Not So Fun? The Challenges of Applying Gamification to Smartphone Measurement

Michael W. Link[1], Jennie Lai[2], and Kelly Bristol[3]

[1] Nielsen, 3784 Ardsley Ct, Marietta, GA, USA, 30062
[2] Independent, 105 Duane St, New York, NY, USA, 10007
[3] Nielsen, 501 Brooker Creek Blvd, Oldsmar, FL, USA 34677
{michael.link,Kelly.Bristol}@nielsen.com, Lai.Jennie@gmail.com

Abstract. Gamification and engagement techniques (points, status, virtual badges, and social-sharing) are applied to a mobile and on-line data collection tool to determine if these approaches can improve respondent compliance with a requested task: recording their television viewing over the course of several weeks by increasing their engagement with the app. In a series of tests, we demonstrate that virtual badges appear to be a salient and positively viewed technique for app engagement among teens and younger adults. However, not all of these approaches have positive impact especially with older adults and, in the end, do not improve compliance with the primary task.

Keywords: smartphones, mobile apps, gamification, motivation, user engagement.

1 Introduction

"Gamification" can be broadly viewed as the concept of applying game-design thinking to non-game applications to make them more engaging and improve motivation for targeted behaviors. For measurement researchers, "gamification" may be viewed through a more focused lens as the process of applying the psychological and sociological drivers of game play to behavioral and attitudinal measurement tasks in order to improve respondent engagement and compliance[1]. Gamification has proliferated in recent years throughout the marketing industry, and by extension, to market research -- less so in more traditional survey research markets. But while both disciplines seek to promote engagement for their respective needs, marketing uses gamification to keep consumers returning to a product or service (e.g., airline loyalty programs), while opinion and behavioral research is exploring its use to engage panelists to respond and comply fully with an information-gathering task. The collection of valid and reliable information from the public requires that respondents be motivated to both participate in and comply fully with the tasks involved in the measurement process. Traditionally, respondents have been offered monetary incentives. While perhaps still necessary to initially capture the attention of prospective respondents, monetary incentives do little to engage them in the longer-term. This is where game-based approaches may offer a new and innovative way of motivating participants.

There are two key concepts at work in gamification [2]. First, "game mechanics," which refers to the actions, tactics, or mechanisms used to create an engaging and compelling experience for respondents. This can include use of points, levels, challenges, leaderboards, virtual goods, gifts/charity, etc. to drive the desired behavior for engagement with and ultimately achievement of a given task. For example, virtual badges have been shown to have a range of utilities, from reinforcing goal setting, increasing compliance with instructions, denoting reputation/status/affirmation, and fostering group identification [3]. In contrast, "game dynamics" refers to the motivations that are tapped into as a result of the gamified experience, thereby driving continued participation by the respondent. These include motivational elements such as achievement, self-expression, competition and altruism. The choice of gaming tactics is an important one. Researchers need to understand the motivations they are trying to trigger so that they can utilize the appropriate game mechanics in turn. In doing so, however, researchers need to take care not to change or influence the attitudes, behaviors, or phenomena they are trying to measure.

Although gamification techniques tied to monetary or tangible rewards (such as point systems for panel participants to earn cash or goods over time or lottery drawings for survey participants), little research has looked at the use of gamification techniques to tap into intrinsic motivations -- that is, activating game dynamics without the use of monetary or tangible rewards. One such study applied a "gamified" approach to an online survey, utilizing a respondent avatar operating in a virtual world with the goal of answering survey questions to move the action forward [4]. While many of the participants reported that the approach was engaging and fun, it was also viewed as being more difficult to navigate, took longer, and produced greater survey break-offs than did the more traditional online survey design.

Here we examine a somewhat different set of game mechanics (points, status, virtual badges) and engagement techniques (social sharing) in order to test their effectiveness in motivating and engaging respondents to utilize a data collection mobile app over different timeframes.

2   Methodology

A succession of three experiments was conducted between January 2012 and May 2013, examining the impact of various gamification and engagement techniques on respondents' use of a smartphone/online-based activity diary. The features tested includes points for completing various tasks; "status" upgrades based on the number of points earned; virtual badges for both task completion and longevity in the measurement panel; and ability to post activities to interact with others in a newsfeed, on Facebook, and/or via Twitter. The three studies consisted of an initial convenience sample of "friends and family" using a split-sample design to test basic concepts, followed by two probability samples of the general public (see Table 1 for details).

User engagement was assessed using a number of different techniques across the studies, including (1) analysis of the paradata generated within the app by the users behavior and interaction with the app; (2) structured debrief web survey; and, (3) indepth interviews conducted either via telephone or in-person.

3   Key Findings

The app use studies were conducted in sequence, with learnings from preceding tests used to enhance subsequent versions.   We consider the key findings from each in turn.

Table 1. App Version, Study Design, Game Features, & Evaluation Methods

App Ver.	Respondent Population	Sample Frame / Recruit Mode (n)	Field Period	App Interfaces	Features Tested	Evaluation Methods
V1	Friends & Family (one per home) w Internet or mobile device[1]	Employee listing / Email recruitment (n = 250)	6 weeks (1/17/12-2/27/12)	• iOS app	• Points • Levels • Badges • Social Sharing (Internal Newsfeed)	• App Usage Behavior • Follow-up Survey • Follow-up Telephone Debriefs
V2	Adults aged 18+ (one per home) w Internet or mobile device[1] in Birmingham, AL, USA & Little Rock, AK, USA	Landline & Cellphone Listings / Telephone Recruitment (n = 150)	2 weeks (8/2/12 - 8/15/12)	• Android app • iOS app • Web	• Badges • Social Sharing (Facebook Newsfeed)	• App Usage Behavior • Follow-up Survey
V3	Adults aged 15+ (all in home) with Internet or mobile device[1] in Birmingham, AL, USA & Dallas, TX, USA	Home address Listing / Mail & Telephone (n = 464)	2 weeks (5/9/13-5/24/13)	• Android App • iOS app • Web	• Badges • Social Sharing (Facebook, Twitter)	• App Usage Behavior • Follow-up Survey • In-person Interviews

3.1   Version 1

The initial study utilized only an iOS-based smartphone app with a convenience sample of "friends and family" of Nielsen employees across the United States and was conducted over a 6 week time period [5]. A split sample design was used in which half of the respondents were provided with a full-feature app on the first day, which allowed users to earn points & advance "levels" by completing specific activities such as accessing the app or responding to push notifications; and earn virtual badges within the app for specific viewing activities such as entering their first 'live' viewing or retrospectively recalling entries they did not report in the moment (they were not told, however, explicitly how to earn badges -- this was to keep a sense of "mystery" and encourage the users to explore the app features.), and an internal newsfeed for sharing their viewing with other app users. We call this the "Full App" group. The second group started the study with an app that had no gamification features. They used this version for the first two weeks, and then were sent an app upgrade which activated the game mechanics (points, status, & badges). Two weeks later, they received a second upgrade which activated the social sharing newsfeed. We refer to this as the "Sequential App" group.

Looking at the number of television viewing hours recorded on average by week for the two groups, three key findings emerge. First is a general decline in participation across the 6 week study by all participants -- with a greater decline between weeks 1 and 2 and a more gradual fall-off from weeks 2 through 6. Based on data from Nielsen's ongoing TV ratings services, we would expect the 6 week trend to be relatively flat with some minor variation. The conclusion, therefore, is that gamification and engagement elements did not appear to drive long-term compliance with the primary data collection task.

Second, the number of hours recorded for the Full App and Sequential App groups was nearly identical for weeks 1 and 2 (week 1: 10.3 hours Full App, 10.2 hours Sequential App; Week 2: 7.3 hours for both). This indicates that neither the game nor social features appears to have had an effect in engaging respondents to record their viewing. In other words, it made no difference whether respondents did (Full App) or did not (Sequential App) have these features in their version of the app (though the age group of these users was skewed older than the general U.S. population given the convenience sample of Nielsen employees).

Third, and perhaps most important, when the game features were introduced to the Sequential App group at the start of week 3, their average viewing per week increased (7.3 hours to 8.0 hours), while those in the Full App group declined slightly (7.3 hours to 7.2 hours). During the subsequent weeks (3 thru 6) both groups saw viewing hours decline, but the Sequential App group drop-off was less -- especially once the social sharing elements were added at the start of week 5. By week 6, those with the Full app recorded, on average, 4.8 hours of viewing per week compared to 6.4 hours for the Sequential App group. The takeaway is that by introducing game and social mechanics sequentially, the user experience was "freshened" over time compared to those who received these features at the start and as a result appears to have slowed the rate of decline in compliance with the viewing entry task.

Virtual badges were viewed by respondents as the most liked game component, with younger respondents (aged 18-29, 54%; aged 30-39, 48%) being more likely to say in a follow-up debrief survey that they loved or liked the feature, compared to older adults (aged 40+, 28%). In terms of race/ethnicity, positive reactions (loved or liked) to virtual badges were higher among African American respondents (83%) and Asian Americans (73%), than among Hispanics (54%), or non-Hispanic Whites (39%). Moreover, a majority (60%) of Asian Americans said that the virtual badges were very or somewhat encouraging their continued participation in the study. When asked about their experience in using the app during the debriefing interviews, these users wanted to learn how to earn points or badges and opportunity to earn them over the course of the study period otherwise these features would be viewed as "pointless." More importantly, they preferred to be rewarded for watching more TV or a specific program rather than their reporting behavior (they cited other apps that rewarded them for watching specific programs as intended for marketing purposes). This approach, however, would have a clear biasing effect by influencing the behavior we were trying to measure. One of the potential complexities of using gamification approaches is, therefore, how to apply them effectively to secondary tasks such that there is respondent motivation to comply with the primary task as well -- yet not change the actual behavior or attitudes being measured.

In contrast to the more positive views of some of the gamification elements, the social sharing newsfeed was utilized by very few respondents: approximately 20% accessed the feature and only 3% posted any viewing content to the feed. In debrief interviews, respondents remarked: "I want to invite friends and family so I can share viewing and compete for badges, points, etc." and "I don't want to share what I'm watching with people I don't know."

3.2   Version 2

Given the lack of apparent efficacy of the points and status approach in Version 1, we dropped these elements and focused solely on virtual badges as the gamification technique in Version 2 of the app [6]. In particular, using virtual badges to "reward" individuals for engaging with the app rather than the amount of viewing they recorded.. Respondents could earn badges by 1) completing instructional tasks such as completing the app tutorial or accessing the 'info' button customized for specific app activities; and 2) exploring various components of the app such as the home screen or checking their viewing entries (in hopes of jogging their memory of entering any viewing they may have forgotten). Taking into consideration the feedback from the first version, respondents were provided instructions on how to earn the badges once they accessed the   "virtual trophy case" within the app and also shown the number of badges they could potentially earned over the course of the two-week study period.

In terms of social-sharing, we eliminated the internal app-user only newsfeed and instead allowed respondents to post a generic message to Facebook about their televisions viewing (e.g., "Bob Smith is watching Comic book Heroes"). This approach was an attempt to allow individuals to share their activity with people in their own social network while minimizing any potential biasing effects this might have on their

viewing behavior by not allowing those network friends to comment within the app -- any comments were kept with the individual respondent's Facebook account.

Looking at the virtual badges earned during this field period, all respondents received the badge for registering the app ("Rock the Registration") as a way of introducing them to the badge concept. The next two highest earned badges were for accessing all four major modules within the app ("Hide & Seek," 65% of users) and completing their first (of four) brief, five-question survey posed during the field period ("Reporter," 64%). Both of these related to first-time, high-level exploratory behaviors. The lowest earned badges involved checking their viewing entries for accuracy ("Editor-in-Chief, 21%) and completing 3 of the 4 brief surveys posed during the two week collection ("Correspondent," 13%). These reflected either more detailed types of behaviors or repeated activities. It appears, therefore, that a majority of respondents did partake in the initial, high-level exploration of the app and its features, yet only a much smaller percentage had deeper, longer-term interactions with many of the app features.

Age differences in acceptance of the badge feature were clear in the surveys conducted after the first badge was earned and after the last badge was earned. All (100%) of the respondents aged 18 - 35 said they were "excited" about receiving the badge in both surveys. Among those 36 and older, 87% said they were excited after the first badge, however, this number dropped to 72% when asked about the last badge -- so interest was more on the wane among the older respondents.

In terms of social-sharing, the results were similar to those seen in Version 1 with very few respondents accessing this feature and far fewer still pushing content to their Facebook social feed.

3.3   Version 3

The final version of the television viewing app continued the focus on virtual badges, but with an eye towards encouraging younger respondents to engage with the app from start to end of a two-week study period (with emphasis on those aged 15 to 24, the demographic which is often the most difficult to get to comply with data collection tasks) -- the group seen in the first two tests as most likely to be engaged by this approach. A total of 15 badges were designed focusing on obtaining greater compliance with high-value activities such as registering the app at the start of the study, completing instructional tasks and reporting their viewing regularly or completing custom surveys during the study in hopes of sustaining their motivation throughout the two-week period.

Social-sharing options were also included in this version given the increasing popularity and growing number of social media sites. Like version2, respondents could post (but not receive within the app) a message about their viewing activity on Facebook. Version 3 also included a similar option to post to Twitter and allowed respondents to tailor their messages rather than a generic message from the first two versions.

Assessing the impact of the gamification features, younger age groups (particularly those aged 15-24) found the virtual badges to be of greater value and interest than did older adults. Just under 95% of this younger group indicated that the virtual badges

helped them to understand how the app worked, compared to 60% of those aged 50 and older. Likewise, 82% of respondents aged 15 to 24 years indicated that they consciously tried to earn badges, versus just 34% of adults aged 50 or older. Younger respondents also said that the badges motivated them to enter television viewing more regularly -- the critical task for the study: 82% versus 26%. A majority across all age groups felt that earning badges was not difficult.. This finding aligns with in-person debrief results that found younger age groups were very engaged in the badges; however, most of other age groups did not see badges as a very useful engagement feature.

In terms of social-sharing, 68% of users did not think the social sharing feature (Facebook or Twitter) were important to them, while only 3% of respondent thought that they were very important. About 20% of users did not even have FB/Twitter account. These findings were fairly consistent across age groups. The figures indicated that Facebook and Twitter were considered less important in the TV viewing as the motivation to participate in the study (they preferred not to share what programs they watch or how much TV they watch with their social network online).

Based on the in-person debriefing interviews conducted with 15 participating households in Birmingham and Dallas, younger respondents aged 15-18 in multi-generational homes were far more engaged with the badge feature than the adults (i.e., parents or grandparents). The minors across these homes consistently reported they deliberately tried to earn all the badges in the app and were "frustrated" when they couldn't earn them all. However, some also admitted while they liked earning the badges but that did not necessarily motivate them to report their viewing consistently throughout the study period. In fact, the parents played a key role in reminding them to enter what they watched in the app. While this version of the app achieved the objective of using badges as an instructional tool by helping the respondents learn the features of the app but it did not seem to influence their reporting behavior.

4   Conclusions

Despite the finding that younger participants reported liking the virtual badges and found them useful for instructional purposes, there was no corresponding improvement in television viewing recording for this group compared to others age groups (looking simply at the regularity of their app entries, not the amount of viewing). In this case, therefore, respondent perception and behavior were not apparently aligned and as a result, the gamification features did not have the desired effect of improving consistency in reporting. Given the evidence across the three studies, therefore, the final recommendations were to either remove all gamification features or to add monetary rewards when badges are earned (e.g., incremental incentive) -- a recognition that any intrinsic motivation generated by the use of virtual badges was insufficient in and of itself to motivate respondents to comply fully, over time with the primary data collection task.

Based on these studies we can draw the following initial conclusions about the use of gamification techniques to improve respondent engagement:

Techniques Work Differently across Populations: People differ in the types of game mechanics and dynamics that motivate them. In particular younger adults, Blacks, Asians and Hispanics responded more favorably to some of the gaming techniques (in particular virtual badges) than did older adults and Whites. No single approach, therefore, will work equally well across different subgroups of the population.

Gamification Appears to Work Better with a Longer-Term Panel Survey than a One-Time Survey: Game mechanics appear to have their optimal impact in terms of potentially changing the pace or "freshening" a longer-term measurement experience. Although the techniques could be employed for a stand-alone survey, it is doubtful that the time and cost of developing these techniques in such an instance would be worthwhile.

It is Important to Measure the Desired Outcome Behaviors not Just Assess Respondent Engagement of Gamification Features: In all three studies, when asked, younger adults and teens tended to give certain gamification mechanics (in particular virtual badges) high marks. They even stated that they felt motivated by receipt of the virtual badges. This did not, however, translate into more consistent compliance with the primary task they were asked to undertake.

Techniques Should Motivate Respondents to Achieve the Measurement Tasks, but not to Drive or Change the Behavior or Attitudes that are being Measured: This is an obvious, but critical insight. While marketers can utilize gamification techniques to directly drive consumer purchase behavior, researchers need to motivate respondents to complete the measurement tasks without inadvertently encouraging or changing what it is they are trying to measure.

Gamification is a powerful motivational tool in marketing. While it may one day have a similar impact in attitude and behavioral measurement, the goals and mechanics need to be thought through carefully to ensure that respondents are motivated to comply with the measurement tasks and are not driven simply to win the "game." There are promising results for engagement, particularly with traditionally hard-to-engage groups (i.e., younger adults, racial/ethnic minorities); however, there are also challenges to effectively implementing these techniques to influence compliance with primary (rather than simply secondary) data collection tasks and do so without jeopardizing the critical information being gathered.

References

1. Link, M.W., Donato, P.: Engaging Consumer Panelists Through Gamification. Marketing News by the Advertising Marketing Association, http://www.marketingpower.com/ResourceLibrary/MarketingNews/Pages/2013/2-13/consumer-panels-purchasing-behaviors-gamification-behavior-motivations-consumer-measurement.aspx (February 2013)
2. Lai, J., Link, M.W.: Emerging Techniques for Respondent Engagement: Leveraging Gamification Techniques for Mobile Application Research. CASRO Journal 2012-13 (2013), http://bit.ly/123gmpr

3. Antin, J., Churchill, E.: Badges in Social Media: A Social Psychological Perspective. Paper presented at 29th CHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada (2011)
4. Downes-Le Guin, T., Baker, R., Mechling, J., Ruyle, E.: Myths and Realities of Respondent Engagement in Online Surveys. International Journal of Marketing Research 54, 613–633 (2012)
5. Link, M.W., Lai, J., Vanno, L.: Smartphone Applications: The Next (and Most Important?) Evolution in Data Collection. Paper presented at the 67th Annual Conference of the American Association for Public Opinion Research, Orlando, FL, May 17-20 (2012)
6. Lai, J., Bristol, K., Link, M.W.: Unlocking Virtual Badges as the Key for Respondent Engagement in Mobile App Surveys. Paper presented at the Annual Meeting of the Midwest Association for Public Opinion Research, Chicago, IL (2012)