their particular purposes.

However, the main contribution to knowledge provided by this thesis is the definition of a tool to guide the design games taskification for crowdsourcing, namely a theoretical framework that sums the relevant aspects of both games and crowdsourcing to sustain the seamless design of additive crowdsourcing systems into the game structure. The tool relies on interdisciplinary knowledge and in particular combines two frameworks: the MDA (Hunicke et al., 2004) and Simperl's (Simperl, 2015) framework for designing crowdsourcing. Moreover, the diegetic connectivity approach (Lane and Prestopnik, 2017) relates the task with the MDA. The so-formed framework was augmented with the relevant topics derived from the guidelines review, i.e. the umbrella terms used to describe the identified clusters. The tool demonstrates great potential to reach its aim, namely guiding designers into the taskification process to combine games and crowdsourcing systems. Some aspects may be further improved, but the results are already great.

5.3 Directions for future research

As suggested in par. 5.1.1, there are some aspects that can benefit from further study and experimentation. In particular, there are three areas of intervention that emerge from the current investigation. Those areas are:

1. structure;
2. process;
3. evaluation.

The first area is **structure**. It describes the components of the framework, its readability and their relations. Probably, the most interesting aspect to study in this area is the definition of a framework for task analysis and design. A member of team C, in Pilot 3 (par. 4.2.3), explained that the framework supports the game analysis and design through the MDA, while it describes the task as a single element. The element could be expanded to create a framework for task analysis and design providing interesting features to think about for designers. Moreover, a greater focus on the task instead than on the crowdsourcing system as a whole could improve the design process. Beyond that, some elements could be further explored and others could be added to the framework. "Ethics" and "Community" were little employed in the workshop. A greater exploration of those elements may provide new knowledge on how to design a taskification. The relevance

of two new elements emerged from the data analysis, namely the reward system and control groups. These elements could improve the framework by guiding game designers to detail more in-depth their taskifications. However, this is a suggestion: these elements could change for the better the current structure of the tool,but they would have to result useful to game designers as well as the other elements in the framework. Hence, it would be necessary to test the relevance of reward system and control groups for design taskification in a dedicated set of pilots. Future research can investigate this hypothesis.

The second area is **process**. It questions how to employ the framework in the design process and exploit it to shape and boost design activities. It is surely crucial to test the framework with mixed groups and in a professional pipeline. The framework would be extremely useful if it could support discussion between different specialists. It does contain and bridge knowledge from several fields of study. For this reason, it might be fertile ground for encounters between various professionals. Future research could study whether it actually enables discussion and confrontation. At the same time, it is necessary to test the framework with professionals and/or on an actual project. It is a different environment from the one tested in the current study and could provide additional insights into the usage of the tool. Professionals could consider pipelines and other factors that unveils other critical aspects of game design for taskification. Surely an aspect that could be tested in this regard is the effectiveness in guiding an iterative design process, common in professional environments. The workshop tested the framework in a linear process, but an iterative process has great advantages as explained in par. 5.1.1. Testing the framework in this environment would provide the most accurate data on the effectiveness of the tool in guiding the taskification design process.

The third area is **evaluation**. It questions how to assess the quality of a taskification to help designers go through various iterations. The actual framework guides the design of taskification, but it does not provide feedback on the quality of it, namely its actual playability. Surely, the playtesting phase can shed light on this concern. However, the scope of this study was to design a tool to support the ideation of taskification, which is one of the steps of an iterative process (ideate, test, measure). It was not tested in this study and it has already been said that applying the framework in an iterative approach, which contains even the playtesting phase, should be investigated in the future. Alongside the analysis of such a condition, the expansion of the taskification practice and literature could produce even more knowledge to guide the design of taskification. It is important to keep the framework updated and collect and organize this knowledge into its structure. Experimentation even on extreme cases, like games and crowdsourcing projects which seem inconsistent, can produce interesting results. However, future research should pay attention to the efficiency of taskification design. As discussed during the focus group of Pilot 3 (par. 4.2.3), system efficiency is particularly important. A taskification is efficient if low resources employed to taskify the game – such as time, money, effort – result in a high number of contributions. A design choice is more efficient than another if it lowers the resources or if it increases the number of contributions. It is surely a crucial aspect to consider when designing a taskification. Future research could shed light on how to understand the potential efficiency of a game-based crowdsourcing system so that designers can attempt to find the best solution.

What emerged from the study on the ground of this research is that there are still a few examples of taskified games and little comprehension of the phenomenon. This research is a first step toward the definition of tools and theories to understand, analyze and shape taskified games. It surely does not expect to be exhaustive: as the phenomenon grows, research should investigate it more.

Augmenting games and turning them into crowdsourcing systems that convey and direct human power is a thrilling idea and uncovers a realm of relevant possibilities for various scientific fields. The debate on taskification is just at the beginning, and it has a lot to say.

References

Andersen, M., 2009. Four crowdsourcing lessons from the Guardian's (spectacular) expenses-scandal experiment [WWW Document]. Nieman Journal. Lab. URL https://www.niemanlab.org/2009/06/four-crowdsourcing-lessons-from-the-guardians-spectacular-expenses-scandal-experiment/ (accessed 11.7.19).

Anderson, G., 2015. Designing for social impact [WWW Document]. O'Reilly Media. URL https://www.oreilly.com/content/designing-for-social-impact/ (accessed 3.19.20).

Barker, P., 2018. How EVE Online's Project Discovery is helping real-world astronomy [WWW Document]. www.redbull.com. URL https://www.redbull.com/int-en/eve-online-project-discovery-interview (accessed 3.9.20).

Bayrak, A.T., 2019. Games as a Catalyst for Design for Social Innovation. Unlocking legendary tools. Des. J. 22, 1409–1422. https://doi.org/10.1080/14606925.2019.1594966

Berdichevsky, D., Neuenschwander, E., 1999. Toward an ethics of persuasive technology. Commun. ACM 42, 51–58. https://doi.org/10.1145/301353.301410

Bertolo, M., Mariani, I., 2014. Game design: gioco e giocare tra teoria e progetto. Pearson, Milano; Torino.

Bogost, I., 2013. UX Week 2013 | Ian Bogost | Fun.

Bogost, I., 2011a. Persuasive Games: Exploitationware [WWW Document]. Gamasutra. URL https://www.gamasutra.com/view/feature/134735/persuasive_games_exploitationware.php (accessed 1.7.20).

Bogost, I., 2011b. Gamification is Bullshit [WWW Document]. Bogost.com. URL http://bogost.com/writing/blog/gamification_is_bullshit/ (accessed 2.5.20).

Bogost, I., 2007. Persuasive games: the expressive power of videogames. MIT Press, Cambridge, MA.

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V., Shirk, J., 2009. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. BioScience 59, 977–984. https://doi.org/10.1525/bio.2009.59.11.9

Boons, M., Stam, D., Barkema, H.G., 2015. Feelings of Pride and Respect as Drivers of Ongoing Member Activity on Crowdsourcing Platforms. J. Manag. Stud. 52, 717–741. https://doi.org/10.1111/joms.12140

Bowser, A., Hansen, D., He, Y., Boston, C., Reid, M., Gunnell, L., Preece, J., 2013. Using gamification to inspire new citizen science volunteers, in: Proceedings of the First International Conference on Gameful Design, Research, and Applications - Gamification '13. Presented at the the First International Conference, ACM Press, Toronto, Ontario, Canada, pp. 18–25. https://doi.org/10.1145/2583008.2583011

Bowser, A., Hansen, D., Preece, J., He, Y., Boston, C., Hammock, J., 2014. Gamifying citizen science: a study of two user groups, in: Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW Companion '14. Presented at the the companion publication of the 17th ACM conference, ACM Press, Baltimore, Maryland, USA, pp. 137–140. https://doi.org/10.1145/2556420.2556502

Brabham, D.C., 2013. Crowdsourcing, The MIT Press essential knowledge series. The MIT Press, Cambridge, Massachusetts ; London, England.

Brabham, D.C., 2008. Moving the crowd at iStockphoto: The composition of the crowd and motivations for participation in a crowdsourcing application. First Monday 13.

Brito, J., Vieira, V., Duran, A., 2015. Towards a Framework for Gamification Design on Crowdsourcing Systems: The G.A.M.E. Approach, in: 2015 12th International Conference on Information Technology - New Generations. Presented at the 2015 12th International Conference on Information Technology - New Generations (ITNG), IEEE, Las Vegas, NV, USA, pp. 445–450. https://doi.org/10.1109/ITNG.2015.78

Caillois, R., 2001. Man, play, and games. University of Illinois Press, Urbana.

Connolly, T.M., Boyle, E.A., MacArthur, E., Hainey, T., Boyle, J.M., 2012. A systematic literature review of empirical evidence on computer games and serious games. Comput. Educ. 59, 661–686. https://doi.org/10.1016/j.compedu.2012.03.004

Cooper, S., Baker, D., Popović, Z., Treuille, A., Barbero, J., Leaver-Fay, A., Tuite, K., Khatib, F., Snyder, A.C., Beenen, M., Salesin, D., 2010a. The challenge of designing scientific discovery games, in: Proceedings of the Fifth International Conference on the Foundations of Digital Games - FDG '10. Presented at the the Fifth International Conference, ACM Press, Monterey, California, pp. 40–47. https://doi.org/10.1145/1822348.1822354

Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z., players, F., 2010b. Predicting protein structures with a multiplayer online game. Nature 466, 756–760. https://doi.org/10.1038/nature09304

Coronavirus Spike Protein Binder Design [WWW Document], n.d. URL https://fold.it/portal/node/2008926 (accessed 9.5.20).

Creswell, J.W., 2009. Research design: qualitative, quantitative, and mixed methods approaches, 3rd ed. ed. Sage Publications, Thousand Oaks, Calif.

Curtis, V., 2015. Motivation to Participate in an Online Citizen Science Game: A Study of Foldit. Sci. Commun. 37, 723–746. https://doi.org/10.1177/1075547015609322

Daniel, A., Flew, T., 2010. The Guardian Reportage of the UK MP Expenses Scandal: a Case Study of Computational Journalism 11.

Deng, X. (Nancy), Joshi, K.D., Galliers, R.D., 2016. The Duality of Empowerment and Marginalization in Microtask Crowdsourcing: Giving Voice to the Less Powerful Through Value Sensitive Design. MIS Q. 40, 279–302. https://doi.org/10.25300/MISQ/2016/40.2.01

Deterding, S., Dixon, D., Khaled, R., Nacke, L., 2011. From Game Design Elements to Gamefulness: Defining "Gamification" 7.

Eiben, C.B., Siegel, J.B., Bale, J.B., Cooper, S., Khatib, F., Shen, B.W., Players, F., Stoddard, B.L., Popovic, Z., Baker, D., 2012. Increased Diels-Alderase activity through backbone remodeling guided by Foldit players. Nat. Biotechnol. 30, 190–192. https://doi.org/10.1038/nbt.2109

Estellés-Arolas, E., González-Ladrón-de-Guevara, F., 2012. Towards an integrated crowdsourcing definition. J. Inf. Sci. 38, 189–200. https://doi.org/10.1177/0165551512437638

Eustace, A., 2011. A fall spring-clean. Off. Google Blog. URL https://googleblog.blogspot.com/2011/09/fall-spring-clean.html (accessed 11.4.19).

Flanagan, M., Punjasthitkul, S., Seidman, M., Kaufman, G., Carini, P., 2013. Citizen Archivists at Play: Game Design for Gathering Metadata for Cultural Heritage Institutions 13.

Flew, T., Spurgeon, C., Daniel, A., Swift, A., 2010. The Promise Of Computational Journalism. Journal. Pract. 6, 157–171. https://doi.org/10.1080/17512786.2011.616655

Geiger, D., Schader, M., 2014. Personalized task recommendation in crowdsourcing information systems — Current state of the art. Decis. Support Syst. 65, 3–16. https://doi.org/10.1016/j.dss.2014.05.007

Geri, N., Gafni, R., Bengov, P., 2017. Crowdsourcing as a business model: Extrinsic motivations for knowledge sharing in user-generated content websites. J. Glob. Oper. Strateg. Sourc. 10, 90–111. https://doi.org/10.1108/JGOSS-05-2016-0018

Gray, D., Brown, S., Macanufo, J., 2010. Gamestorming: a playbook for innovators, rulebreakers, and changemakers, First edition. ed. O'Reilly, Beijing Cambridge Farnham Köln Sebastopol Tokyo.

Hamari, J., Koivisto, J., Sarsa, H., 2014. Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification, in: 2014 47th Hawaii International Conference on System Sciences. Presented at the 2014 47th Hawaii International Conference on System Sciences (HICSS), IEEE, Waikoloa, HI, pp. 3025–3034. https://doi.org/10.1109/HICSS.2014.377

Hicks, 2009. EASY: The Guardian's Crowdsource Game. News Games Ga. Tech.

Huizinga, J.H., 1980. Homo ludens. Routledge.

Hundal, S., 2017. Games for Social Good [WWW Document]. Medium. URL https://medium.com/babble-on/games-for-social-good-7e8e122aa102 (accessed 3.15.20).

Hunicke, R., LeBlanc, M., Zubek, R., 2004. MDA: A Formal Approach to Game Design and Game Research 5.

Huotari, K., Hamari, J., 2016. A definition for gamification: anchoring gamification in the service marketing literature. Electron. Mark. 27, 21–31. https://doi.org/10.1007/s12525-015-0212-z

Iacovides, I., Jennett, C., Cornish-Trestrail, C., Cox, A.L., 2013. Do games attract or sustain engagement in citizen science?: a study of volunteer motivations 6.

Interview with Jennifer Couch and Dave Miller, 2018. . Citiz. Sci. Games. URL https://citizensciencegames.com/nih-funding-games/ (accessed 3.25.20).

Jenkins, H., 2006. Convergence culture: where old and new media collide. New York University Press, New York.

Jenkins, H., Ford, S., Green, J., 2013. Spreadable media: creating value and meaning in a networked culture, Postmillennial pop. New York University Press, New York ; London.

Jorgensen, D., 1989. Participant Observation. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States of America. https://doi.org/10.4135/9781412985376

Kafai, Y.B. (Ed.), 2008. Beyond Barbie and Mortal Kombat: new perspectives on gender and gaming. MIT Press, Cambridge, Mass.

Kaufman, G., Flanagan, M., Punjasthitkul, S., 2016. Investigating the Impact of "Emphasis Frames" and Social Loafing on Player Motivation and Performance in a Crowdsourcing Game, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16. Presented at the the 2016 CHI Conference, ACM Press, Santa Clara, California, USA, pp. 4122–4128. https://doi.org/10.1145/2858036.2858588

Khatib, F., Cooper, S., Tyka, M.D., Xu, K., Makedon, I., Popovic, Z., Baker, D., Players, F., 2011. Algorithm discovery by protein folding game players. Proc. Natl. Acad. Sci. 108, 18949–18953. https://doi.org/10.1073/pnas.1115898108

Khatib, Firas, DiMaio, F., Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., Zabranska, H., Pichova, I., Thompson, J., Popović, Z., Jaskolski, M., Baker, D., 2011. Crystal structure of a monomeric retroviral protease solved by protein folding game players. Nat. Struct. Mol. Biol. 18, 1175–1177. https://doi.org/10.1038/nsmb.2119

Kim, T.W., Werbach, K., 2016. More than just a game: ethical issues in gamification. Ethics Inf. Technol. 18, 157–173. https://doi.org/10.1007/s10676-016-9401-5

Lafourcade, M., Joubert, A., Le Brun, N., 2015. Games with a Purpose (GWAPS) 160.

Lane, N., Prestopnik, N.R., 2017. Diegetic Connectivity: Blending Work and Play with Storytelling in Serious Games, in: Proceedings of the Annual Symposium on Computer-Human Interaction in Play. Presented at the CHI PLAY '17: The annual symposium on Computer-Human Interaction in Play, ACM, Amsterdam The Netherlands, pp. 229–240. https://doi.org/10.1145/3116595.3116630

Laurel, B. (Ed.), 2003. Design research: methods and perspectives. MIT Press, Cambridge, Mass.

Lazzaro, N., 2008. The four fun keys.

LeBlanc, M., 2000. Game Developers Conference.

Leifsson, H., Bjarkason, J.Ö., Sigurþórsson, H., 2015. Project Discovery - Advancing scientific research by implementing citizen science in EVE Online 32.

Lévy, P., 1995. Collective Intelligence: Mankind's Emerging World in Cyberspace. New York: Plenum.

Maler, E., 2019. Games that Change Lives: Social Impact Done Right.

Manzini, E., 2014. Making Things Happen: Social Innovation and Design. Des. Issues 30, 57–66. https://doi.org/10.1162/DESI_a_00248

Mariani, I., 2016. Meaningful negative experiences within games for social change. Designing and Analysing Games as Persuasive Communication Systems.

McAloon, A., 2017. Round 2 of Project Discovery seeks to use Eve Online to discover actual planets [WWW Document]. www.gamasutra.com. URL www.gamasutra.com/view/news/292247/Round_2_of_Project_Discovery_seeks_to_use_Eve_Online_to_discover_actual_planets.php (accessed 3.10.20).

McGonigal, J., 2011. Reality is broken: why games make us better and how they can change the world. Penguin Group, New York.

Michael, D., 2006. Serious games: games that educate, train and inform. Thomson Course Technology, Boston, Mass.

Morschheuser, B., Hamari, J., Koivisto, J., Maedche, A., 2017. Gamified crowdsourcing: Conceptualization, literature review, and future agenda. Int. J. Hum.-Comput. Stud. 106, 26–43. https://doi.org/10.1016/j.ijhcs.2017.04.005

Mulgan, G., Tucker, S., Ali, R., Sanders, B., University of Oxford, Skoll Centre for Social Entrepreneurship, 2007. Social innovation: what it is, why it matters and how it can be accelerated. Young Foundation, London.

Murray, J.H., 2017. Hamlet on the holodeck: the future of narrative in cyberspace, Updated edition. ed. The MIT Press, Cambridge, Massachusetts.

Murray, R., Caulier-Grice, J., Mulgan, G., 2010. THE OPEN BOOK OF SOCIAL INNOVATION 224.

Narula, H., 2019. A billion new players are set to transform the gaming industry. Wired UK.

O'Reilly, T., 2005. Web 2.0: Compact Definition... 1.

Paharia, R., 2012. Gamification Means Amplifying Intrinsic Value. interactions, Gamification: designing for motivation 19, 14. https://doi.org/10.1145/2212877.2212883

Prestopnik, N., Crowston, K., 2012. Purposeful Gaming & Socio-Computational Systems: A Citizen Science Design Case 10.

Prestopnik, N., Crowston, K., Wang, J., 2017. Gamers, citizen scientists, and data: Exploring participant contributions in two games with a purpose. Comput. Hum. Behav. 68, 254–268. https://doi.org/10.1016/j.chb.2016.11.035

Quinn, A.J., Bederson, B.B., 2011. Human computation: a survey and taxonomy of a growing field, in: Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11. Presented at the the 2011 annual conference, ACM Press, Vancouver, BC, Canada, p. 1403. https://doi.org/10.1145/1978942.1979148

Raddick, J.M., Bracey, G., Gay, P.L., Lintott, C.J., Cardamone, C., Murray, P., Schawinski, K., Szalay, A.S., Vandenberg, J., 2013. Galaxy Zoo: Motivations of Citizen Scientists. Astron. Educ. Rev. 12. https://doi.org/10.3847/AER2011021

Randhawa, K., Wilden, R., West, J., 2019. Crowdsourcing without profit: the role of the seeker in open social innovation: Crowdsourcing without Profit. RD Manag. 49, 298–317. https://doi.org/10.1111/radm.12357

Ridge, M., 2011. Museum Crowdsourcing Games: Improving Collections Through Play (and some thoughts on the future of museums) 25.

Ritterfeld, U., Cody, M.J., Vorderer, P. (Eds.), 2009. Serious games: mechanisms and effects. Routledge, New York.

Robertson, M., 2010. Can't Play, Won't Play [WWW Document]. Kotaku. URL https://kotaku.com/cant-play-wont-play-5686393 (accessed 1.6.20).

Robson, K., Plangger, K., Kietzmann, J.H., McCarthy, I., Pitt, L., 2015. Is it all a game? Understanding the principles of gamification. Bus. Horiz. 58, 411–420. https://doi.org/10.1016/j.bushor.2015.03.006

Rogers, S., 2009. How to crowdsource MPs' expenses. The Guardian.

Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D., Jacobs, D., 2012. Dynamic changes in motivation in collaborative citizen-science projects, in: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work - CSCW '12. Presented at the the ACM 2012 conference, ACM Press, Seattle, Washington, USA, p. 217. https://doi.org/10.1145/2145204.2145238

Ryan, R.M., Deci, E.L., 2000a. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. Am. Psychol. 67.

Ryan, R.M., Deci, E.L., 2000b. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. Contemp. Educ. Psychol. 25, 54–67. https://doi.org/10.1006/ceps.1999.1020

Ryan, R.M., Rigby, C.S., Przybylski, A., 2006. The Motivational Pull of Video Games: A Self-Determination Theory Approach. Motiv. Emot. 30, 344–360. https://doi.org/10.1007/s11031-006-9051-8

Saini, A., 2008. Solving the web's image problem.

Salen, K.T., Zimmerman, E., 2004. Rules of play: game design fundamentals. MIT Press, Cambridge, Mass.

Sandovar, A., Braad, E., Streicher, A., Söbke, H., 2016. Ethical Stewardship: Designing Serious Games Seriously, in: Dörner, R., Göbel, S., Kickmeier-Rust, M., Masuch, M., Zweig, K. (Eds.), Entertainment Computing and Serious Games, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 42–62. https://doi.org/10.1007/978-3-319-46152-6_3

Schrier, K., 2016. Knowledge games: how playing games can solve problems, create insight, and make change, Tech.edu. Johns Hopkins University Press, Baltimore.

Seaborn, K., Fels, D.I., 2015. Gamification in theory and action: A survey. Int. J. Hum.-Comput. Stud. 74, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006

Silvertown, J., 2009. A new dawn for citizen science. Trends Ecol. Evol. 24, 467–471. https://doi.org/10.1016/j.tree.2009.03.017

Simperl, E., 2015. How to Use Crowdsourcing Effectively: Guidelines and Examples. Liber Q. 25, 18. https://doi.org/10.18352/lq.9948

Skarlatidou, A., Hamilton, A., Vitos, M., Haklay, M., 2019. What do volunteers want from citizen science technologies? A systematic literature review and best practice guidelines. J. Sci. Commun. 18. https://doi.org/10.22323/2.18010202

Spradley, J.P., 1980. Participant observation. Holt, Rinehart and Winston, New York.

Standing, S., Standing, C., 2018. The ethical use of crowdsourcing. Bus. Ethics Eur. Rev. 27, 72–80. https://doi.org/10.1111/beer.12173

Steinkhuehler, C., 2017. Keynote - Ten Important Findings From The Research On Games For Impact.

Stokes, B., O'Shea, G., Walden, N., Nasso, F., Mariutto, G., Hill, A., Burak, A., 2016. Impact with Games: A Fragmented Field.

Suits, B., 2005. The grasshopper: games, life and utopia. Broadview Press, Peterborough, Ont.

Sullivan, D.P., Winsnes, C.F., Åkesson, L., Hjelmare, M., Wiking, M., Schutten, R., Campbell, L., Leifsson, H., Rhodes, S., Nordgren, A., Smith, K., Revaz, B., Finnbogason, B., Szantner, A., Lundberg, E., 2018. Deep learning is combined with massive-scale citizen science to improve large-scale image classification. Nat. Biotechnol. 36, 820–828. https://doi.org/10.1038/nbt.4225

Sun, Y., Fang, Y., Lim, K.H., 2012. Understanding sustained participation in transactional virtual communities. Decis. Support Syst. 53, 12–22. https://doi.org/10.1016/j.dss.2011.10.006

Sun, Y., Wang, N., Peng, Z., 2011. Working for one penny: Understanding why people would like to participate in online tasks with low payment. Comput. Hum. Behav. 27, 1033–1041. https://doi.org/10.1016/j.chb.2010.12.007

Surowiecki, J., 2004. The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations, 1st ed. ed. Doubleday, New York.

Sutton-Smith, B., 2001. The ambiguity of play, 2. printing, 1. Harvard Univ. Press paperback ed. ed. Harvard Univ. Press, Cambridge, Mass.

The rundown on coronavirus [WWW Document], 2020. URL https://fold.it/portal/node/2008963 (accessed 3.6.20).

Travis, D., 2016. Desk research: the what, why and how [WWW Document]. URL https://www.userfocus.co.uk/articles/desk-research-the-what-why-and-how.html (accessed 9.4.20).

von Ahn, L., 2005. Human Computation. Carnegie Mellon University.

von Ahn, L., Dabbish, L., 2008. Designing games with a purpose. Commun. ACM 51, 57. https://doi.org/10.1145/1378704.1378719

von Ahn, L., Dabbish, L., 2004. Labeling images with a computer game, in: Proceedings of the 2004 Conference on Human Factors in Computing Systems - CHI '04. Presented at the the 2004 conference, ACM Press, Vienna, Austria, pp. 319–326. https://doi.org/10.1145/985692.985733

Wang, X., Goh, D.H.-L., Lim, E.-P., 2020. Understanding Continuance Intention toward Crowdsourcing Games: A Longitudinal Investigation. Int. J. Human–Computer Interact. 1–10. https://doi.org/10.1080/10447318.2020.1724010

Werbach, K., 2014. (Re)Defining Gamification: A Process Approach, in: Spagnolli, A., Chittaro, L., Gamberini, L. (Eds.), Persuasive Technology. Springer International Publishing, Cham, pp. 266–272. https://doi.org/10.1007/978-3-319-07127-5_23

Who coined the term "gamification"? [WWW Document], 2010. URL https://www.quora.com/Who-coined-the-term-gamification (accessed 2.6.20).

Wolf, M.J.P., 2013. Building imaginary worlds: the theory and history of subcreation. Routledge, New York.

Ye, H. (Jonathan), Kankanhalli, A., 2017. Solvers' participation in crowdsourcing platforms: Examining the impacts of trust, and benefit and cost factors. J. Strateg. Inf. Syst. 26, 101–117. https://doi.org/10.1016/j.jsis.2017.02.001

Yee, N., 2006. Motivations for Play in Online Games. Cyberpsychol. Behav. 9, 772–775. https://doi.org/10.1089/cpb.2006.9.772

Zagal, J.P., Björk, S., Lewis, C., 2013. Dark Patterns in the Design of Games 8.

Zheng, H., Li, D., Hou, W., 2011. Task Design, Motivation, and Participation in Crowdsourcing Contests. Int. J. Electron. Commer. 15, 57–88. https://doi.org/10.2753/JEC1086-4415150402

Zichermann, G., Cunningham, C., 2011. Gamification by design: implementing game mechanics in web and mobile apps, 1st. ed. ed. O'Reilly Media, Sebastopol, Calif.