Designer? User? Player! – A Framework of Opportunities to Improve Interpersonal Dynamics During Usability Testing Through Gamification

Yating Li
Anhalt University of Applied Sciences,
SAP Germany
leahliya0825@gmail.com

Katja Thoring
Anhalt University of Applied Sciences,
Delft University of Technology
katja@thoring.com

Mito Mihelic
Anhalt University of Applied
Sciences
mitja.mito@gmail.com

Abstract

Usability testing is an important technique in the design process. However, most existing research takes usability testing as a research method, not as the research topic itself. This paper explores the relationships and the interpersonal dynamics between participants and researchers during usability testing. In this study, we use interviews and observations to uncover interpersonal dynamics that might be hidden in the complex multilevel structure of the testing process. Our findings suggest that interpersonal dynamics impact the testing experience and results: participants in a nervous or tense atmosphere may not perform as in real life or conceal their real thoughts. To improve the usability testing experience and results, we developed a framework of opportunities for improving interpersonal dynamics through gamification.

1. Introduction

Usability testing, a well-known technique in the design process, is an activity that focuses on observing users working with a product, performing tasks that are real and meaningful to them, with the goal to identify possible problems of the system [1]. This can be seen as an irreplaceable usability practice since it provides direct input on how real users use the system.

However, the quality of the testing experience and results might be affected by different factors, such as whether the tasks and the questionnaire design are reasonable, and whether the participants can perform the task in a real-life manner [1]. In addition, due to funding restrictions or other reasons, many companies don't have enough user researchers, so in the real field, many usability testings are conducted by designers [4], who possibly lack expertise in the user research aspect.

Consequently, we will use the terms "designer" and "participant" throughout this paper. The "designer" is the person conducting the usability test, while the "participant" is the user who is invited as the testee.

Most existing research takes usability testing as a research method [2], [3], not as the topic to investigate.

The research about the method improvements usually focused on different aspects, such as testing tasks [5], [6], or the application of the method in the enterprise [7], [8]. There is some existing research about psychological and emotional aspects of usability testing, for example, Kara [9] has investigated the conversation with participants during usability testing. Still, research on interpersonal dynamics between designers and users is scarce. One of the main goals of this paper is to study whether interpersonal dynamics are a factor that might affect the experience and results in usability testing. Therefore, our first research question is:

RQ 1: How do interpersonal dynamics between designers and participants impact the quality and experience of usability testing?

To answer this research question, we conducted interviews with professional designers and observed an exemplary usability testing process.

The second goal of this paper is to develop a framework that summarizes different possibilities to improve interpersonal dynamics during usability testing through gamification. The concept of gamification refers to the use of game elements in non-game contexts [10]. Gamification is still in the process of methodological verification and academic validation, but it has raised much interest in the industry [11]. And its unique "gameful" experience can increase enjoyment, freedom, collaboration, and engagement in the real-world experience. We assume that some of its features are suitable for improving interpersonal dynamics. Therefore, our aim with this study is to investigate how gamification could enhance the usability testing method while keeping the original setup of this method intact. Here arises the second research question:

RQ 2: How can we improve the interpersonal dynamics between designers and participants during the usability testing by gamification?

To answer this research question, we first tried to identify opportunity areas within the usability testing process. Then, we matched the potential game elements with different improvement goals. Our results are

HǐCSS

summarized in a framework that can guide designers to create their own game activities to improve interpersonal dynamics in usability testing.

2. Literature Review

To answer our first research question, we started by conducting a literature review using relevant keywords and concepts. We divided the identified sources into three categories: literature on "usability testing", on "interpersonal dynamics" and on "gamification".

The first section "usability testing" includes the background research of the usability testing method itself and some existing tools that support usability testing. The section "interpersonal dynamics" includes the definition of interpersonal dynamics from a psychological perspective. The section on "gamification" includes the basic concepts of gamification, existing toolkits or frameworks of gamification, and some existing gamification designs with similar goals as ours.

2.1. Literature on Usability Testing

Usability testing is an activity that focuses on observing users working with a product or system [1]. The literature distinguished between in-person usability testing and remote usability testing [5]. In this research, we only discuss the interpersonal dynamics in in-person usability testing.

The goal of usability testing is to evaluate the product or system by observing how users use and interact with it. The core elements in the in-person usability testing are the facilitator, the tasks, and the participant [1],[5]. As mentioned before, in this research we will use the terms "designer" instead of "facilitator" or "researcher" for the person conducting the usability testing throughout this paper.

During the usability-testing process, the designer asks the participant to perform tasks, usually using specific user interfaces or a product system. While the participant completes each task, the designer observes the participant's behavior and listens for feedback [1]. The participants are usually real or exemplary users of the product or service being studied. In usability testing, they are often asked to perform tasks and think out loud (called the "think-aloud method"). These are typically tasks that the participants might perform in real life with the product. In this way, the designers can understand the participants' behaviors, goals, ideas, and motivations to find critical points and potential for improving the product or system [5], [6].

The question of whether the participants can perform the task in a real-life manner will affect the

quality and credibility of the testing and results. However, today's usability testing is usually carried out by designers or even non-designers [4] rather than user researchers who have expertise in user research. There are many reasons for this phenomenon. For example, the companies don't have (enough) user researchers or designers who want to participate in the testing process of their own designs.

Most existing research takes usability testing as a research method, not the topic. The research about the method improvements usually focused on different aspects such as testing tasks [5], [6], or application of the method in the enterprise [7], [8].

There is some existing research about psychological and emotional aspects of usability testing, for example, Kara's research about the conversation with participants during usability testing, which proposed three techniques (echo, boomerang, and Columbo) for interrupting or answering users while facilitating a usability test or other behavioral research study [9].

In summary, the literature shows that the relationships and interpersonal dynamics between participants and researchers during usability testing may affect the process and results of the testing. At present, research or design for this direction is scarce, which warrants our goal with this paper to contribute to this field.

2.2. Literature on Interpersonal Dynamics

There are two different meanings of dynamics: structural dynamics and temporal dynamics. Structural dynamics, introduced by Freud, has a rich history in psychoanalysis. It is described as the "play of forces in the mind" [12]. Temporal dynamics, stemming from systems theory, is described as the "unfolding of behavior over time" [13].

The interpersonal dynamics in usability testing can be classified as temporal dynamics. It is at a smaller timescale, people engage in many distinct interactions over a case of an event, and each individual's interpersonal behavior varies in important ways across these interactions [14]. The authors present an example of the interpersonal dynamics between the client and the therapist during a psychotherapy session.

The social interaction between individuals reveals the dynamics of interpersonal relationships. According to [14], there are some models and frameworks that can be used to evaluate these moment-to-moment patterns in social interactions.

We chose one of these models that may be more suitable for the process of usability testing to analyze the relationship between designers and participants. The model was proposed by Brewin in 1988 [15]: the

interpersonal circumplex. This model is for conceptualizing, organizing, and evaluating interpersonal behavior, traits, and motivations. As shown in Figure 1, the interpersonal circumplex is defined by two orthogonal axes: a vertical axis and a horizontal axis. Therefore, each point in the interpersonal surrounding space can represent the temporal interpersonal dynamic.
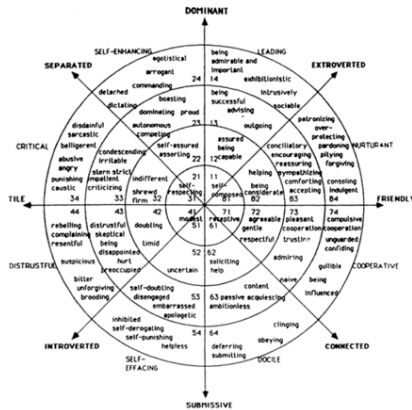

Figure 1. The interpersonal circumplex

We used a simplified version of the interpersonal circumplex model in our studies to help designers describe the type and degree of the experienced interpersonal dynamics in their usability testing (as described in Section 3).

2.3. Literature on Gamification

Gamification is generally defined as "the use of game design elements in non-game contexts" [10]. Even though gamification is still in the process of conceptual definition, methodological verification, and academic validation, it has raised much interest in the industry as a new way of making work more fun and motivating [11]. As the understanding of gamification has been solely based on the act of adding systemic game elements into services, Hamari [11] proposed a gamification understanding from the perspective of service marketing: a process of enhancing a service with affordances for gameful experiences in order to support the user's overall value creation. This definition emphasizes the experiential nature of games and gamification instead of the systemic understanding, which is from the user's perspective.

In our study, we can use both definitions of gamification to understand our goal: usability testing can be considered as a core service, and we are trying to design a gamified service to enhance this core service and improve interpersonal dynamics; by implementing game design elements in usability testing (i.e. non-game contexts), the participants are directed towards a more "gameful" experience, which leads to an experience with increased enjoyment, freedom, collaboration, and engagement.

Most of the existing gamification approaches are limited to either motivational or educational aspects. However, there exist some gamification studies and designs with a similar goal as ours. For example, there is research from SAP's central user research team that wanted to enhance method literacy and empathic involvement in multidisciplinary teams [7]. Therefore, they introduced an emotional component into the process by applying slight modifications to user research methods by using game principles. Another research focuses on improving the collaboration between team members to deliver optimal outcomes in teamwork [16]. They present a framework to assist in

Table 1. Potential game types and game elements

Possible Game Types	Apply in Usability Testing Scenario			
Quiz games	A quiz game can be useful for helping participants understand the usability testing methods before they start the formal testing.			
Role-playing games	By exchanging the roles, participants and designers can not only get closer, but they can also rethink usability testing from the perspective of others and understand others' feelings.			
Icebreaker/Warm-up	An icebreaker or warm-up game can be useful for helping designer and participants get to know each other before the formal testing.			
Cooperative games	As the name suggests, cooperative games stress cooperation over competition. It can be useful for guiding designers and participants with game rules towards cooperative behavior.			
Adventure games	An Adventure game can be used to connect the entire usability testing. Participants and designers need to complete the task together to unlock the next stage, which provides attractive activities and cooperative environments.			
Possible Game Elements				**Apply in Usability Testing Scenario**
Rules	Quests	Hints	Road Maps	Content Unlocking	*Rules can control and ritualize player activities, positively affecting their experiences -- guiding designers and participants with game rules towards cooperative behavior.*
Objective	Levels	Points	Leaderboards	Achievements	*An Objective can strongly influence player's interest, commitment, and desire -- activities can be designed with multiple linked goals, which can influence participants' or designers' pleasure when they are doing tasks.*
Feedback	Rewards	Badges	Scores	Progress Bar	*Players need to be aware of their activity options and the rewards /feedback they will obtain -- design more fast, frequent, and clear feedback during the usability testing for both designers and participants*
Social interaction	Avatars	Teams	Social graph	Goods Exchanging	*Player's emotions in social context activities are generally more intense and pleasurable -- Provide different ways of communication, expression, interaction for designers and participants*

the analysis and gamification design of teamwork situations.

Based on existing research on game element categories and game types [17]–[19] and research on most used game elements in gamification [20], we selected some game types and game elements that have the potential to improve the interpersonal dynamic in usability testing. Table 1 shows a preliminary analysis of how they might apply in usability testing scenarios.

Our analysis suggests that gamification could be useful for improving interpersonal dynamics in usability testing in three ways. First, it can structure the process with goals and guide players with game rules towards cooperative behavior. Secondly, it can make it easier for participants to generate positive emotions, such as feedback and meaningful outcomes. Thirdly, it can help designers and participants to establish emotional connections faster, such as teamwork and social interaction.

3. Methodology

3.1. Qualitative Interviews

For our interview study, we selected six professional designers with more than two years of experience in user research and design, and two junior designers with less than two years of experience.

The questions of the interview were designed into three parts: (1) description of interpersonal dynamics, (2) effect of interpersonal dynamics, and (3) space to improve, in order to gradually understand the current interpersonal dynamics in usability testing.
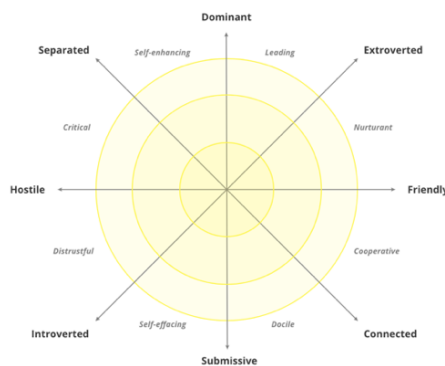

Figure 2. Simplified interpersonal circumplex

We tried to use the interpersonal circumplex model in interviews to help interviewees describe the interpersonal dynamic in an easier way and express their true experience more accurately. Since the concept of interpersonal dynamics is not familiar to people, they might not be able to accurately summarize it. We simplified the interpersonal circumplex [15] by deleting

specific points and descriptions in the circle. The simplified model (Fig. 2) only has four axes with descriptions of each direction and three concentric circles representing increasing degrees. This simplification made it easier for the designers to describe their experienced interpersonal dynamics using this model. In the beginning of the interview, the designers were asked to describe the interpersonal dynamics from their usability testing experiences in their own words. Because designers might have observed different interpersonal dynamics with different participants, we asked them to draw those different interpersonal dynamics as dots on the interpersonal circumflex, and to draw a range to represent the interpersonal dynamics they have experienced in general. This process resulted in qualitative insights expressed verbally by the interviewed designers, and a set of eight diagrams depicting a range of interpersonal dynamics they had experienced.

3.2. Observation

To triangulate the data from the previous step with an additional data source, we observed one of the interviewed designers during an actual usability test. We obtained consent from the designer and four participants to observe the entire testing process. The designer recorded the testing process with a camera. During the observation, we sat next to the designer and observed and took notes on our observations. Whenever designers or participants made noteworthy behaviors or their facial expressions showed some emotion during the testing, we immediately wrote down the scene and their behavioral expressions. After they completed the entire usability testing, we conducted brief interviews with the designer and the participants based on our observation notes to better understand their thoughts and moods at the time. Then we contacted the designer and got the corresponding screenshots of the video taken by the recording camera at that time to show those noteworthy moments.

4. Results

4.1. Results of the Interviews

Based on the interview structure, we divided the results of each interview into three main parts: (1) description of interpersonal dynamics, (2) effect of interpersonal dynamics, and (3) potentials for improving interpersonal dynamics. We listed the key insights with interviewees' quotes from each part. The

Table 2. Result of the interviews

Interviewees	Experience	Part I. Description	Part II. Effect	Part III. Improvement
No.1 Professional designer	3 years as UX designer	'' I feel like I was provide them with services, like a robot.''	''Sometimes I worry that they(participants) will be too scrupulous and afraid to express, so the results will be not objective or critical enough.''	''I think the open-ended stages such as Feedback stage needs more improvement. Sometimes participants are not that willing to talk and don't give too much feedback.''
No.2 Professional designer	5 years as UI/UX designer	''We don't communicate much, like teammates who complete their tasks separately.''	'' The participants usually want to be helpful, so they will be too cautious about each steps, afraid of giving us wrong results...they're not confident about their decisions. ''	'' Because we won't have too much interaction and communication during the test stage, the feedback is important. We can receive more useful feedback and inspiration from more active participants.''
No.3 Professional designer	3 years as UX designer	''I sometimes feel I'm the tutor...like leading the students do some activities.''	'' I usually put my position as a tutor,like leading the participants to do tasks...but I think this somehow gives the participants stress and they they will try to do the Right things. ''	'' The testing stage needs to be improved most, because I think our behaviors will have influence on their performance.''
No.4 Professional designer	4 years as UI/UX designer	'' I always feel I'm asking them for a favor,so I will try to be more friendly.''	''Sometimes I will worry that if atmosphere is too casual or relax, might make the participants feel that this is not formal. We want them to think that what they do is valuable so that they will take this test process seriously.''	'' I think the introduction stage needs to be improved most, cause this is the beginning to build our emotional connection and communication with participants. ''
No.5 Professional designer	2 years as UI/UX designer	'' Businesslike - I will try to show my professional side.''	''Some participants will tell/do what you want to hear/see, especially when you pay them for testing.''	''The preparation stage, the way we select the participants might affect the interpersonal dynamics. It always depends on people.''
No.6 Professional designer	4 years as UX designer	''Sometimes I feel that as an observer, I give them a sense of oppression, like I am invigilating an exam.''	''I think we sometimes prefer to see the outcome we want, and the participants can feel our preferences during the testing, so they will just lie to us.''	''The 'think-aloud'(testing) part needs to be improved, so that participants can express themselves more naturally and consciously, without being nervous because of our observations.''
No.7 Junior designer	1 years as UI/UX designer	'' I feel like we are asking for a favor from the participants, they are helping us. ''	''I am also nervous to communicate with participants, because I don't know what state and emotions I should use to communicate with them, so I always maintain a relatively distant communication.''	''I think the introduction stage needs to be improved, because that's when i feel awkward and don't know how to communicate with participants.''
No.8 Junior designer	1 years as UI/UX designer	''Most participants are very friendly and try to be helpful...sometimes we can even become friends after testing.''	''Once, a participant kept saying ' I'm sorry ' and showed anxiety. Under this kind of observed situation, some participants are afraid of making mistakes, so they will be nervous and unable to do the task as in daily situations.''	'' The introduction stage. I think a good begin is very improtant for the testing next. I think if I build a good connection with participants from the beginning, they can feel more relax in the testing.''

results of the interview with the eight designers are summarized in Table 2.

In previous interviews, we used a combination of inductive and deductive analysis approaches. We used the interpersonal circumplex model as a structure to guide the interviewees' responses, but we were also open to emerging new insights from the responses. Since each designer and different participants may have different interpersonal dynamics, the designers first described some memorable interpersonal relationships they have experienced and then tried to conclude the dynamic in general. The following are several common interpersonal dynamics described by designers:

Friendly: All designers had the experience of friendly interpersonal dynamics with participants. Some designers said they tend to be more friendly in the testing because they felt they are asking for help from participants, and if they build a more friendly atmosphere the participants will provide more feedback.

Dominant/Guiding: Four designers have felt that they were in their interactions with participants. One designer sometimes thinks she is just guiding participants to finish tasks like a teacher.

Connected: Three designers have felt they had a good connection with participants. One designer said if he can build a good emotional connection with a participant at the beginning, the participant will be more active and might provide more feedback.

Distance/Separate: Two designers expressed the experience of keeping distance to participants. One designer who also felt that he was like a 'tutor' in the interactions with participants, said that he tried to keep distance after he realized that he was too dominant in the test. He thinks that keeping distance is more professional because it can reduce their impact on participant choices.

In summary, the interpersonal dynamics most designers described are *'friendly'* and '*dominant'*. We found that the designers themselves are in fact also not sure how they should communicate with the participants, or what kind of relationship they want to build with participants: to be like professional testers and mentors, or rather being like asking for a favor.

Most designers claim that they are swinging between these dynamics and want to find a better balance between "friendly, connected" and "dominant, separated".

Mapping these insights from the interviews and the drawn interpersonal circumplexes, we can distinguish between "positive impact" and "negative impact" of interpersonal dynamics.

The positive impact is basically related to the communication atmosphere and can bring more insights or feedback:

- Designers think when they have a friendly interpersonal dynamic, participants are more likely to share more feedback so designers can get more inspiration from their communication.
- When the participants could feel more relaxed in testing, they can follow their own selection in the testing without worrying about making mistakes.

The negative impact is mainly that the participants' behavior and feedback in the testing might be unable to reflect their real thoughts:

- When the participants feel stressed, they will be too nervous in testing and afraid of being wrong or making mistakes in testing.
- In addition, when the designers provide too much guidance in the test, the participants won't be objective or critical enough.
- Designers also found that if in a too formal atmosphere, the participant may "pretend" or "lie" in

the test— doing what the designer wants to see or hear, because they think this is helping the designer complete their work.

By analyzing the current interpersonal dynamics and their impact, we can see that improving the interpersonal dynamics during usability testing is relevant:

Most designers are now swinging between "friendly/connected" and "separated/serious" dynamics and want to find a better balance between them.

Both interpersonal dynamics can have a positive and negative impact on the testing results: The "friendly and connected" interpersonal dynamics can positively affect participants to share more feedback. But it can also cause participants to be not objective or critical enough: do/say what the designer wants to see or hear. The "separated and serious" interpersonal dynamics can help participants remain objective in the testing. On the other hand, it can cause participants to be too nervous and afraid of making mistakes.

4.2. Results of the Observation

Drawing on insights from the observation of the four participants' usability testing, we were able to separate the usability testing process into five core stages: (1) preparation, (2) introduction, (3) testing, (4) feedback, and (5) post-stage. This procedure allowed us to better show the results on a timeline. Then we summarized the key interaction points in different stages and listed the designer's and participants' noteworthy behaviors and feedback corresponding to the stage at the
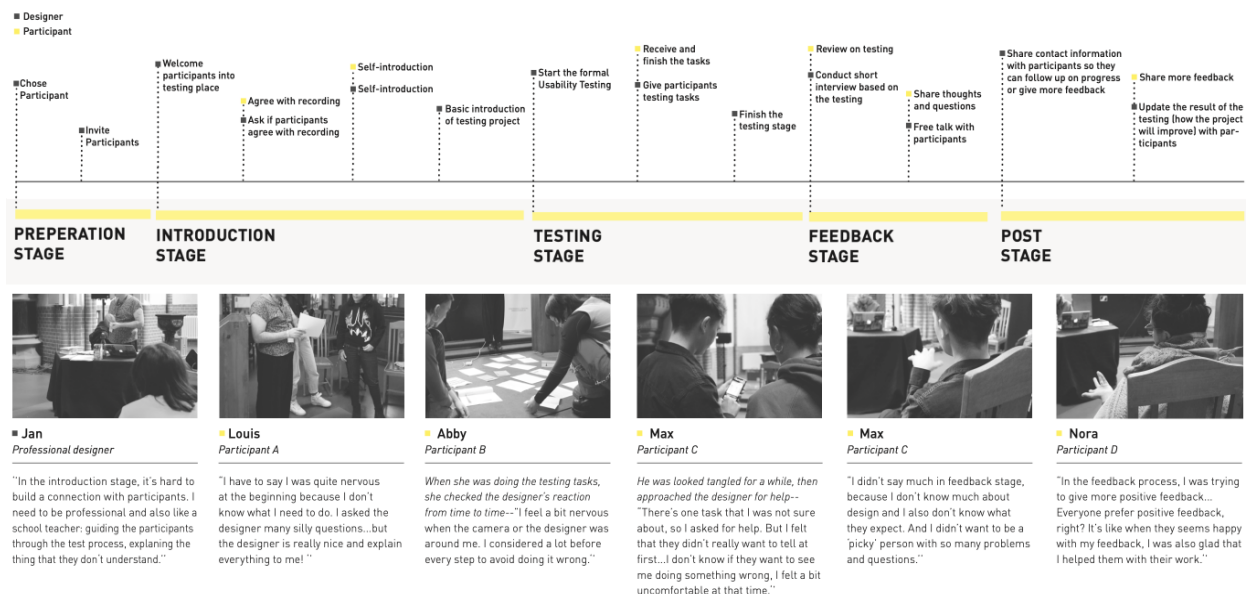
Figure 4. Results of the observation

bottom of the timeline. The results of the observation of the four participants are shown in Figure 4.

In the next steps, we use this 5-stage timeline to identify opportunity areas for introducing gamification elements. This process shall lead to the development of our "framework of opportunities for improving interpersonal dynamics during usability testing" which we will present in the next section.

5. Framework of Opportunities to Improve Interpersonal Dynamics

To answer our *RQ 2: How can we improve the interpersonal dynamics between designers and participants during the usability testing by gamification?*, we developed a framework to help designers improve the current interpersonal dynamics and find a balance between the "friendly, connected" and "separated, serious" interpersonal dynamics.

5.1. Mapping of Problem Space and Solution Space

We first identified opportunity areas and improvement space to be included in the framework. Then, based on the potential game types and elements table created in the literature review, we go through the insights from interviews and observations to match those game types and elements with different improvement goals.

In our observation result, we have divided the usability testing process into five core stages: preparation, introduction, testing, feedback, and post-stage. As we can see in Table 3, we list the potential improvement space based on the interview results. Then we use stars to indicate how many designers mentioned

potential improvement directions ("friendly, connected" or "separated, serious") in interpersonal dynamics at different stages.

Based on this table, we summarized six goals that designers want to achieve in these five stages to improve the interpersonal dynamic:

1. Help participants understand what's usability testing for
2. Remain neutral and do not interfere participants
3. Support participants to express their real thoughts
4. Establish good connections with participants
5. Have a comfortable communication atmosphere
6. Keep connection with participants to follow up more easily

Then we go through the goals and the insights we have from the previous research, to match the game elements or game types with those improvement goals.

For goal 1 we have found out that participants feel nervous and afraid to make mistakes because they think this testing is to evaluate their performance and they assume designers will prefer the positive result. To help them understand the goal of usability testing is to evaluate the product, not participants, the game types such as 'role-play' and 'quiz' can be useful to help participants understand the usability testing method in a more interesting and attractive way.

For goal 2 and 3 we have found out that some designers already realized they might subconsciously show their own presets for the results or give participants too much guidance. However, during the real testing, they may still show a preference for the participant's choice due to the subjective view of their own design. The game elements like 'Rules', 'Rewards', and 'Teams' can be useful to help create some activities with goals and positive reward to remind

Table 3. Opportunity areas

| Stages of usability testing	Potential improvement space	Improvement direction of interpersonal dynamics	
		Friendly & connected	Separated & serious
Preparation Stage	Some designers think how to select and reach out to the participants might affect the first impression and the interpersonal dynamics. It's also important to let participants understand the meaning of this testing.	★★	
Introduction Stage	Many designers said that the introduction stage is very important to build emotional connection and a good communication atmosphere with participants. Also can be useful to provide an explanation both on the topic and the method. And there is indeed more space for improvement at this stage.	★★★★★★	
Testing Stage	In this stage, there will be different forms depending on the project. Most of the designers realized that it's important to support participants express their real thoughts, but they need to be reminded sometimes to remain neutral and not interfere with participants.	★	★★★★
Feedback Stage	Designers want to have a good communication atmosphere with participants in this stage, but it is difficult because they lack communication and interaction in the testing stage. Besides, they also need to be neutral in the conversation.	★★★	★★
Post Stage	In this stage, designers said if they can keep their connection with participants, the participants will be more willing to follow up on the project or give more feedback.	★★★	
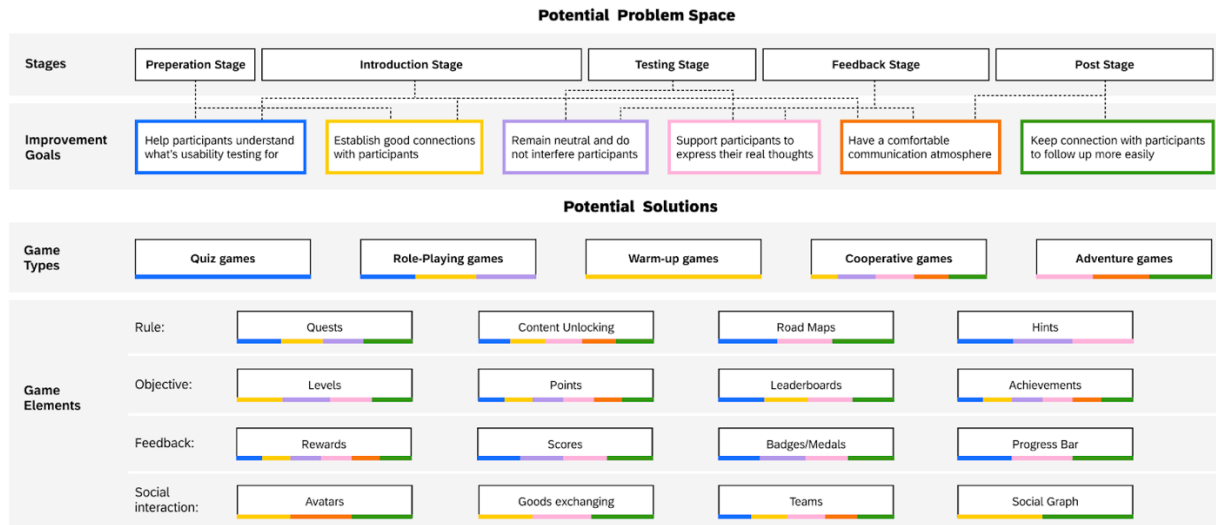
Figure 5. The framework of opportunities for improving interpersonal dynamics during usability testing

designers to be neutral regardless of whether results meet their presets.

For goal 4, 5, and 6 we learnt that if designers and participants feel comfortable and have emotional connections with each other, they might perform better. Game types like 'warm-up games', role-playing' and 'adventures' and game elements like 'teams', 'progress bar', and 'avatar' are more helpful in creating a fun and comfortable atmosphere.

5.2. The Framework

Based on our previous analysis of improvement goals and potential solutions (game types and game elements), we suggest a framework that designers can use to create their own gamified activities to improve interpersonal dynamics in usability testing.

The framework we suggest is shown in Figure 5. There are two main categories in the framework: **'Potential Problem Space'** and **'Potential Solutions'**.

The selection of the first category 'potential problem space' is divided into two rows, which can determine the general direction of the result.

First row is **'Stage':** according to the testing environment or the stage that the designer thinks need to be improved, they can select one or more stages.

 Second row is **'Improvement goal':** one or more improvement goals can be selected according to the testing stages and the actual situation.

The second category is 'potential solutions', which is divided into two big rows. The offered choices all have a highlight area with different colors. According to the color of the goal, designers can find the choice with

the same color highlight as a suggestion for creating a potential solution to this goal.

First row is **'Game type':** one or more types can be selected according to time and budget.

Second big row is **'Game elements'** with four types of game elements: rule, objective, feedback, and social interaction. It is recommended to select at least one game element in each row.

The way to use this framework is to select appropriate options from top to bottom in each row. First, select 'potential problem space', and then select 'potential solutions' based on the selected improvement goals with the color code suggestions. Then, combine the selected game elements in the selected game type to create a gamified activity. Consider your individual situation (such as budget, time limit, etc.) to arrive at an appropriate solution for your usability testing goals.

5.3. Exemplary Design Solution

To show how to use this framework in real scenarios, we tried to create one exemplary design solution in usability testing.

Figure 6. Exemplary solution space identified with the developed framework (black highlight)

We first selected the potential problem space from the framework, then we selected the potential solutions. The result of our selection process is shown in Figure 6.

Based on the combination of the game elements and game types, we designed a prototype with the role-exchange cooperative game setting.

Figure 7. Exemplary design solution

As illustrated in Figure 7, the developed game consists of task cards and rewards (in the form of rubber stamps) within a box. In the introduction stage, designers and participants need to follow the instructions in the box to finish a role exchange game: participants and designers exchange their roles and simulate the simple version of usability testing with some classic design interface to win the rewards together. In this activity, they need to cooperate and can see the usability testing method from the respective other's perspective. Participants and designers will receive rewards no matter what result they get in this testing, that way participants can understand there's no right or wrong in the testing and be encouraged to express their real thoughts in the formal testing. In the post-stage, they will use the rewards (stamps) they won in the role-exchange game to co-create an eco-bag for the participant as a gift. The stamps depict some usability-related symbols and cartoons to create a fun and memorable souvenir.

6. Discussion

6.1. Contribution

The work presented in this paper explores interpersonal dynamics between participants and researchers (in this case designers) during usability testing. Based on the interviews and observations, we explored possible solutions to improve current dynamics by gamification. To be specific, we suggested a framework that will help researchers identify opportunities to improve interpersonal dynamics to receive more objective results in usability testing.

6.2. Feedback and Iteration

To obtain designers' opinions on this framework, we shared it along with an exemplary design solution with five designers we had interviewed before. Most of them think the framework has a lot of potentials and there is space for further development.

All five designers expressed very positive feedback. They appreciate how this framework can adapt to different test scenarios. They also like how it provides designers with possible solutions while leaving room for further design. One designer said that the exemplary design solution we developed using the framework is interesting. She found this example very attractive and would like to apply it to her own future usability tests to get the optimal solution.

At the same time, some designers also pointed out some aspects that we can further develop in the framework. For example, the game elements and game types we currently provide are not rich enough, they wonder if there are more game elements that can be used. In addition, the framework lacks definition, explanations, and instructions for each game element or game type.

6.3. Limitations

The work presented in this paper has some limitations. Because our research and interviews began at the beginning of the Covid-19 pandemic, people just started to work remotely. Therefore, most of the designers we interviewed had only an in-person usability testing experience. Hence, our research is only based on in-person usability testing. Research on remote usability testing, which is increasingly common, was not part of our study. The next step of this research will need to focus on interpersonal dynamics in remote usability testing. Only then we can improve and develop our framework further.

Another limitation lies in the fact that we only validated the framework based on one exemplary solution. Additional configurations that the framework allows have not yet been explored. Further research is needed to investigate these configurations with different participants and in different contexts.

6.4. Future Work

Based on the positive feedback we received, we believe that this framework has the potential to be of use for other researchers and designers. However, for future development, there are some aspects we think could be improved.

We envision an interactive framework (for example, a web page), that can suggest 'potential solutions' according to the selection results in 'potential problem space'. The digital application would allow clicking on each game element/type to view its detailed definition and usage suggestions. Optional game elements and types could also be continuously added and updated. Moreover, based on some related literature regarding well-established methods like focus-groups [21]. We think we can optimize our solution and try to integrate it with these established social psychology research methods in our future work.

7. Conclusion

In this paper we introduce a novel framework outlining the problem space of interpersonal dynamics during usability testing and a possible solution space including several gamification elements. The framework was developed based on an extensive literature review and qualitative interviews and observations of actual usability tests.

Our findings suggest that interpersonal dynamics have an impact on the testing results: participants in a friendly atmosphere will give more insights and feedback, while participants in a nervous or tense atmosphere may not perform as in real life or may not conceal their real thoughts in the interview or feedback stage. Hence, improving the interpersonal dynamics in a usability test may contribute to better research results and better-designed products and services.

By mapping the potential problem space with potential gamification element solutions, we provide a framework of opportunities to improve interpersonal dynamics during usability testing.

8. References

[1] B. Carol, Usability Testing Essentials. 1st Edition. Morgan Kaufmann, 2010.

[2] N. Ahmad, M. W. Boota, and A. H. Masoom, 'Smart Phone Application Evaluation with Usability Testing Approach', Journal of Software Engineering and Applications, vol. 07, no. 12, 2014, doi: 10.4236/jsea.2014.712092.

[3] A. Hinchliffe and W. K. Mummery, 'Applying usability testing techniques to improve a health promotion website', Health Promot J Aust, vol. 19, no. 1, pp. 29–35, 2008, doi: 10.1071/he08029.

[4] H. Wilke, P. Badke-Schaub, and K. Thoring, 'The Healthcare Design Dilemma: Perils of a Technology-Driven Design Process for Medical Products', Proceedings of the DESIGN Conference, vol. 1, pp. 2217–2226, May 2020, doi: 10.1017/dsd.2020.133.

[5] M. Alshamari and P. Mayhew, 'Examining the role of tasks in usability testing for dynamic Websites', In *IHCI 2009 Conference in Rome, November.* 2009.

[6] N. Ranade, 'Conditional usability testing for UX optimization', presented at the SIGDOC 2019 - Proceedings of the 37th ACM International Conference on the Design of Communication, 2019. doi: 10.1145/3328020.3353906.

[7] E. Rügenhagen and T. Held, 'Game design techniques in user research methods - A new way to reach the high score in development teams', Lecture Notes in Computer Science, vol. 8518, pp. 754–762, 2014, doi: 10.1007/978-3-319-07626-3_71.

[8] A. Woodcock, S. Fielden, and R. Bartlett, 'The user testing toolset: A decision support system to aid the evaluation of assistive technology products', Work, vol. 41, no. SUPPL.1, pp. 1381–1386, 2012, doi: 10.3233/WOR-2012-0328-1381.

[9] P. Kara, 'Talking with Users in a Usability Test', Nielsen Norman Group, Jan. 26, 2014. Accessed: Jul. 05, 2020. [Online]. Available: www.nngroup.com/articles/talking-to-users/

[10] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, 'From game design elements to gamefulness: Defining "gamification"', Proceedings of the 15th international academic MindTrek conference, 2011, pp. 9–15. doi: 10.1145/2181037.2181040.

[11] K. Huotari and J. Hamari, 'Defining gamification - A service marketing perspective', Proceeding of the 16th international academic MindTrek conference, 2012, pp. 17–22. doi: 10.1145/2393132.2393137.

[12] M. H. Erdelyi, Psychoanalysis: Freud's cognitive psychology. New York, NY: W H Freeman/Times Books/ Henry Holt & Co, 1985, pp. xv, 303.

[13] S. Salvatore and W. Tschacher, 'Time dependency of psychotherapeutic exchanges: The contribution of the Theory of Dynamic Systems in analyzing process', Frontiers in Psychology, vol. 3, 2012, doi: 10.3389/fpsyg.2012.00253.

[14] Pincus, A. L., Sadler, P., Woody, E., Roche, M. J., Thomas, K. M., & Wright, A. G. C., Multimethod assessment of interpersonal dynamics. In C. J. Hopwood & R. F. Bornstein (Eds.), *Multimethod clinical assessment* (pp. 51–91). The Guilford Press, 2014.

[15] C. R. Brewin, Cognitive Foundations of Clinical Psychology (Psychology Revivals). Psychology Press, 2013.

[16] N. Vegt, V. Visch, H. de Ridder, and A. Vermeeren, 'Designing gamification to guide competitive and cooperative behavior in teamwork', in Gamification in Education and Business, 2015, pp. 513–533. doi: 10.1007/978-3-319-10208-5_26.

[17] M. Aparicio and C. Costa, Gamification: Software Usage Ecology. The Online Journal of Science and Technology, 8(1), 2018, pp. 92-100.

[18] J. G. Cover, The creation of narrative in tabletop role-playing games. Jefferson, N.C. : McFarland & Co., 2010.

[19] A. Rollings and E. Adams, Andrew Rollings and Ernest Adams on Game Design. New Riders Games, 2003.

[20] G. Guillen, J. Hamari, and J. Quist, Gamification of Sustainable Consumption: a systematic literature review. Proceedings of the 54th Hawaii International Conference on System Sciences, 2021, p. 1345. doi: 10.24251/HICSS.2021.163.

[21] M. C. Tremblay, A. R. Hevner, and D. J. Berndt, "Focus Groups for Artifact Refinement and Evaluation in Design Research," CAIS, vol. 26, no. 27, pp. 599–618, 2010, doi: 10.17705/1CAIS.02627.