ONLINE APPENDIX

Personalized Game Design for Improved User Retention and Monetization in Freemium Mobile Games

A1 Experimental design – Further details about the degree of difficulty

Table A1 shows the difficulty adjustment intensity for each group of players, depending on the number of rounds that they have played in the previous 7 days. If a user played 20 or more rounds, the difficulty remained intact, at the default level denoted by 5. If a user played between 15 and 19 rounds, the difficultly was set to 4, which implies that the probability of connectivity in the board increased by 10% (i.e., the game was made easier for that player). As the user played fewer rounds in the last 7 days, the chances of connecting pieces were made easier.

# rounds in last 7 days	Difficulty level	Increase of "connectivity"
>= 20	5 (more difficult)	0%
< 20	4	10%
< 15	3	20%
< 10	2	30%
< 5	1 (least difficult)	45%

Table A1: Difficulty adjustment intensity by level of play during the past 7 days.
Difficulty 5 is the default option in the game design.

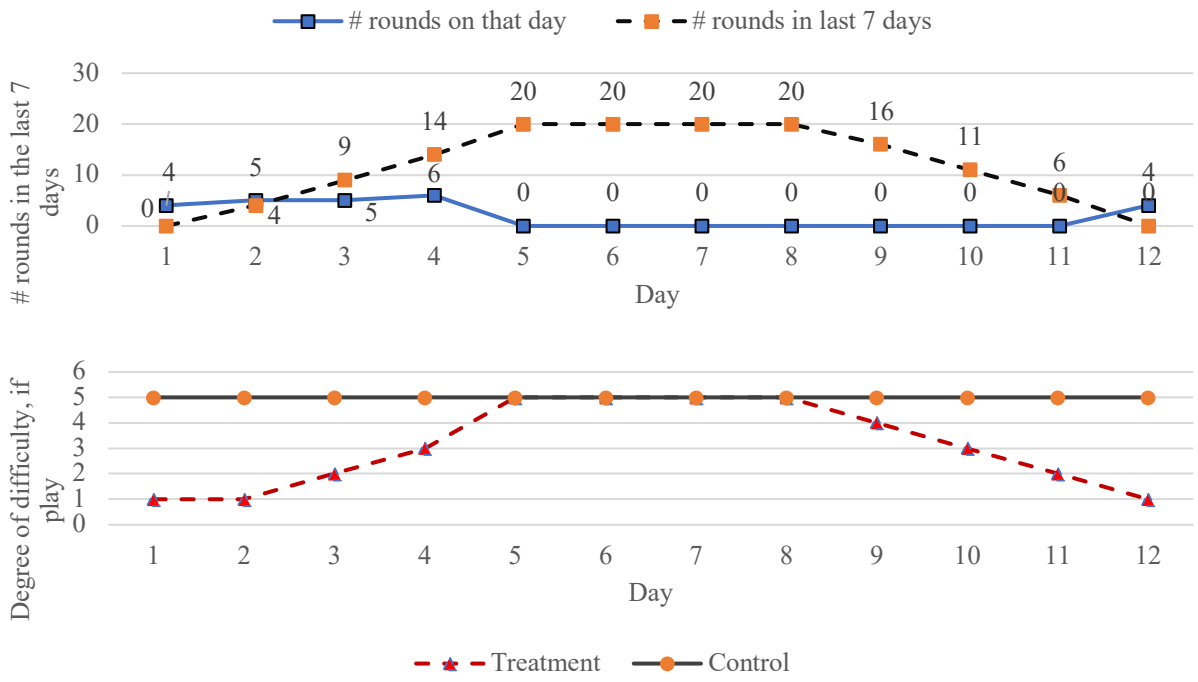

Figure A1: Hypothetical example to illustrate the level of difficulty (bottom figure) that correspond to different levels of past play. *The red line (dashed) corresponds to treated users, whose difficulty was altered based on the amount of play in the previous 7 days, and the back line (solid) corresponds to control users whose level of difficulty never changed.*

Figure A2 shows two difficulty scenarios. On the left the player is facing the highest level of difficulty ("default" in the game design) whereas the player on the right is facing the easiest scenario with low difficulty levels.

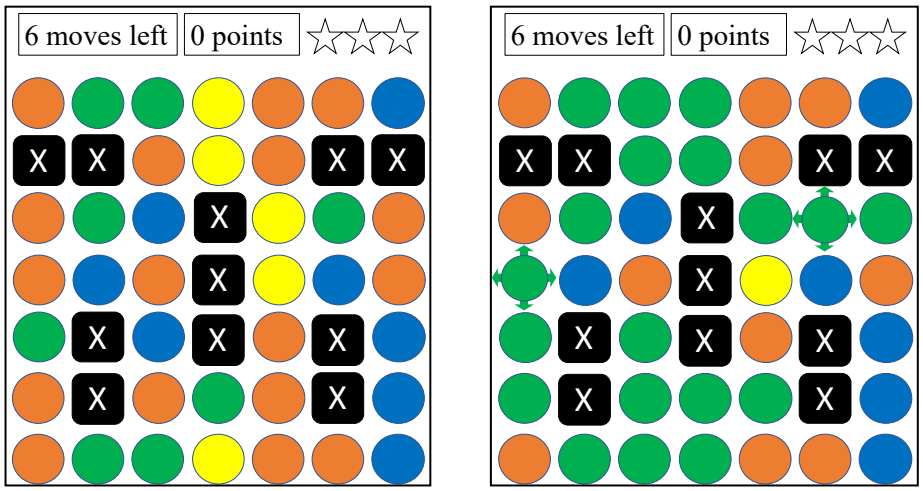

Figure A2: Example of difficulty manipulation. *Image of a level of the game, with two different levels of difficulty. On the left, the most difficult scenario which corresponds with the default difficulty of the game. On the right, the player is facing an easier scenario in which there are much higher chances of creating long combinations (of green color) and has two "special chips," identified by the arrows around the two pieces.*

A2 Data – Descriptive statistics and randomization checks

Table A2 shows the descriptive statistics and randomization checks for the full set of user-level variables. Level20 variables are measured when users pass level 20 (and remains constant over time) and all other variables are measured at the moment of the intervention.

	N=	329,999					190,863	139,136		
		Mean	SD	p25	p50	p75	Control	Treatment	Difference	p-value
level20_rounds		45.83	36.91	29	36	49	45.85	45.81	-0.04	0.810
level20_days		8.691	15.74	1	3	9	8.67	8.719	0.049	0.373
level20_stars		39.80	4.497	37	39	42	39.81	39.78	-0.03	0.110
level20_coins_collected		40.75	75.93	0	0	70	40.82	40.64	-0.18	0.499
level20_did_use_extra		0.926	0.261	1	1	1	0.927	0.926	-0.001	0.668
level20_did_use_coin		0.513	0.5	0	1	1	0.513	0.514	0.001	0.343
level20_did_purchase		0.011	0.103	0	0	0	0.011	0.011	0.000	0.080
age_rounds		189.6	217.7	62	113	225	190.0	189.2	-0.8	0.315
age_days		45.85	31.91	19	37	66	45.78	45.94	0.16	0.168
age_level		37.74	16.97	24	39	40	37.78	37.69	-0.09	0.101
age_stars		75.19	36.68	49	67	87	75.3	75.04	-0.26	0.045
age_coins		24.0	74.0	0	0	30	24.04	23.96	-0.08	0.752
age_distance_to_gate		6.353	6.561	0	4	11	6.484	6.174	-0.31	0.000
rfm_rec		13.88	17.9	3	7	16	13.85	13.92	0.07	0.276
rfm_week		5.728	6.68	0	2	12	5.735	5.718	-0.017	0.476
rfm_ratio		3.719	2.952	1.667	3.077	4.836	3.721	3.716	-0.005	0.605
stuck_rounds		26.72	62.45	2	7	25	26.79	26.61	-0.18	0.403
stuck_days		17.25	21.59	4	9	21	17.22	17.3	0.08	0.347
stuck_playdays		3.054	3.804	1	2	4	3.054	3.055	0.001	0.930
stuck_days_in_gate		23.47	23.38	8	15	32	23.32	23.67	0.35	0.033
stuck_broke_gate		0.0218	0.146	0	0	0	0.0217	0.022	0.0003	0.645
yesterday_progress		1.225	2.564	0	0	2	1.223	1.227	0.004	0.703
yesterday_win_prop		0.308	0.341	0	0.2	0.545	0.309	0.307	-0.002	0.289
skill_rounds_per_level		4.41	3.653	2.227	3.286	5.318	4.413	4.406	-0.007	0.568
skill_stars_per_level		1.988	0.257	1.821	1.957	2.149	1.988	1.987	-0.001	0.036

Table A2: Descriptive statistics (left-most columns) and randomization checks (last 4 columns). *We acknowledge that four variables (out of 25) show significant differences across conditions. Given the experimental set up, we attribute these differences to chance, and not to any systematic intrinsic difference between treatment and control users.*

A3 Manipulation checks by difficulty adjustment intensity

Table A3 shows the remaining metrics examined in the manipulation checks.

	Points (by level of difficulty)				Snake Length (by level of difficulty)			
	4	3	2	1	4	3	2	1
Treatment	**1275**	**3225**	**6126**	**10741**	**0.018**	**0.027**	**0.112**	**0.269**
	(128.1)	(137.9)	(157.4)	(91.2)	(0.009)	(0.009)	(0.009)	(0.005)
Constant	35818	35862	35648	35047	5.119	5.148	5.144	5.179
	(82.7)	(87.6)	(95.8)	(50.8)	(0.006)	(0.006)	(0.007)	(0.003)
N	370137	302836	281827	1055166	370137	302836	281827	1055166
R-squared	0.1%	0.6%	2.0%	4.9%	0.0%	0.0%	0.2%	1.1%

Table A3: Manipulation checks by degree of difficulty. *OLS of the round outcome against a treatment dummy using all rounds on the first day of the experiment. Standard errors are clustered at the user level. Treatment variable in bold indicates p-value<0.05.*

A4 Robustness checks for the main analyses

Day of intervention

One potential concern with the main results presented in **Table 6** is that the intervention does not occur simultaneously for each treated player. **Figure A3** shows the number of users assigned to each condition over time.

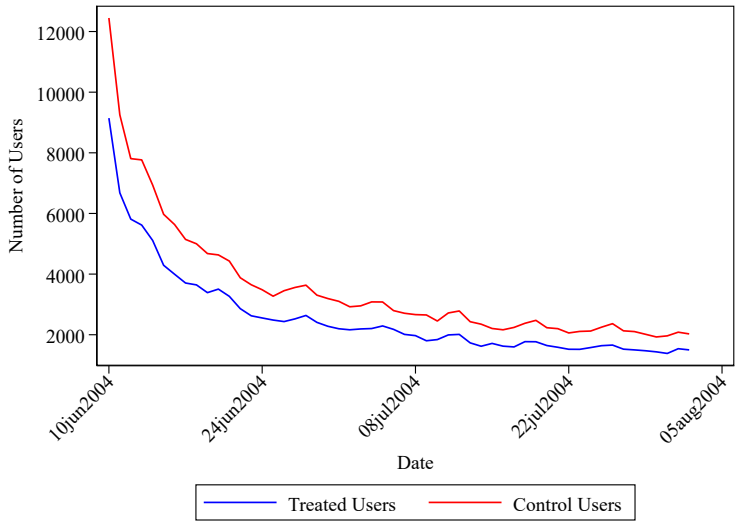

Figure A3: Number of users, by condition, qualifying for the experiment in each day.

Although time-varying factors (e.g., July 4th, weekday vs. weekend) are not confounds due to our experimental design—for every treated user, we have control users who would have been treated on the exact same day but were allocated to control instead—these factors might introduce some variation that could potentially reduce the efficiency of our ITT estimates. To address this, we re-run the main regressions, including time-related variables as controls to capture the potential differences arising from the non-uniform treatment timing. Specifically, we added 'day of the week', 'month', and 'July 4th' as controls. In **Table A4**, we replicate the results presented in Table 6 of the main manuscript, now controlling for variables that capture the day on which users received the treatment. This includes 'day of the week', 'month of the year', and 'July 4th', all incorporated as dummy variables. All results are consistent with those obtained in the main analysis.

	# Rounds played	Progress made	Engagement	Retention 1	Retention 7	Retention 14
Treatment	**1.247**	**0.746**	**0.0212**	**0.027**	**0.0244**	**0.0196**
	(0.0251)	(0.00882)	(0.000729)	(0.00164)	(0.0017)	(0.00156)
p-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Constant	5.256	0.65	0.812	0.279	0.632	0.721
	(0.0326)	(0.0108)	(0.00109)	(0.0023)	(0.00244)	(0.00225)
# obs	329,999	329,999	1,867,849	326,472	316,257	308,003

Table A4: Robustness check including day-level controls. *OLS of the behavior of interest against the treatment variable and 'day of intervention' dummies. Standard errors are clustered at the user level. Treatment variable in bold indicates p-value<0.05.*

Maximum level achieved prior to the intervention

It is possible that users who have previously achieved higher levels in the game respond differently to the DDA intervention compared to newer players at lower levels (recall that in order to be eligible for treatment (or control) users had to pass at least level 20). If this is the case, the results in the main manuscript might reflect a subset of users rather than the entire user base. In **Table A5**, we replicate the results from **Table 6**, incorporating the interaction between

treatment and the maximum level achieved by the user prior to the experiment (the variable 'Max Level' has been standardized). The treatment effect results remain consistent with those in the main manuscript. The main effect of 'Max Level' aligns with expectations: more experienced users play more rounds, achieve less progress, and exhibit higher engagement and retention levels. All interaction effects are positive, indicating that more experienced users (i.e., those who had reached higher levels before the intervention) react more strongly to the difficulty adjustment. Importantly, the magnitude of the interaction effect is much smaller than the treatment effect, confirming that lowering difficulty positively impacts user outcomes regardless of their previous maximum level.[13]

	# Rounds played	Progress made	Engagement	Retention 1	Retention 7	Retention 14
Treatment	**1.263**	**0.748**	**0.0211**	**0.0277**	**0.0251**	**0.0202**
	(0.0253)	(0.00877)	(0.000729)	(0.00164)	(0.0017)	(0.00156)
p-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Max Level	**0.136**	**-0.288**	**0.0059**	**0.04**	**0.0465**	**0.0363**
	(0.0146)	(0.00384)	(0.000487)	(0.00114)	(0.00106)	(0.000944)
p-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Treatment * Max Level	**0.332**	**0.093**	**0.00329**	**0.0102**	**0.00506**	**0.0035**
	(0.0272)	(0.00945)	(0.000699)	(0.00178)	(0.00162)	(0.00142)
p-val	0.0000	0.0000	0.0000	0.0000	0.0017	0.0139
Constant	5.57	0.773	0.818	0.297	0.644	0.744
	(0.0132)	(0.00375)	(0.000501)	(0.00105)	(0.00111)	(0.00103)
# obs	329,999	329,999	1,867,849	326,472	316,257	308,003

Table A5: Robustness check including max-level achieved prior to intervention. *OLS of the behavior of interest against the treatment variable and (standardized) 'maximum level achieved' prior to the intervention. Standard errors are clustered at the user level Standard errors are clustered at the user level. Treatment variable in bold indicates p-value<0.05.*

[13] Recall that all users in our sample had achieved level 20 prior to the intervention.

Degree of difficulty adjustment

The DDA condition sets a difficulty level based on the number of rounds a user played in the last 7 days. Consequently, the level of difficulty adjustment varies among users, both on the first day of the intervention and in subsequent periods. It is expected that the treatment effect is stronger in groups where the difficulty adjustment was more significant and weaker where the adjustment was minimal. Although the degree of difficulty adjustment is an endogenous metric determined by the number of rounds in previous days (which is likely influenced by earlier adjustments), we can utilize this metric on the first day of the experiment since prior consumption could not have been affected by the intervention. We conduct this analysis in two ways: First, we create a standardized continuous variable named 'Intensity,' which increases as the level of difficulty adjustment grows (i.e., when users are allocated to difficulty=1) on the first day of the intervention. We replicate the results in **Table 6**, adding both 'Intensity' and its interaction with 'Treatment' (**Table A6**). All results are as expected: Average treatment effects are consistent with those presented in **Table 6**, the main effect of intensity of difficulty adjustment is negative for all outcomes of interest, consistent with the notion that lower usage in the last 7 days predicts lower usage in the future, and the interaction terms are all positive, corroborating that stronger difficulty adjustment leads to stronger treatment effects. See results in **Table A6.**

	# Rounds played	Progress made	Engagement	Retention 1	Retention 7	Retention 14
Treatment	**1.219**	**0.725**	**0.0217**	**0.0269**	**0.0246**	**0.0202**
	(0.0249)	(0.00868)	(0.000732)	(0.00162)	(0.00161)	(0.00149)
p-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Intensity	**-0.578**	**-0.0925**	**-0.0104**	**-0.106**	**-0.15**	**-0.123**
	(0.0137)	(0.00409)	(0.000487)	(0.00109)	(0.000985)	(0.000892)
p-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Treatment * Intensity	**0.515**	**0.361**	**0.0101**	**0.00924**	**0.0136**	**0.0112**
	(0.0245)	(0.00823)	(0.00072)	(0.0017)	(0.00151)	(0.00135)
p-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Constant	5.597	0.79	0.818	0.301	0.65	0.747
	(0.0132)	(0.0039)	(0.000505)	(0.00103)	(0.00105)	(0.000982)
# obs	329,999	329,999	1,867,849	326,472	316,257	308,003

Table A6: Robustness check including difficulty adjustment intensity on the day of the intervention. *OLS of the behavior of interest against the treatment variable and difficulty adjustment 'intensity' (standardized) on the first day of the intervention. Standard errors are clustered at the user level Standard errors are clustered at the user level. Treatment variable in bold indicates p-value<0.05.*

Second, we treat difficulty adjustment as a categorical variable and replicate the results of the main manuscript separately for each level of difficulty adjustment cohort (see **Table A7** for levels of adjustment). As expected, and consistent with the manipulation analyses (**Table 5**), the effect of the intervention is stronger when the difficulty adjustment is larger.

	Intensity of Difficulty Adjustment			
	Largest (Difficulty = 1)	Large (Difficulty = 2)	Small (Difficulty = 3)	Smallest (Difficulty = 4)
# Rounds played	**1.639**	**1.279**	**0.781**	**0.334**
	(0.0352)	(0.0674)	(0.0619)	(0.0554)
p-val	0.0000	0.0000	0.0000	0.0000
Progress made	**1.041**	**0.659**	**0.381**	**0.154**
	(0.0132)	(0.0215)	(0.0196)	(0.0169)
p-val	0.0000	0.0000	0.0000	0.0000
Engagement	**0.0307**	**0.0204**	**0.0113**	**0.00582**
	(0.00104)	(0.0019)	(0.00181)	(0.00164)
p-val	0.0000	0.0000	0.0000	0.0004
Retention 1	**0.0333**	**0.0353**	**0.0181**	**0.0095**
	(0.00199)	(0.00459)	(0.00463)	(0.00431)
p-val	0.0000	0.0000	0.0001	0.0276
Retention 7	**0.0367**	**0.0258**	**0.00513**	0.00579
	(0.00242)	(0.00417)	(0.00368)	(0.00302)
p-val	0.0000	0.0000	0.1640	0.0558
Retention 14	**0.031**	**0.016**	0.00531	**0.00552**
	(0.00236)	(0.00354)	(0.00303)	(0.00243)
	0.0000	0.0000	0.0802	0.0232

Table A7: Impact of difficulty adjustment by cohorts, determined by levels of difficulty adjustment on the first day of the intervention. *OLS of the behavior of interest against the treatment variable by group of intensity of difficulty adjustment on the first day of the intervention. This table presents the parameter estimates of the variable of interest ('Treatment'). Standard errors are clustered at the user level. Treatment variable in bold indicates p-value<0.05.*

User persistence in the game

A potential limitation of the round-level analysis (as seen in **Table 7** and **Table 8**) is that the users who persist in playing for extended durations—and hence feature more prominently in this analysis—might exhibit systematic differences between the treatment and control groups. This phenomenon represents a survival bias, where users who continued gameplay after being treated with a reduced difficulty level during the initial round (or subsequent few rounds) may differ inherently from those in the control group who persevered despite facing more challenging game settings. To address this concern, we conducted two robustness checks. Firstly, we replicated the

round-level analyses considering only the initial five rounds post-treatment for each player.[14] The rationale for choosing the first five rounds is that it offers a conservative approximation of the number of rounds most users would likely play, irrespective of their treatment assignment. Given that all users start each day with a set of five lives, even those who encounter the most challenging difficulty levels can engage in up to five rounds without having to pause or purchase additional extras.

Secondly, we replicated all ITT analyses while controlling for all user attributes outlined in Section 3.4. It is important to note that these attributes capture historical user competencies in gameplay, activity levels, and prior tendencies to utilize extras, coins, and monetary resources. A detailed breakdown of the outcomes from these robustness checks can be found in Appendix A4. In brief, barring the exceptions of coins and money expended on extras within the inaugural five rounds (where notable differences were not identified), all other results align with those delineated in **Table 7** and **Table 8**.

	Continue playing	Use extras	Use coins	Use Money	Coin extra	Coin gate	Money extra	Money gate
Treatment	**0.0145**	**0.0012**	**0.0004**	**0.0002**	-0.00026	**0.00235**	0.00002	**0.00066**
	(0.0008)	(0.0004)	(0.0002)	(0.0001)	(0.00014)	(0.00036)	(0.00006)	(0.00016)
p-val	0.0000	0.0016	0.0269	0.0155	0.06340	0.00000	0.74800	0.00004
Constant	0.7920	0.0261	0.0095	0.0015	0.00503	0.00875	0.00078	0.00163
	(0.0005)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0)	(0.0001)
# obs	1136686	1177102	1177102	1177102	1177102	315600	1177102	315600

Table A8: Robustness check using first 5 rounds per user. *OLS of the behavior of interest against a treatment dummy using the first five observations per user. Standard errors are clustered at the user level. The gate-related outcomes are conditioned on the user being at a gate Treatment variable in bold indicates p-value<0.05.*

Similarly, **Table A9** shows the results when we add multiple controls per user. Specifically, we control for the activity-related variables that are observed at the moment in which the user

[14] One way to overcome this concern entirely would be to replicate the analysis using only the very first round of data. However, doing so discards a lot of useful information, dramatically reducing the sample size and thus the power to capture the observed effects.

receives the experiment for the first time. These include recency, frequency, past purchases (if any), number of round in current level, whether the user is stuck at a gate, number of days since they broke the previous gate, and various summaries of the level of activity on the previous day they played the game. All results are consistent with those presented in **Tables 6**, **Table 7**, and **Table 8** of the main analyses.

	Continue playing	Use extras	Use coins	Use Money	Coin extra	Coin gate	Money extra	Money gate
Treatment	**0.0202**	**-0.0016**	**-0.0006**	-0.0001	**-0.00062**	-0.00010	**-0.00012**	0.00001
	(0.0007)	(0.0003)	(0.0001)	(0.0001)	(0.00011)	(0.00019)	(0.00005)	(0.00008)
p-val	*0.0000*	*0.0000*	*0.0001*	*0.0660*	*0.00000*	*0.58000*	*0.01040*	*0.93200*
Constant	0.7850	0.0033	0.0021	-0.0109	-0.00821	0.03160	-0.00889	-0.00976
	(0.0083)	(0.004)	(0.003)	(0.0013)	(0.0021)	(0.00478)	(0.00103)	(0.00347)
# obs	1867811	2009921	2009921	2009921	2009921	652565	2009921	652565

Table A9: Robustness check adding controls. *OLS of the behavior of interest against a treatment dummy using all rounds on the first days of the intervention. Standard errors are clustered at the user level. The gate-related outcomes are conditioned on the user being at a gate Treatment variable in bold indicates p-value<0.01.*

A5 Implications for user targeting: Full set of results

In this appendix, we provide further details on the analysis conducted to explore the groups of customers who would be more responsive to the intervention (Section 4.3). First, we describe how each of the variables has been computed and report their summary statistics and correlations among them. Then, we present the full set of results.

Table A10 provides the summary statistics and the operationalization of each variable, while **Table A11** presents the pairwise correlations. The variables 'early progress,' 'progress prone,' 'spender,' and 'progression decline' are computed based on behavior prior to surpassing Level 20, whereas the 'frustrated' and 'distance' variables are derived from data at the time of the intervention.

Variable	Measure	Mean	St.Dev	P25	P50	P75
Early progress	# days to achieve Level 20	-8.69	15.74	-9.00	-3.00	-1.00
Progress prone	Proportion of rounds the user kept playing after winning	0.92	0.08	0.90	0.95	1.00
Spender	Whether user spent anything before Level 20	0.01	0.10	0.00	0.00	0.00
Frustrated	# unsuccessful rounds in current level	27.27	62.43	3.00	8.00	25.00
Distance	# levels until next gate	6.50	6.69	0.00	4.00	12.00

Table A10: Summary statistics for moderating variables.

Variables	(1)	(2)	(3)	(5)	(6)
(1) Early progress	1.000				
(2) Progress prone	0.206*	1.000			
	(0.000)				
(3) Spender	0.002	0.016*	1.000		
	(0.279)	(0.000)			
(4) Frustrated	0.095*	0.095*	0.003	1.000	
	(0.000)	(0.000)	(0.058)		
(5) Distance	0.146*	0.082*	0.020*	-0.110*	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	
** shows significance at p<.05*					

Table A11: Pairwise correlations among moderating variables.

Next, we report the results from the full model. Given the large number of variables in the regressions, we split the results by treatment and interaction effects (**Table A12**), and all sets of three-way interactions (**Table A13**).

	Playdays	Rounds	Progress	Purchases	Pay extra	Pay gate	Coins	Coin extra	Coin gate
Treatment	**0.631**	**8.664**	**2.097**	**0.0106**	**0.00415**	**0.00483**	**0.026**	**0.00919**	**0.631**
	(0.0203)	(0.249)	(0.0632)	(0.0023)	(0.0011)	(0.0004)	(0.0021)	(0.0018)	(0.0203)
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Early progress	**0.0946**	**1.504**	**0.93**	**0.0061**	**0.0031**	**0.0017**	**0.0127**	**0.0075**	**0.0946**
	(0.0228)	(0.295)	(0.0742)	(0.002)	(0.0011)	(0.0004)	(0.0023)	(0.0019)	(0.0228)
p-val	0.000	0.000	0.000	0.002	0.006	0.000	0.000	0.000	0.000
Progress prone	0.0052	**0.88**	**0.202**	0.00332	0.00107	0.000497	0.00344	0.0025	0.0052
	(0.0206)	(0.236)	(0.0611)	(0.0025)	(0.0011)	(0.0004)	(0.0022)	(0.0018)	(0.0206)
p-val	0.800	0.000	0.001	0.184	0.346	0.219	0.114	0.165	0.800
Spender	0.00722	-0.0527	-0.0199	**0.0277**	**0.0151**	**0.00503**	0.012	**0.0125**	0.00722
	(0.0213)	(0.243)	(0.0787)	(0.0114)	(0.0051)	(0.001)	(0.0066)	(0.006)	(0.0213)
p-val	0.735	0.828	0.800	0.015	0.003	0.000	0.067	0.036	0.735
Frustrated	**0.225**	**2.462**	0.0234	0.00215	0.00126	**0.00101**	0.00175	**0.00413**	**0.225**
	(0.0432)	(0.626)	(0.159)	(0.0019)	(0.0011)	(0.0004)	(0.0023)	(0.0019)	(0.0432)
p-val	0.000	0.000	0.883	0.264	0.269	0.011	0.441	0.027	0.000
Distance	**0.171**	**2.933**	**0.154**	**0.00587**	0.00112	**0.00385**	**0.0149**	0.00216	**0.171**
	(0.0203)	(0.256)	(0.063)	(0.0023)	(0.0012)	(0.0005)	(0.0022)	(0.0018)	(0.0203)
p-val	0.000	0.000	0.014	0.009	0.328	0.000	0.000	0.231	0.000
# obs	329999	329999	329999	329999	329999	329999	329999	329999	329999

Table A12: Results for the (simple) interaction effects between each moderator variable and the treatment effect. *OLS with treatment, all moderators, their interaction effects, and all possible three-way interactions. All moderator variables have been standardized before computing the interaction terms. Robust standard errors reported in parentheses. Interaction variable in bold indicates that p-value<0.05.*

	Playdays	Rounds	Progress	Purchases	Pay extra	Pay gate	Coins	Coin extra	Coin gate
Early progress x	**0.0742**	**1.11**	**1.213**	**0.00323**	**0.00149**	**0.000611**	**0.00513**	**0.00376**	**0.0742**
Progress prone	(0.0113)	(0.125)	(0.0342)	(0.0008)	(0.0005)	(0.0002)	(0.0011)	(0.0009)	(0.0113)
p-val	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000
Early progress x	-0.0238	-0.302	0.132	0.0148	0.00688	-0.000552	0.00251	0.0033	-0.0238
Spender	(0.0164)	(0.181)	(0.0798)	(0.0093)	(0.0056)	(0.0011)	(0.0081)	(0.0074)	(0.0164)
p-val	0.147	0.095	0.098	0.113	0.219	0.628	0.756	0.657	0.147
Early progress x	0.00357	0.782	**1.494**	**0.00414**	**0.00234**	0.000641	**0.00808**	**0.00479**	0.00357
Frustrated	(0.0434)	(0.621)	(0.144)	(0.0014)	(0.0008)	(0.0004)	(0.0023)	(0.002)	(0.0434)
p-val	0.934	0.208	0.000	0.004	0.005	0.081	0.001	0.016	0.934
Early progress x	**-0.108**	**-0.552**	**-0.191**	**0.00446**	**0.0021**	**0.00147**	**0.00612**	**0.00471**	**-0.108**
Distance	(0.0192)	(0.247)	(0.0595)	(0.0018)	(0.001)	(0.0004)	(0.0022)	(0.0018)	(0.0192)
p-val	0.000	0.025	0.001	0.013	0.043	0.000	0.005	0.007	0.000
Progr. prone x	0.0144	**0.405**	**0.13**	0.00125	-0.00431	0.00144	-0.00214	-0.0021	0.0144
Spender	(0.0161)	(0.165)	(0.0559)	(0.0091)	(0.0045)	(0.0009)	(0.0054)	(0.0051)	(0.0161)
p-val	0.372	0.014	0.020	0.891	0.342	0.112	0.690	0.679	0.372
Progr. prone x	**-0.202**	**-1.699**	**-1.06**	-0.00102	-0.00099	-0.000225	-0.00187	-0.00149	**-0.202**
Frustrated	(0.0312)	(0.44)	(0.104)	(0.0014)	(0.0009)	(0.0004)	(0.0024)	(0.002)	(0.0312)
p-val	0.000	0.000	0.000	0.478	0.278	0.540	0.437	0.455	0.000
Progr. prone x	**-0.0879**	**-0.719**	**-0.671**	0.00105	0.000451	0.000801	-0.00037	-0.0012	**-0.0879**
Distance	(0.0172)	(0.202)	(0.0493)	(0.0018)	(0.0009)	(0.0004)	(0.0018)	(0.0014)	(0.0172)
p-val	0.000	0.000	0.000	0.552	0.619	0.049	0.836	0.398	0.000
Spender x	**-0.0454**	**-0.601**	**-0.163**	**-0.0157**	**-0.00832**	**-0.00103**	**-0.011**	**-0.0111**	**-0.0454**
Frustrated	(0.0197)	(0.275)	(0.0359)	(0.0061)	(0.0033)	(0.0004)	(0.0043)	(0.0043)	(0.0197)
p-val	0.021	0.029	0.000	0.010	0.011	0.019	0.011	0.010	0.021
Spender x	0.0186	0.385	**0.229**	0.0174	**0.00941**	**0.00288**	**0.0179**	**0.0153**	0.0186
Distance	(0.0178)	(0.242)	(0.0666)	(0.0109)	(0.0048)	(0.001)	(0.0063)	(0.0059)	(0.0178)
p-val	0.296	0.112	0.001	0.108	0.047	0.005	0.004	0.010	0.296
Frustrated x	**0.122**	**1.365**	0.0603	0.00329	0.00199	**0.000873**	0.00124	**0.00375**	**0.122**
Distance	(0.0294)	(0.493)	(0.0981)	(0.0019)	(0.0011)	(0.0004)	(0.0019)	(0.0015)	(0.0294)
p-val	0.000	0.006	0.539	0.084	0.079	0.020	0.506	0.012	0.000
# obs	329999	329999	329999	329999	329999	329999	329999	329999	329999

Table A13: Results for the three-way interaction. *OLS with treatment, all moderators, their interaction effects, and all possible three-way interactions. All moderator variables have been standardized before computing the interaction terms. Robust standard errors reported in parentheses. Interaction variable in bold indicates p-value<0.05.*