When more is less: Designing and Testing the Usability of a Gamified Survey to Capture Relationship Data

Jhon Bueno-Vesga, University of Missouri
Joe Griffin, University of Missouri

Abstract: *Relevate Sign Up* is a gamified patient intake survey designed to collect demographic and relationship data from players to customize relationship research dissemination. In this paper, we report the details of the design choices and usability testing of the game. We also discuss lessons learned.

Introduction

Surveys are still one of the preferred ways to capture data for social sciences research and marketing purposes (Oliveira & Paula, 2020). Government entities, private companies, and universities used to invest a considerable amount of resources to collect data using surveys. More recently, the Internet has made delivering surveys to potential respondents easy and cost-effective. However, online surveys are usually tedious. Guin et al. (2012) identified four factors making respondents experience boredom and fatigue when answering surveys: the length of the survey, the effort required to answer it, the stress associated with the questions, and the number of surveys they are asked to answer. These factors may trigger respondent behaviors like speeding, random responding, premature termination, and lack of attention (Harms et al., 2014). These negative effects of conventional surveys have pushed researchers to devise creative ways to improve user experience and increase levels of engagement in their participants when filling surveys. One way to improve user experience is to add game design elements to traditional systems in an approach called "gamification" (Triantoro et al., 2019). Gamification has been defined as "the use of design elements characteristic for games in non-game contexts" (Deterding et al., 2011, p.10). Research shows that game features can make filling out a questionnaire a more enjoyable experience (Harms et al., 2014). Gamified surveys have also shown benefits like better user experience and increased motivation, resulting in higher engagement, more feedback, and better data quality (Cechanowicz et al., 2013; Dolnicar et al., 2013; Guin et al., 2012). Even in cases showing no improvements in terms of engagement, gamified surveys were seen as more attractive and easier to answer than the traditional ones (Oliveira & Paula, 2020). Recent comparative studies have also found that gamified surveys are also more enjoyable than traditional ones without compromising the integrity of the data collected (Triantoro et al., 2019).

Relationship Data Collection

One of relationship researchers' primary interests lies within the study of the complexity of interpersonal relationships. Relationship data allows targeting specific populations with specific needs. For instance, someone who is single might not be interested in information about interacting with romantic partners, or adopted children may have different needs than children that were raised by their biological parents. It is possible that if relationship researchers can elicit better quality data about relationships, they could better address the specific needs of their patients more effectively.

Although gamified surveys might be a solution for relationship researchers, the use of gamified surveys for relationship data collection has received little attention and application in family science research. Since family science has started an era of evaluation and innovation (Hamon & Smith, 2014), this study becomes a significant contribution to achieving that goal. In this paper, we describe the design and usability testing of a preliminary prototype of a gamified patient intake survey designed to capture relationship data by incorporating game elements (e.g. a fun story and dialogues with a fictional character) into an online survey solution. We also describe lessons learned in the process to be applied in future iterations of the game and for the benefit of the research community. This is the first iteration

of a design-based research where a solution to the problem is put to use to test how well it works (Brown, 1992). The solution may be adapted and re-tested to gather more data in subsequent iterations. This iteration included two stages: Expert feedback and game design choices (Stage 1) and a usability test (Stage 2).

Related Work

Literature shows that gamified surveys have been used in several knowledge domains spanning from human-computer interaction (Zagel et al., 2018), health (Wimmer et al., 2018), and academic research (Barwick et al., 2018; Dolnicar et al., 2013; Harms et al., 2015), to business with special emphasis on market research (Bailey et al., 2015; Dorcec et al., 2019; Guin et al., 2012). Gamification has also been applied to surveys intended to collect public opinion (Aubert & Lienert, 2019). However, to our knowledge, there are no applications in the field of family science.

We found two studies particularly relevant due to similarities with the characteristics of our survey and the results we want to achieve. The first study examines a gamified survey designed to explore how much children consider the law as an empowering force in their lives (Barwick et al., 2018). The authors designed a tablet-based game with scenarios common to children and asked law-related questions about specific incidents shown using animations. They used a "child alien" as a key character in the game to ask questions to the player. The game proved to be an effective way to collect data from children by keeping them engaged. As found in Barwick et al. (2018), we believe that using a fictional character to establish a fun dialogue with the survey respondents could enhance their levels of engagement and break the usual boredom associated with long surveys. The second study we found particularly relevant detailed a gamified survey for a context in which collecting data would be exceptionally difficult (Dorcec et al., 2019). The results of the UX analysis suggested that the use of gamification is a good approach for situations where real-world data are difficult to acquire. It also proved the gamified survey superior to the classic questionnaire in terms of attractiveness, stimulation, and novelty. We also found that collecting relationship data could be a difficult task in real life. Relationship data is very private, intimate, and sometimes embarrassing. As found in Dorcec et al. (2019), we believe that using a gamified survey can help us collect data by using a neutral, fictional, and fun surveyor that allows the participant to feel more relaxed and engaged. Additionally, using a richly designed gamified interface can bring superior attractiveness and novelty to the players.

Stage 1 – Expert Feedback and Game Design Choices

The survey was designed to collect relationship and demographic data from participants so they can receive more personalized disseminated research information tailored to their specific needs. The researchers defined a set of relationships questions categorized into four different relationship areas: a) questions about social relationships (6 questions), b) questions about parenting and family relationships (20 questions), c) questions for single people regarding romantic relationships (7 questions), and d) questions for people who are currently in a romantic relationship (11 questions). They also included questions to capture participants' detailed demographic data (27 questions). After deciding to use gamification as a strategy to increase levels of engagement in the participants when answering questions about their relationships, we decided to use a small number of game elements in an approach called "light gamification" (Bailey et al., 2015). Three game elements were included: story, a pedagogical agent, and input mechanics.

The researchers decided that the survey was going to be conducted by a fictional character that would engage the participant in a conversation taking place in a world of fantasy, similar to what happens in role-playing games. They conceptualized three ideas for the fictional character to be chosen: 1) A

newborn toddler created with AI who knows nothing about people or their relationships. The player is expected to train the AI toddler by feeding it with his/her personal relationship information. The AI toddler responds by asking follow-up questions, often wrong and misguided while it learns from the player. 2) An alien that has been sent to research humans. The alien interviews the player using a hypothetical technical manual that is unfortunately off base and needs frequent correction. The alien is confused due to its lack of knowledge about the human race. The alien uses sarcasm with a dry sense of humor to respond to the player's information. 3) A broken mechanical fortune teller tries to guess the fortune of the player. To get the predictions right, he asks vague and leading questions forcing the player to correct him so he can "get it right" and pretend he knew it all along. After a round of expert reviews, the subject matter experts felt the alien surveyor would be the best main character for the story, just as in Barwick et al. (2018), relationship researchers chose the alien surveyor as the main character of the story. It was named "SmiggleFord" and contextualized as shown in Figure 1. SmiggleFord was written to be slightly sarcastic. The following is an example of the type of dialogue it uses: *"This is researcher SmiggleFord. I've got … let me look at my holopad here… Human #3092309812 of… sighs...of batch 23935-10934851-45715491467. … yawn… you see that pod race on the holotube last parsec? Flibblesnap got his tentacle caught again and… oh… this thing is on?!"*

Figure 1
"SmiggleFord", Alien Character Conducting the Survey

The researchers explored the possibility of using an innovative mechanism for users to answer multiple-choice questions that would make it easy to choose options in a touch screen device (smartphone, tablet, or tactile screens). The team prototyped a sliding mechanic using the picture of the alien in two (left and right) to four (left, right, up, and down) directions to choose one out of 2, 3, or 4 options depending on the question (See Figure 2). The currently chosen direction is highlighted in green to show the player his/her potential choice. If one or more of the options (directions) is disabled it is shown highlighted in red. The game would also display a "*Next*" button to go to the next question, an "*Exit*" button to exit the survey at any time after finishing one category of questions, and an "*End*" button to leave the game when all questions are answered.

Since questions are grouped into four main categories (Family, Social, Single, Romantic); after the demographic questions, the player must choose one of the remaining categories to start answering questions from that chosen category. A screenshot of the whole interface displaying the category selector is shown in Figure 3. Once the player finishes answering all questions from one category, they come back to the same screen. The direction corresponding to the finished category is then highlighted in red signaling that it is no longer selectable. When all questions are answered, the button "*Next*" is changed to "End" to finish the survey and exit.

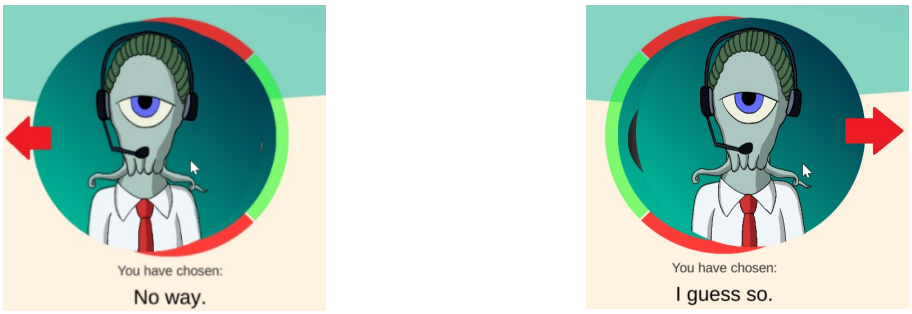

Figure 2
Alien Being Dragged to the Left (Left Image) and Right (Right Image) to Choose One of Two Options

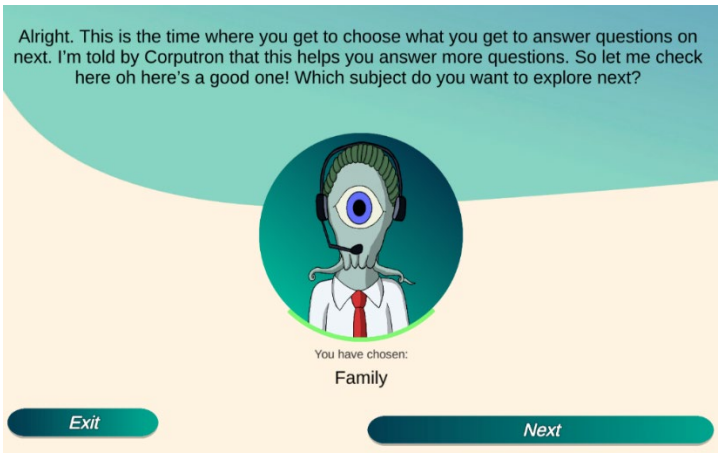

Figure 3
Category Selector Screen

Stage 2 – Usability Testing Methods
This study included a remote usability test via Zoom. The researchers chose concurrent think-aloud because this method allows getting an insight into the participants' thinking process. The study also included a system usability survey – SUS (Brooke, 1996), and a short individual follow-up-interview.

Participants
Hwang and Salvendy (2010) estimated that the minimum number of test users needed to reach an 80% discovery rate when using a think-aloud testing method was 9. We invited university students to join the study via email. Ten participants expressed their willingness to join the study. They all were in the range of 24 to 50 years of age. All of them used a Google Chrome browser. All of them expressed having professional internet experience. All of them were graduate students. Three of them said they belong to programs of the Human Development and Family Science department.

Test Protocol and Instruments
Participants were asked to test the game while expressing aloud their thoughts about their experience (think-aloud). Users' behavior while interacting with the game was observed. Observations were performed and recorded via Zoom. Participants were asked to share their screens and to turn their cameras on. After a brief explanation of the purpose of the game and the expectations related to the test, participants were asked to open the game in their browser and start their gameplay session. Their behavior was observed to identify functionality in need of improvement. The test was completed when

they finished answering all the questions in the gamified survey without help. After the interaction, participants completed the System Usability Survey (Brooke, 1996) and answered a few open-ended interview questions related to their perception of the experience.

Stage 2 – Usability Testing Findings
All ten participants finished the test successfully. The average time to complete the test was 26 minutes.

Observation
The most prevalent issues in all observations were:

- Most participants (8/10) struggled to understand the directions provided about the new mechanism to answer multiple-choice questions or could not figure it out. Most of them (4/10) needed help to understand there were more than two options in every multiple-choice question and the way to find them (by dragging the character up and down). Some of them (2/10) finished the test without exploring all the answer options for some multiple-choice questions.
- Half of them (5/10) did not find an appropriate option to answer at least one of the questions. Thus, they had to choose any answer just to make the software let them continue.
- Some of them (3/10) expressed how they enjoyed interacting with the alien surveyor. Conversely, some of them (2/10) showed discomfort with the alien. They expressed that its language was hard to understand, and its conversation irrelevant, distracting, or offensive.
- Some of them (3/10) didn't realize that they had to scroll down to have access to other answer options in selection list questions.
- Some of them (2/10) found the text animation distracting and time-consuming. They said it was inappropriate given the amount and extent of the texts. Mac users (3/10) could not skip the text animation with the space bar or right-click.
- None of them lost their focus on the game or prematurely quit the survey. However, two of them showed speeding or random responses.

System Usability Scale - SUS
The average score obtained using the System Usability Scale - SUS was 41/100. A SUS score below 68 is considered below average. The low average score obtained shows the concern of the users about the effectiveness, efficiency, and/or their satisfaction with the game (Brooke, 1996).

Interview
The interview confirmed some of the observations. This is a summary of the most prevalent topics in the interview:

Q1: Tell me what you liked about the experience:

- Most of them (7/10) found the alien character cute, playful, and/or its conversation fun. In contrast, some others (2/10) found it too sarcastic or too talkative.
- Some of them (2/10) said they liked nothing about the experience.

Q2: Tell me what you disliked about the experience:

- Some of them (4/10) said they disliked the mechanism to answer multiple-choice questions.
- Some of them (4/10) said they disliked the text animation.
- Some of them (3/10) disliked scrolling down to see more options in checkbox list questions.
- Some of them (2/10) disliked the alien character or its narrative.
- One player disliked the limited options of multiple-choice questions.
- One player pointed out that using green and red would be challenging for color-blind people.

Q3: Was there anything confusing about the experience?

- Half of them (5/10) said that making choices by dragging the alien was initially confusing.
- Some of them (3/10) found the conversation with the alien confusing.
- Some of them (2/10) were confused because some of the questions in the survey were not consistent with previous answers provided.

Q4: If you had a magic wand, what would you change about the experience?

- Some of them (5/10) would start the experience by stating the purpose of the game.
- Most of them (4/10) expressed the need to eliminate the text animation.
- Some of them (2/10) said they would change the alien character.

Discussion

In this study, we designed and implemented a gamified survey to collect relationship data from participants. We performed usability testing to detect improvement areas for the next iteration.

Our results showed that most of the participants liked the alien character conducting the interview, although three of them expressed very negative feelings about the character. None of them exited the interview prematurely, got distracted, or skipped any of the question sets in the game although the exit button was always available after a set was finished. This indicates good levels of engagement considering that the average time of the interaction with the game was 26 minutes. Observational data also showed levels of enjoyment with some of the participants even laughing out loud during the dialogue or uttering expressions like *"I like SmiggleFord, it's so sassy"* or *"It's so funny"*.

The test also surfaced serious usability problems. The first and possibly highest impact issue was the use of a novel mechanic to choose options in multiple-choice questions. The problem was not with the mechanic itself, but with the fact that it was unknown and not as intuitive for the participants as expected. Although one of Nielsen´s heuristics (Nielsen, 1994) indicates that interfaces should follow industry standards; when innovative mechanics are used, this can be solved by providing an illustrative tutorial at beginning of the experience. The mechanic used in our gamified survey app is very simple and could be easily exemplified with an animation demonstrating its use. The second major problem was the use of animation to display text on every screen. The use of text animations is very pervasive in games. Although game players had the option to skip the animation and/or immediately display all the text, many of them expressed their discomfort with the animation. The idea of using a text animation to incrementally display questions and dialogues with the alien surveyor seemed initially attractive but, given the large number of questions of the survey and the extensive character length of some of them, the animation became too time-consuming and annoying to some of the participants.

Conclusion

The idea of a gamifying a relationships data collection survey showed good results in terms of engagement and high likeability of the fictional character performing as a surveyor. But, the use of innovative but novel methods for multiple-choice questions and text display, as well as the use of scrolling in checkbox list questions among others, caused usability problems in the game. We plan to iterate and correct these issues based on the findings of this study. We will continue this research with future rounds of expert reviews and usability testing before releasing the game to a public audience.

References

Agnew, C. R., & VanderDrift, L. E. (2009). Survey Methods in Relationship Research. In H. Reis & S. Sprecher, *Encyclopedia of Human Relationships*. SAGE Publications, Inc.

Aubert, A. H., & Lienert, J. (2019). Gamified online survey to elicit citizens' preferences and enhance learning for environmental decisions. *Environmental Modelling & Software*, 111, 1–12.

Bailey, P., Pritchard, G., & Kernohan, H. (2015). Gamification in Market Research: Increasing Enjoyment, Participant Engagement and Richness of Data, but what of Data Validity? *International Journal of Market Research*, 57(1), 17–28. https://doi.org/10.2501/IJMR-2015-003

Barwick, J., Watkins, D., Kirk, E., & Law, E. (2018). Adventures with Lex: The gamification of research? Convergence, 24(3), 229–250. https://doi.org/10.1177/1354856516677682

Brock, R. L., & Laifer, L. M. (2020). Family Science in the Context of the COVID-19 Pandemic: Solutions and New Directions. *Family Process*, 59(3), 1007–1017. https://doi.org/10.1111/famp.12582

Brooke, J. (1996). *SUS: A "Quick and Dirty" Usability Scale*. In Usability Evaluation In Industry. CRC Press.

Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *Journal of the Learning Sciences*, 2(2), 141–178.

Cechanowicz, J., Gutwin, C., Brownell, B., & Goodfellow, L. (2013). Effects of gamification on participation and data quality in a real-world market research domain. *Proceedings of the First International Conference on Gameful Design, Research, and Applications*, 58–65.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification." *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. https://doi.org/10.1145/2181037.2181040

Dolnicar, S., Grün, B., & Yanamandram, V. (2013). Dynamic, Interactive Survey Questions Can Increase Survey Data Quality. *Journal of Travel & Tourism Marketing*, 30(7), 690–699.

Dorcec, L., Pevec, D., Vdovic, H., Babic, J., & Podobnik, V. (2019). How do people value electric vehicle charging service? A gamified survey approach. *Journal of Cleaner Production*, 210, 887–897.

Guin, T. D.-L., Baker, R., Mechling, J., & Ruyle, E. (2012). Myths and Realities of Respondent Engagement in Online Surveys. *International Journal of Market Research*, 54(5), 613–633.

Hamon, R. R., & Smith, S. R. (2014). The Discipline of Family Science and the Continuing Need for Innovation. *Family Relations*, 63(3), 309–322. https://doi.org/10.1111/fare.12068

Harms, J., Wimmer, C., Kappel, K., & Grechenig, T. (2014). Gamification of online surveys: Conceptual foundations and a design process based on the MDA framework. *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, 565–568.

Harms, J., Seitz, D., Wimmer, C., Kappel, K., & Grechenig, T. (2015). Low-Cost Gamification of Online Surveys: Improving the User Experience through Achievement Badges. *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, 109–113.

Hwang, W., & Salvendy, G. (2010). Number of people required for usability evaluation: The 10±2 rule. *Communications of the ACM*, 53(5), 130–133. https://doi.org/10.1145/1735223.1735255

Nielsen, J. (1994). *10 Usability Heuristics for User Interface Design*. Nielsen Norman Group. Retrieved January 26, 2022, from https://www.nngroup.com/articles/ten-usability-heuristics/

Oliveira, K. W. R., & Paula, M. M. V. (2020). Gamification of Online Surveys: A Systematic Mapping. *IEEE Transactions on Games*, 1–1. https://doi.org/10.1109/TG.2020.3004366

Triantoro, T., Gopal, R., Benbunan-Fich, R., & Lang, G. (2019). Would you like to play? A comparison of a gamified survey with a traditional online survey method. *International Journal of Information Management*, 49, 242–252. https://doi.org/10.1016/j.ijinfomgt.2019.06.001

Wimmer, C., Biegler, S., Harms, J., Kappel, K., & Grechenig, T. (2018). SQT: A tool for the automated measurement of respondent behaviour and response quality in health-related gamified online surveys. *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)*, 1–8. https://doi.org/10.1109/SeGAH.2018.8401377

Zagel, C., Piazza, A., Petrov, Y., & Bodendorf, F. (2018). SciencOmat: A Gamified Research Platform for Evaluating Visual Attractiveness. In L. E. Freund & W. Cellary (Eds.), *Advances in The Human Side of Service Engineering* (pp. 50–60). Springer International Publishing.