cbess
Centre for Behavioural and Experimental Social Science

2017-10

IT'S NOT ALL FUN AND GAMES: FEEDBACK, TASK MOTIVATION, AND EFFORT

Sheheryar Banuri

Katarína Danková

Philip Keefer

It's not all fun and games: Feedback, task motivation, and effort

Sheheryar Banuri, Katarína Danková, and Philip Keefer[1]

November 2017

DRAFT: Please do not cite without permission

Abstract

Performance feedback is a pervasive element of education, management, marketing, and the entire gaming industry. Prior research on feedback focuses on what information individuals receive, and how frequently. The gaming industry, though, is built upon the premise that how feedback is delivered matters and, in particular, that "context" – narrative and story – are key. However, even as organizations increasingly adopt gaming elements into their feedback systems, prior research offers little guidance about whether standard performance feedback, combined with gaming elements, yields greater effort. We report the results of experiments that identify the impact of feedback through gamification, through a novel experimental design that introduces narratives into the task. Compared to standard performance feedback, gamification significantly increases effort in a "real effort" task. However, consistent with past research showing that intrinsic and extrinsic incentives interact, gamification has a positive impact on effort when extrinsic incentives are low, but no impact when they are high. The introduction of narrative – storyline development – induces the greatest effort, even compared to Leaderboards, a gaming element that often features in performance feedback systems.

Keywords: Intrinsic motivation, extrinsic motivation, performance, feedback, gamification, effort
JEL Classification Codes: J24, M12, C91

---

[1] Banuri: School of Economics, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK (e-mail: sbanuri@gmail.com); Dankova: School of Economics, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK (e-mail: katarina.danko@gmail.com); Keefer: Inter-American Development Bank, 1300 New York Ave, NW, Washington, DC, (e-mail: pkeefer@iadb.org). The authors have no relevant or material financial interests that relate to the research described in this paper. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. We wish to thank Maroš Servatka (MGSM Sydney), Oana Borcan (UEA Norwich) and participants of the 4th Lund Conference on Games, Interaction, Reasoning, Learning and Semantics 2016 and ESA European meeting 2016 in Bergen for valuable comments and suggestions. Financial support was provided by the University of East Anglia. Any errors remain our own.

"The opposite of play is not work, it is depression" – Brian Sutton-Smith

Introduction

A persistent question in education, marketing, management and behavioral economics is how best to use performance feedback to increase effort. Most social science research in this area focuses on the type of feedback (absolute versus relative performance, for example) and its frequency. However, a central principle of the gaming industry is that context matters: the way in which feedback is delivered also affects effort. We ask whether gamification – the integration of game elements into standard performance feedback – motivates individuals to exert more effort. These elements turn out to have a substantial effect, and, indeed, context does matter. Our experimental setting also allows us to contribute to a large literature on the interaction of intrinsic and extrinsic motivations: extrinsic incentives reduce the effect of gaming elements on effort; task motivation *increases* them.

Game developers invent elaborate point systems, symbolic rewards, leaderboards and storylines, none of which change players' tasks, to elicit interest in their games.[2] Video games ask players to drive a car or pursue an opponent, but embed these activities in different stories. Many board games involve rolling the dice and moving a token a corresponding number of squares, but in some games this effort yields a *Get out of Jail Free* card and in others it sends the opponent's token back to Start. These gaming elements obviously have an impact: not only do people exert substantial effort in games, whether measured in time or exertion; they also spend billions of dollars for the right to play them.[3]

Our experimental design considers the impact on effort of each of these four basic gaming elements (point systems, symbolic rewards, leaderboards, and storylines). The addition of all four gaming elements to performance feedback has large effects on performance. These are significantly stronger among individuals who enjoy the underlying task and for whom pecuniary rewards are not tied to effort. Most strikingly, narrative context, the gaming element that is perhaps the greatest focus of the gaming industry, though not of the literature on feedback or gamification, has the largest effect.

Gaming elements have already found their way into applications outside of the gaming industry. Marketers have introduced gaming elements in a variety of contexts to increase demand, for example for social networking services (Facebook), games themselves (Angry Birds) and location-based services (Foursquare) (Hamari 2013).[4] Large private sector firms, such as Uber, have utilized gaming techniques to elicit effort from their employees: Uber drivers are given badges for achieving certain quality benchmarks.[5] This phenomenon raises a natural question in the context of a large body of research on non-pecuniary motivation and effort

---

[2] These are also highlighted in the gamification literature (Deterding, et al. 2011).

[3] According to Newzoo's (a leader in games, e-sports and mobile intelligence) Global Games market report (accessible at: https://cdn2.hubspot.net/hubfs/700740/Reports/Newzoo_Free_2016_Global_Games_Market_Report.pdf) an estimated the global gaming market to be worth nearly 100 billion USD in 2016, with an estimated 2.2 billion gamers (https://newzoo.com/insights/articles/the-global-games-market-will-reach-108-9-billion-in-2017-with-mobile-taking-42/). Furthermore, according to the Nielsen 360º gaming report, the average US gamer spent over 6 hours per week playing video games in 2013 (http://www.nielsen.com/us/en/insights/news/2014/multi-platform-gaming-for-the-win.html).

[4] Gamification has already been applied in the promotion of greener energy consumption (EcoIsland), taking care of one's health (Fitocracy) and even to tracking one's life aspirations (Mindbloom).

[5] See more of Uber's use of gamification here: https://www.nytimes.com/interactive/2017/04/02/technology/uber-drivers-psychological-tricks.html

choices within organizations (e.g. Falk et al., 1999; Fehr and Falk, 1999).[6] Does the integration of game elements into organizational tasks – gamification – motivate individuals to work harder (i.e. exert more effort) on those tasks?

Research on performance feedback argues that clear goals and frequent performance feedback, unrelated to compensation, improves performance. The feedback could come in the form of evaluations of individual performance relative to some absolute standard, as in the management literature (e.g. Bloom and Van Reenen, 2007), or relative to others (Eriksson, Poulsen and Villeval, 2009). Psychologists argue that the central element of feedback is that it focuses on individuals' effort on the task rather than on their ability in general (Kluger and DeNisi (1996).[7] The considerable literature in economics on feedback and performance generally finds that feedback has heterogenous effects across agents (Eriksson et al. 2009; Azmat and Iriberri, 2010; Lizzeri et al., 2002; Ederer, 2008; Bandiera et al., 2007; Blanes-i-Vidal and Nossol, 2011; List and Rasul, 2011; Barankay, 2012; Dubey and Geanakoplos, 2005, 2010).

Games certainly incorporate the goal-setting and performance feedback attributes that are central to the feedback literature. Relative performance comparisons, such as leaderboards, are central to both games and feedback. Games go further, however, by introducing gaming elements that do not change the task, the goals, nor the performance information provided to players. Instead, they add other intangible elements that are also expected to increase performance by making goal achievement more interesting or "fun".[8] Greater effort not only permits players to accumulate points, but can change the rate at which they accumulate them. Effort not only allows players to achieve goals, but to receive symbolic rewards, such as badges or Pokémons, when they achieve them; and persistent effort not only raises output, but it advances a storyline that reveals itself more completely when players exert more effort.

Despite the success of these elements in encouraging people to purchase and play games, little previous research explicitly examines the effects of gamification on effort. Hamari, et al. (2014) identify 24 papers in their review of research on the effects of gamification. Substantively, most research focuses on the effects of gamification on learning in educational contexts. Methodologically, the research is largely descriptive (e.g., lacking control groups) and yields ambiguous evidence on the effects of gamification on effort.[9] No prior literature has examined a central gaming element that we study here, a narrative context that changes with effort. The gaming elements we examine relate to, but are distinct from, previous literature that has examined non-pecuniary motivations to engage in a task (e.g., Kosfeld and Neckermann,

---

[6] Motivation refers to a desire to exert effort, and comes from a variety of sources, including external controls, incentives, punishments, rewards, etc. (e.g., Herzberg, 1966; Porter & Lawler, 1968; Staw, 1977).

[7] In education, scholars take a broader view of feedback. For example, they debate whether more thorough instructional feedback matters more than rewards for performance (Hattie and Timperley 2007). The former conveys significantly more information than the latter, so we would expect it to have larger effects on learning. In our experiments, we focus on feedback that speaks only to results and effort – the least powerful form of feedback according to the meta-review of Hattie and Timperly (2007) and hold the information content of feedback fixed across experiments using different gamification elements.

[8] Psychologists debate what precisely makes these elements "fun" and engaging. A common thread running through the debate overlaps with the feedback literature: gaming elements are a sophisticated form of feedback: they provide accelerated feedback, clear goals and challenging tasks (Hamari 2013). However, recognizing the importance of context and narrative in games, others argue that gaming elements give "meaning" to a particular task (Deterding, 2011; McGonigal, 2011).

[9] An important exception is research conducted by Mekler, et al. (2013). They allow effort to affect whether subjects accumulate points, shift their position on leaderboards and achieve levels of achievement and ask whether these gaming elements affect effort in a real effort task, image annotation. Relative to a no-feedback baseline, gamification increases performance. They do not address the interaction of gamification, extrinsic and task motivation, the focus of this paper.

2011; Ashraf et al. 2014; Bradler et al. 2016; Ariely, Kamenica, and Prelec, 2008; Banuri, Keefer, and de Walque 2017). This work particularly examines the effects of pro-social or task motivation, but also status and peer effects. For example, Kosfeld and Neckermann (2011) show that effort increases when top performers on a task are promised a congratulatory card signed by the managing director of the organization that contracted the task. This award conveys social recognition and status, precisely the motivations that the researchers seek to examine. Gaming feedback elements, in contrast, generally do not invoke pro-social missions; they carry no additional status (leaderboards, for example, reveal only a player's relative position but not the identity of other players), and they are separate from the task.

We also consider an important issue raised by prior literature, whether the effects of new forms of intrinsic motivation depend on the pre-existing extrinsic and intrinsic motivations of individuals (Frey and Oberholzer-Gee 1997; Deci and Ryan 1985). Extrinsic incentives reduce intrinsic motivation in certain contexts (Deci et al, 1999; Frey and Oberholzer-Gee, 1997; Ryan and Deci, 2000; Deci and Ryan, 1985). For example, Ashraf et al., (2014) find that agents who are offered non-financial rewards exert more effort than those offered financial margins or volunteer contracts. Banuri, Keefer and de Walque (2017) examine the interaction of two intrinsic motivations, pro-sociality and task enjoyment, and find that the first has no additional effect on effort when the second is high. Casas-Arce and Martinéz Jerez (2009) study the loss of motivation in tournaments. They find that winning participants decrease their effort as their lead extends, whereas the effort of trailing participants fades only when the gap to a winning position is very large.

We test whether the returns to gamification differ when extrinsic incentives vary, furthering our understanding of the effects of non-pecuniary incentives on effort. Eriksson et al. (2009) find that feedback on performance – absent gaming elements – generates higher effort in tournament settings, but not in piece rate settings. We compare feedback with gaming elements across flat salary and piece rate settings. Prior research examines the interaction of different non-pecuniary motivations (for example, Ariely, Bracha and Meier, 2009; Carpenter and Myers, 2010; Friedrichsen and Engelmann, 2017; and Banuri, Keefer, and de Walque, 2017). We also examine how the effects of gamification feedback of a task vary with individuals' task motivation.

Social scientists have expended considerable effort in understanding how non-pecuniary incentives can be marshalled to elicit effort. Much of this has focused on social incentives of various kinds, from peer pressure to mission, and the interaction of social and extrinsic incentives. However, recent evidence suggests that intrinsic task motivation may have a significantly larger effect on effort than social incentives and, where task motivation is significant, social motivation may have a small effect. Other research has emphasized performance feedback, but its effect on effort among the task-motivated is unknown. In addition, researchers have focused on what information is provided to workers, not how it is provided. However, the gaming industry is predicated upon the fact that people are willing to pay large sums to engage in familiar tasks that manipulate performance feedback in novel ways. We extend these inquiries by examining the effect how the introduction of gaming elements into performance feedback influences effort by both motivated and unmotivated subjects.

Our paper makes three main contributions to the literature: First, we shed light on how the returns to feedback (gamification) interact with extrinsic rewards: gamification has a positive impact on effort when extrinsic incentives are low, but no impact when extrinsic incentives are high. Second, we identify the mechanism through which gamification operates: gamification (as feedback) interacts through task motivation. Task-motivated individuals exert significantly higher effort when gamification is present. Third, we evaluate the effectiveness of each gaming element on effort separately. Competition within individuals (Leaderboards) and explicit

additional output (Context) significantly increase effort. The results also reveal heterogeneous effects: competitive individuals drive the effects of Leaderboards, but the effects of context are independent of preferences.

To our knowledge, this study is the first to provide evidence for the motivational mechanisms at work when manipulating the way that feedback and goals are communicated, and the effects of this manipulation on performance in labor markets where effort is and is not contractible (extrinsic incentives are high- or low-powered). We find that selection matters for the returns to gamification: those who do not enjoy the task are immune to its motivational effects. Our results contribute to research seeking to disentangle the effect of intrinsic and extrinsic incentives on motivation. Our results are important for the efficiency of operations in organizations, and for managers seeking to enhance motivation for work tasks.

Experimental design and procedures

An illustrative model fixes ideas. Following Banuri, Keefer, and de Walque (2017), there are diminishing returns to motivation, both in the sense that additional task or pecuniary motivation has a diminishing effect on effort, but also that when one type of motivation is substantial, changes in the other type have a smaller effect on effort. The utility from effort is then described by (1) below.

$$u_i = w_f + \left(w_p + (1+\delta)\gamma_i\right) ln(e_i) - \frac{1}{2}e_i^2 \qquad (1)$$

Worker contribution to output is given by $ln(e_i)$ – output rises in effort, but there are diminishing returns. Worker utility is separable in the welfare improvements that they experience from engaging in tasks (pecuniary and intrinsic motivations) and the disutility of effort. Worker utility rises with the flat salary, w_f, which is independent of their effort. It rises with effort depending on the piece rate w_p, and the workers' task motivation (γ_i), but at a declining rate. Task motivation can be modified by the presence of feedback (inclusive of gaming elements). The parameter δ establishes the extent to which gaming elements are available to the worker. The feedback and gamification literature implies that feedback (gamification) operates through and are dependent on task motivation. Hence, we assume that the marginal effects of gaming elements fall with task motivation. The exertion required by additional effort reduces utility, at an increasing rate ($-\frac{1}{2}e^2$).

Maximizing utility with respect to effort yields optimal effort $e_i^* = \sqrt{\left(w_p + (1+\delta)\gamma_i\right)}$. This yields the two key comparative statics that we examine below: $\frac{\partial e_i^*}{\partial \gamma_i} = \frac{1+\delta}{2\sqrt{\left(w_p+(1+\delta)\gamma_i\right)}} > 0$; and $\frac{\partial e_i^*}{\partial \delta} = \frac{\gamma_i}{2\sqrt{\left(w_p+(1+\delta)\gamma_i\right)}} > 0$: effort rises in motivation and in feedback (gamification). In addition, the effort effects of both motivation and gaming elements are increasing in the levels of gaming elements and motivation, respectively: $\frac{\partial e_i^{*2}}{\partial \delta \partial \gamma_i} > 0$. On the other hand, the greater the contribution of extrinsic incentives to utility, the lower are the motivational effects of gamification: $\frac{\partial e_i^{*2}}{\partial \delta \partial w_p}, \frac{\partial e_i^{*2}}{\partial w_p \partial \delta} < 0$. Gamification (feedback) has a smaller effect on effort the larger are extrinsic incentives.

We test whether effort rises when gaming elements are introduced into tasks. In addition, we test whether the effort effects of gamification decline with extrinsic incentives. In

fact, gamification significantly increases effort only when extrinsic incentives are low. Effort varies little with gamification when extrinsic incentives are high.

The experimental design is simple: we use a basic real effort task (decoding task) which provides subjects with five numbers and a code which subjects are supposed to use to decode the numbers into words, one letter at a time (see Figure 1). The experiment uses 6 distinct code sets (each randomly generated prior to the experiment) which assigns one number to each character of the alphabet.[10] Each round of the effort task lasts for 2 minutes, during which subjects decode a preset list of 5-letter words from the given numbers. Researchers have frequently employed this effort task to generate endowments in lab experiments (see, for example: Erkal et al., 2011; Neitzel and Saaksvuori, 2013). It is useful for its simplicity, and because it allows us to examine one of the more challenging aspects of gamification: the introduction of context to tasks.[11]

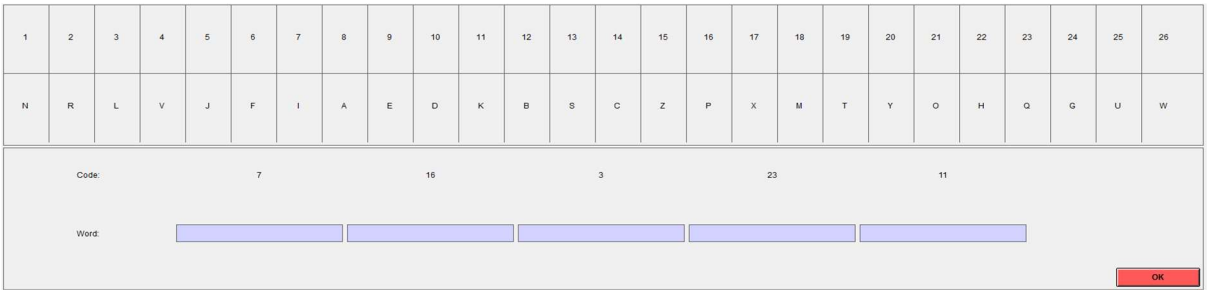


Figure 1: Decoding task

In the baseline ("Flat Salary") treatment, subjects are paid a flat (unconditional) salary of 400 tokens (4 GBP) and asked to undertake the task for at least one round. During and at the end of the first round, subjects are given feedback on their performance.[12] They are then asked whether they would like to continue the task for an additional 2 minutes, or to end the task. As the instructions make clear, subjects understand that ending the task means starting the exit survey, after the completion of which they are free to leave. Subjects are not told how many times they can continue the task, but they can continue for a total of 13 rounds, after which the exit survey automatically begins.[13] Hence, in the baseline treatment, subjects are paid a flat salary and can engage in the task for between 2 and 28 minutes (a minimum of 1 round, and a maximum of 14 rounds). This treatment offers low extrinsic motivation to exert effort (wages are flat) and the only task incentive is provided by precise feedback on performance.

Further treatments modify the baseline treatment by manipulating feedback (adding gaming elements) and extrinsic motivation (providing a piece rate), yielding a 2-by-2 factorial

[10] We use six different code sets to mitigate learning. Using a single code allows subjects to memorize the code and become faster as they continue the task. Resetting the code allows us a way to mitigate this to some degree.

[11] Please see next sub-section on gaming elements for how we operationalize context.

[12] During the round, subjects are provided with a running list of the words that they have decoded correctly, as well as the total number of words coded correctly, the percentage of words coded correctly, as well as the tokens earned. At the end of the round, subjects are provided with a history table containing the round number, the number of words coded correctly, and the percentage of words coded correctly in the round. The history table contains a history of the preceding three rounds.

[13] The decision not to inform the subjects of the maximum number of rounds reduces the impact of focality. Informing subjects of the end may have biased decisions upwards depending on their levels of reciprocity. To avoid this, we did not provide the subjects with a maximum number of rounds.

design. Table 1 displays the full study design. Screenshots and instructions can be found in appendix B.

Table 1: Study design: Number of subjects in each treatment are given in parentheses

		Feedback	
		Feedback without gaming elements	**Feedback with gaming elements**
Extrinsic Motivation	**Low**	Flat Salary (n=43)	Flat Salary + Gaming Elements (all) (n=51)
	High	Piece Rate (n=55)	Piece Rate + Gaming Elements (all) (n=40)

The high extrinsic motivation treatment (Piece Rate) implements a piece rate: it asks participants to engage in the same effort task, but instead of providing subjects with an unconditional payment, subjects are informed that they will be paid a piece rate of 3.6 tokens per word coded correctly.[14] As with the salary treatment, subjects choose whether to continue or end the task at the end of each 2-minute round of decoding.

The final two treatments add gaming elements to the basic effort task presented above, under the flat salary and piece rate treatment conditions. We focus on the four most common elements of gamification: points; badges; leaderboards; and narrative context.[15] We describe each of these elements in detail below. The high feedback treatments ("Flat Salary + Gaming" and "Piece Rate + Gaming") both add gaming elements to the base (flat salary and piece rate) treatments, but leave the extrinsic incentive structure unchanged. However, neither the basic task, nor the payment to the subjects are affected.

Gaming element I – Points: Points provide a numerical score of performance attained while engaging in a task, and so constitute real-time feedback on performance. By themselves, then, points convey the same information as the feedback used in the baseline, which informs subjects of the number of words coded correctly. A principle of many classes of games, however, is that points can elicit greater task motivation when effort affects the rate of point accumulation, and not only point accumulation itself. We therefore award subjects 50 points for each word they decode correctly and then increase the points score by a multiplier that starts at 1 and increases by 0.1 for each consecutive word subjects decode. If subjects make a mistake, the multiplier resets to 1. The points convey no more information about performance than subjects receive in the baseline treatment, but add a gaming element that might increase subject motivation: additional effort increases both the points that they receive and the multiplier applied to those points.

[14] The piece rate was selected to equalize earnings across treatments to minimize differences in behavior due to differences in income. On average, subjects earned 414.52 tokens in the piece rate treatment, which is not significantly different (one-sample t-test: p=0.49) from 400 tokens (earned in the salary treatments).

[15] For more information on gaming mechanics, see Kumar and Herger (2013): https://www.interaction-design.org/literature/book/gamification-at-work-designing-engaging-business-software/chapter-6-58-mechanics

Gaming element II – Badges: Goal-setting is common to both games and traditional feedback. However, gamification can magnify the task motivation of goal-setting, for example by offering symbolic rewards, such as badges. These rewards are unrelated to the task, convey no additional information about goal-achievement, and have no extrinsic value; nevertheless, they may create a desire by individuals to accumulate these rewards, thereby inspiring greater task effort. Badges, typically a graphic or a picture, provide recognition to the individual who has achieved some goal. Different badges are given for greater accomplishment and the hierarchy/importance of individual badges is common knowledge. In this treatment, we introduce a series of (on-screen) rewards based on preset targets. For any round, if subjects decode a minimum of 4 words, they will receive a bronze badge, with a minimum 7 words, a silver badge, and so on. The more words subjects decode (in a given round), the better the badge they receive. Table 2 displays the badges achievable by the subjects.

Table 2: List of targets and badges achievable

Number of words decoded correctly	Level	Medal
0-3	None	None
4-6	Bronze	
7-9	Silver	
10-12	Gold	
13-15	Diamond	
16-18	Platinum	
18+	Titanium	

Gaming element III – Leaderboard: Relative performance feedback is a pervasive element in the feedback literature, and in games. In games, this feedback is achieved by showing player performance rankings on leaderboards. These provide relative performance feedback either among all users (a global leaderboard) or among a certain subgroup (local leaderboard) to induce competition among players, and thereby increase effort.

Subjects are assigned a rank and placed on the leaderboard based on the number of words they decode correctly in a round. With higher performance, subjects achieve higher ranks relative to their peers. The leaderboard was generated based on data from the first three

sessions, and hence constitutes a local leaderboard.[16] Subjects were ranked on the leaderboard according to their highest score. To maintain independence within sessions, the leaderboard that subjects view in the session does not include other subjects in the same session.

Gaming element IV – Narrative Context: Though unknown in the feedback literature, context is an essential element of games that is believed to motivate players to engage in (and purchase) them. The effects of context on effort have not been rigorously tested, however. Context consists of a narrative or storyline that changes the meaning of a task. The other three gaming elements start with the basic performance information that is commonly transmitted in performance feedback and then manipulates how that information is conveyed. In contrast, narrative context allows individuals to generate an additional output – a story line – from their efforts. They receive no additional reward for this (their pecuniary compensation is unrelated to the outcome of the story); nor does this output demand any different or additional effort from them. It is merely the case that if they happen to work harder on the underlying task, they will see more of the story unfold.

We introduce context by developing 14 short stories, taken from a variety of sources, including fables from different cultures. We then edited the stories for similar length and to maximize the number of 5-letter words contained within them. These words were then eliminated from the story. Subjects engage in the same decoding task, but now each decoded word decoded advances the story.[17] Each round contains a different story. The amount of the story that is revealed depends on the number of words that subjects decode correctly. The more words they decode, the more of the story they get at the end of the round.

Here is an example of a story used in the experiment (with the coded words underlined):

> *On a rainy day, a young girl found a small black snake. It spoke to her: "If you kiss me, I will transform into a prince." The girl took the snake and put it in her purse. Again, the snake cried, "I have a power to transform from a snake to a prince!" The girl did not react. The snake got upset and cried aloud, "What is wrong with you, don't you fancy being my queen?" The girl looked at the snake and said, "I am a pilot, I spend sixteen hours a day on a plane. I do not have time for a prince, but a talking snake is fantastic!"*

Additional treatments

As Table 1 indicates, our main results examine the joint effects of all feedback/gaming elements simultaneously. However, the effects of the individual gaming elements are of interest, as well. For example, some (e.g., leaderboards) are familiar from the feedback literature, while others are distant from it (e.g., narrative context). Therefore, in addition to the four treatments used for the main study design, described in Table 1, we conduct four additional treatments that add each feedback (gaming) element independently to the baseline treatment. These treatments are conducted exclusively with low extrinsic incentives -- subjects receive a fixed payment of 400 tokens.

---

[16] The first three sessions implemented the baseline (no gamification), context, and badges separately. The leaderboard was constructed from the effort exerted by all subjects in each of these three sessions. In subsequent rounds in which there was a leaderboard, subjects were given a rank based on how many words they decoded within the round. At the end of every round, subjects were shown their rank on the leaderboard for that round and the previous two rounds.

[17] Note that if a word is repeated in the story, it still only needs to be decoded a single time.

Effort and motivation measures

The measurement of individuals' effort and their task motivation involves several key decisions. This section describes how we address these measurement issues.

Effort

Individuals can respond to enhanced motivation by increasing their efforts on the intensive or extensive margins (e.g., output per hour and number of hours worked). An advantage of our laboratory setting is that we can separately measure these: the number of words that subjects decode per round, the number of rounds that subjects undertake, and then total words decoded. Measures of performance in the feedback literature can typically only measure the last, aggregate measure. For example, Azmat and Irriberri (2010) examine the effects of performance feedback on student test scores, which are a product of effort on the extensive margin (time spent studying) and on the intensive margin (work intensity while studying).

Among these measures, in the analysis below, we focus on effort at the extensive margin, the number of rounds that subjects participate. There are two reasons for this. First, all our treatment effects operate through this channel; subjects do not significantly change the intensity of their decoding efforts across treatments (words per round). Second, the number of words that subjects decode is a function of both their effort and ability. However, subjects are heterogeneous in this ability. Although our measure of ability strongly predicts the number of words decoded per round, it is nevertheless an inevitably imperfect measure. Hence, even controlling for measured ability, total words are a noisier measure of effort than the number of rounds subjects undertake. Despite this noise, we show that our results are robust to using total words decoded rather than total rounds of participation as the measure of effort (Tables A.2 and A.3).

Survey measures – task motivation

The exit survey includes several socio-demographic variables (age, gender, state of personal finances, risk preferences, and academic level). Because we conjecture that the effects on task performance of feedback and gaming elements depend on individuals' intrinsic motivation to undertake the task, we also measure task motivation in the survey, using Ryan's (1982) task evaluation questionnaire. The questionnaire is a part of Intrinsic Motivation Inventory (IMI), which is grounded in Self-Determination Theory, used in assessing the subjective experiences of participants when developing an activity. The interest/enjoyment subscale of IMI serves as a proxy for task motivation. Appendix C provides the questions for the task evaluation questionnaire.

Ideally, we would measure task motivation at the beginning of the experiment so that, by construction, it would have been independent of our gamification treatments. For example, we could have asked the subjects to decode for a round, with no motivational elements present, and then inquired about how much they enjoyed the decoding task. However, a pre-experimental measure could have induced potentially significant experimenter demand effects that varied in unforeseen ways by treatment, biasing our results.

Instead, therefore, we asked about task enjoyment in the exit survey. To offset the risk that the gamification treatments could have influenced task enjoyment, we were careful to word the questions such that they referred repeatedly to the decoding task itself, not to subjects' overall impressions of the activity.

This approach is valid only to the extent that subjects indeed evaluated task enjoyment and ignored the gaming elements associated with feedback about task performance. We present

evidence that supports the assumption that our raw measure of task motivation is independent of the treatments. In addition, our main results are robust to using a different measure of task motivation that, by construction, is orthogonal to the treatments.

Survey measures – control variables

The extensive exit survey filled out by study participants allows us to control for a variety of subject characteristics, ranging from personality traits to risk aversion. The links among preferences, motivation and effort are a longstanding subject of research. In our exit survey, subjects respond to 14 item Competition Index (Smither and Houston, 1992; Houston et al., 2002) designed to measure "a desire to win in interpersonal situations." This item has high reliability and stable properties (Harris and Houston 2010). Risk aversion was measured using subject response to the question "Are you someone who takes risks or do you generally avoid taking risks?" We also controlled for preferences for video games by using subject response to the question "How often do you play video games?" Finally, we use the 10-item measure of Big 5 personality traits developed by Gosling et al. (2003) to control for personality characteristics. Our results are robust to controlling for personality traits.

Implementation

The experiment consists of four main parts. Its structure is depicted in Figure 2. Subjects have a chance to familiarize themselves with the decoding task in a practice round. We then include an "ability measure" prior to the main task, which is incentivized with a piece rate of 10 tokens per word decoded correctly in two minutes. Afterwards, the actual task follows. After each round, subjects decide whether to continue the task or not. The experiment concludes with the exit survey.

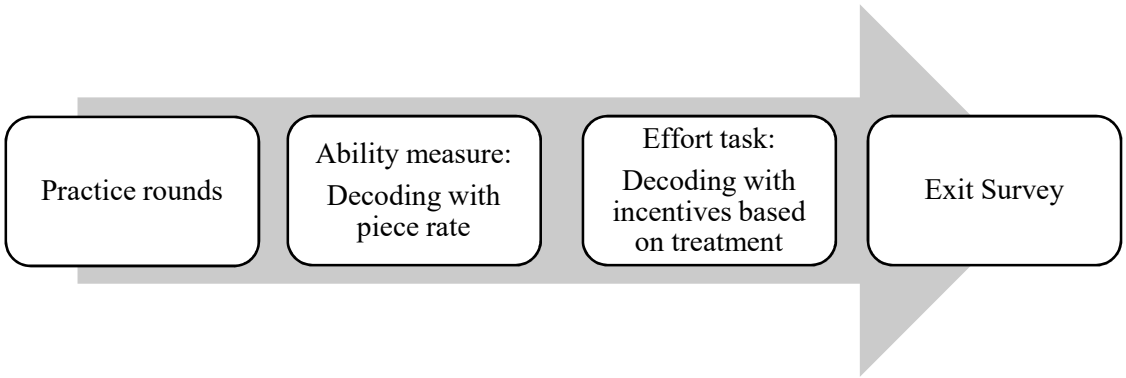


Figure 2: Experiment structure

A total of 362 subjects (undergraduate students at the University of East Anglia) participated in the experiment.[18] Subjects were recruited using the online database system Hroot (Bock et al., 2012). Each subject only participated in a single session of the study. The experiment was programmed and conducted with z-Tree (Fischbacher, 2007). All sessions were run under a single-blind protocol. On average, a session lasted 45 minutes, including the payment, and subjects earned £8.

[18] In addition to the treatments reported in this paper, we conducted one additional treatment with no feedback whatsoever (i.e. subjects received no information on the number of words decoded), as even our baseline treatment has some measure of feedback. We conducted this to understand the effects of no feedback (relative to basic feedback) on effort choices. Previous research finds that basic performance feedback has heterogeneous effects and, for some, those effects can be zero or negative. We find that a baseline with no feedback and a baseline with feedback yield statistically indistinguishable levels of effort.

Before entering the laboratory, subjects drew a ping-pong ball with a number from an opaque bag. They then took their seat in a cubicle with a number corresponding to the number on the ball. Subjects were then welcomed and instructed to ask all questions in private, after raising their hand. All instructions were on the computer screen and subjects could proceed at their own pace. Subjects had to respond to a simple quiz to demonstrate understanding prior to proceeding to the decision-making part of the experiment. After completing the main task, subjects filled out the exit survey. Once they finished, they were instructed to take all their belongings, collect their payment in private at the back of the laboratory, and leave quietly. Table 3 provides summary statistics for our sample.

Table 3: Summary statistics of the entire sample

Observations	362	
Female (%)	59%	
Age (years)	20.08	(2.64)
Risk preferences (5=risk-seeking)	3.09	(0.95)
Competitiveness (5 = Competitive)	3.32	(0.57)
State of personal finances (5 = v. good)	3.04	(0.96)
Gaming frequency (6 = 2+ times/day)	2.69	(1.44)
Ability in coding task	9.22	(2.29)
Task motivation	3.06	(0.82)

To summarize, subjects were asked to engage in the decoding task for as long as they liked. As they decoded the words, the program listed the words at the bottom of the screen. After each round of decoding, subjects are provided performance feedback, which depends on the treatment. At the end of each round, subjects were asked whether they would like to continue with the decoding, or to end the task. The next section presents the results of the experiment.

Results

The treatments allow us to address three questions: Does the addition of gaming elements to enhance performance feedback increase effort relative to the use of simple performance feedback? Does the effect of gaming elements depend on extrinsic motivation? And does it depend on task motivation?

The first two bars of Figure 3 address the first question. We compare effort under the baseline treatment, which offers standard feedback on performance under a flat salary regime, with effort in the presence of all gaming elements, holding constant extrinsic motivation. The figure displays the average levels of effort (measured on the extensive margin, the number of rounds that subjects undertook the task) in the traditional feedback regime and in the presence of gaming elements. Effort is significantly greater after gaming elements have been added. Adding gaming elements to traditional performance feedback increases (more than doubles) the quantity of effort provided by the subjects in the salary treatments (two sample t-test: $p<0.01$).

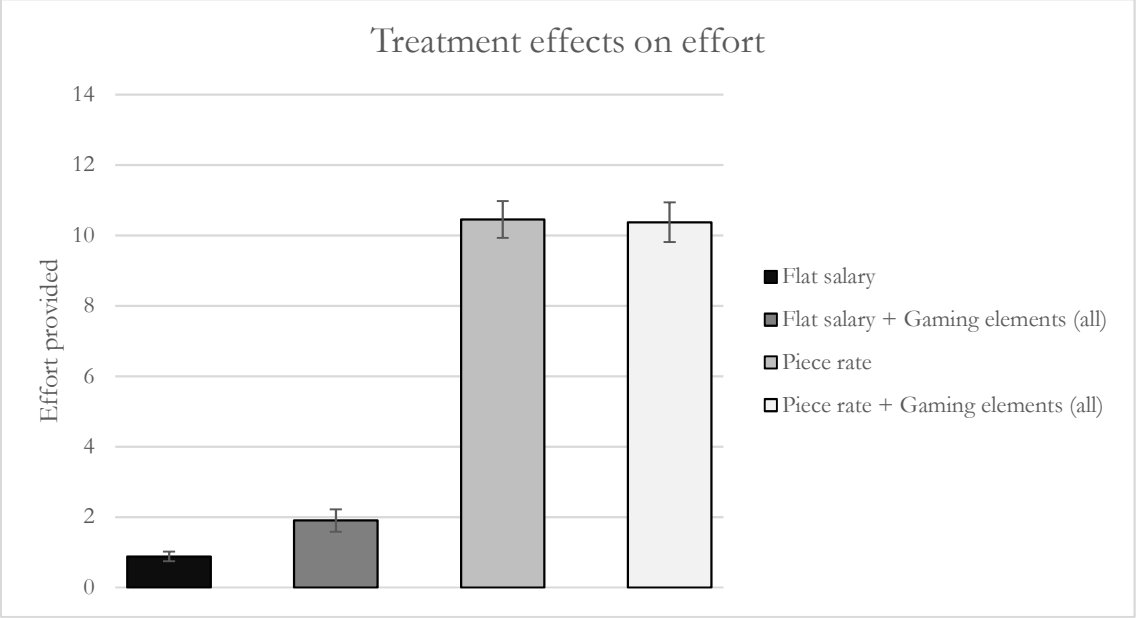


Figure 3: Effects of gaming elements on effort

To verify that the results in Figure 3 are robust to controlling for subject characteristics, we compare the differences in effort holding constant control variables taken from the exit survey. We employ a tobit specification, since subjects could continue the effort task for a maximum of 13 times, censoring our effort data at 13. Column 1 in Table 4 presents the results of a specification that replicates Figure 3, with no controls. As the results in column 2 demonstrate, the effects of gamification remain significant and large in the presence of controls. In addition, we also find a significant relationship between subject self-reported state of personal finances and effort exerted. Subjects that reported better states of personal finances exerted greater effort ($p<0.05$).

Table 4: Effort exerted (rounds continued)

Dependent Variable: Quantity of effort provided (Rounds continued)				
	I	II	III	IV
Treatment: Gaming elements	1.024***	0.897***	1.040***	0.942***
	(0.28)	(0.30)	(0.28)	(0.31)
Treatment: Piece rate			11.22***	10.99***
			(0.64)	(0.59)
Interaction: Gaming elements X Piece rate			-1.458	-1.251
			(1.49)	(1.49)
Gender		0.818		1.200*
1 = Female		(0.59)		(0.68)
Age (in years)		0.027		0.031
		(0.03)		(0.05)
Risk preferences		0.298		0.349
5 = Risk-seeking		(0.24)		(0.32)
Competitiveness		-0.293		0.445
5 = Highly competitive		(0.21)		(0.63)
State of personal finances		0.352**		0.608*

5 = Very good		(0.15)		(0.34)
Gaming frequency		0.210		0.320
6 = More than twice a day		(0.13)		(0.21)
Ability in task		0.084		0.049
		(0.06)		(0.16)
Constant	0.884***	-2.375	0.884***	-6.040*
	(0.19)	(1.77)	(0.18)	(3.09)
Observations	94	94	189	189
Log likelihood	-187.8	-182.5	-416.9	-412.2
Left censors	0	0	0	0
Right censors	1	1	54	54

Notes: Tobit regressions with censoring at 13 (maximum number of times subjects could continue the task). The dependent variable is the number of rounds subjects continued the coding task. Models 1 and 2 focus on the fixed payment (salary treatments) while models 3 and 4 add the piece rate treatments. * 10%, ** 5%, *** 1% significance level. Robust standard errors with sessions-level clusters in parentheses.

Shifting focus to the third and fourth bars of Figure 3, and comparing them to the first two, we can address the second question, the degree to which the effects of gaming elements are sensitive to extrinsic incentives. Comparing the first two bars with the third and fourth allows us to see the combined effect of all gaming elements on the quantity of effort provided when extrinsic incentives are low (flat salary) or high (piece rate). The figure displays the average levels of effort (again, the number of times subjects continued the task) in the flat salary (400 tokens total) and piece rate (3.6 tokens per word) treatments, with and without gaming elements.

Though enhanced feedback (the addition of gaming elements) has a large effect on effort when extrinsic motivation is low (flat salary), it does not increase effort when extrinsic motivation is high ($p=0.92$). Consistent with the model and expectations, gamification increases effort when extrinsic incentives are low, but has no effect when incentives are high.

To see if the results in Figure 3 are robust to controlling for subject characteristics, we regress effort on the gamification treatments, extrinsic incentive treatments, and the interaction of the two. The third and fourth columns of Table 4 report the results of these estimates. The third column, with no control variables, replicates Figure 3. The fourth adds control variables from the exit survey. Estimated treatment effects are unchanged in the presence of numerous controls.

We then turn to the third question: do the effects of gamification depend on task motivation? Our measure of task motivation is the subjects' evaluation of their enjoyment in the task (taken from the Intrinsic Motivation Inventory, Ryan,1982). Figure 4 displays how effort differs between subjects reporting task motivation levels at the median or below, and those reporting above median motivation for each treatment. Task motivated subjects exert higher effort in the baseline (salary) treatment (2-sample t-test: $p<0.1$) and in the salary + gaming elements treatment ($p<0.01$). In both the piece rate treatments, however, more task motivated subjects are not significantly different from the less motivated ($p=0.16$ and $p=0.28$, respectively).

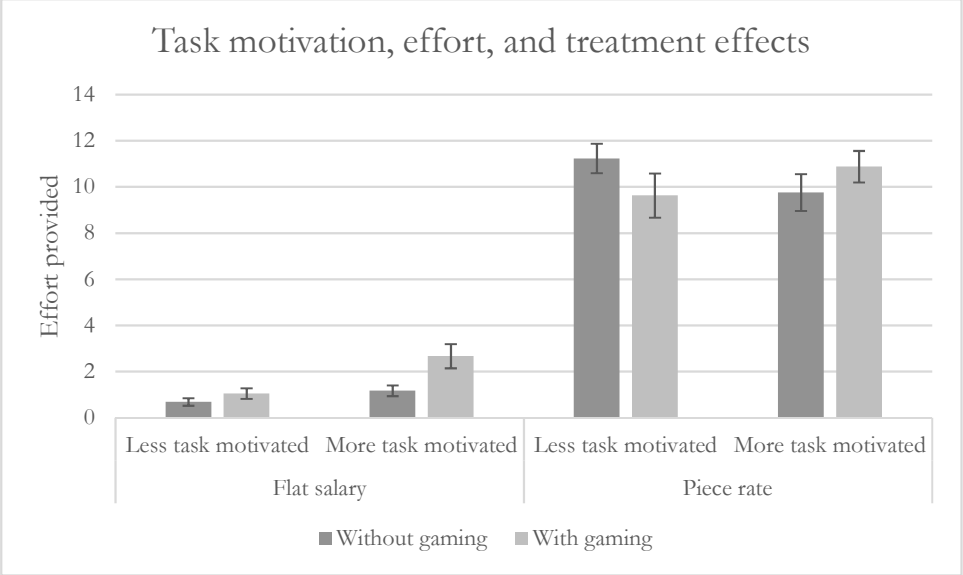


Figure 4: Task motivation, effort, and treatment

We explore these results in greater detail in Table 5, presenting the results of the analysis of the effects of gaming elements, extrinsic incentives, and task motivation, on the quantity of effort supplied. Model 1 simply adds the measure of task motivation to model 4 in table 4. The unconditional effect of task motivation is positive and significant ($p<0.10$) overall. Model 2 adds an interaction term between task motivation and the gaming treatment, to assess whether the average effect of gaming elements, across piece rate and flat salary treatments, is greater when task motivation is higher. Model 3 then introduces an additional test, of whether the effects of task motivation are contingent on extrinsic incentives (piece rate). Finally, model 4 collects all of the interactions, adding a three-way interaction among the gaming treatments, the piece rate treatments, and task motivation.

The results are in line with intuition. The linear task motivation coefficient in Model 1 shows that motivated subjects are significantly more likely to exert greater effort in the baseline ($p<0.10$). The interaction of gaming elements and task motivation in Model 2 is large, positive and highly significant: motivated subjects exert significantly greater effort in the treatments with gaming elements, across both salary and piece rate conditions. Model 3 demonstrates that these effects are, indeed, separately significant under both salary ($p<0.01$) and piece rate conditions ($p<0.01$): the interaction of piece rate and task motivation is insignificant, indicating that more task motivated subjects are not likely to increase effort over the less task motivated under piece rate conditions, consistent with motivation crowding theory (Frey and Oberholzer-Gee, 1997) ($p<0.22$).

From Figure 4 and prior results, we know that extrinsic incentives have a large effect on effort, and that gaming elements have a large effect on effort when task motivation is high, but this effect appears to be driven by subject effort when extrinsic motivation is low (subjects receive a flat salary). Model 4 allows us to examine the effects of extrinsic incentives directly, rounding out the examination of interaction effects by adding a triple interaction. The three-way interaction of motivation, gaming elements and piece rate compensation is large, significant and positive ($p<0.01$).

Model 4 confirms that the effects of gaming elements have the strongest effect on behavior when individuals are task motivated and working under a flat salary. To see this, we first observe that the interaction of piece rate and task motivation is -0.75: going from flat salary to piece rate has a lower effect on effort when individuals are task motivated. The triple

interaction is 1.58, indicating that the addition of gaming elements increases effort, but not enough to yield a significant effect: the addition of gaming elements under piece rate does not yield significantly more effort when individuals are task motivated. Conversely, under a flat salary, the effects of gaming elements unambiguously increase with task motivation (the interaction of the two is significant and positive).

Table 5: Effort, treatments, and task motivation

Dependent Variable: Quantity of effort provided (Rounds continued)				
	I	II	III	IV
Treatment: Gaming elements	0.833***	-3.126**	-3.118**	-1.022**
	(0.26)	(1.30)	(1.31)	(0.51)
Treatment: Piece rate	10.91***	11.03***	11.28***	13.26***
	(0.60)	(0.60)	(1.51)	(1.99)
Interaction: Gaming elements X Piece rate	-1.181	-1.297	-1.284	-6.174***
	(1.48)	(1.42)	(1.47)	(1.98)
Task motivation	0.412*	-0.118	-0.077	0.254*
	(0.23)	(0.31)	(0.31)	(0.13)
Interaction: Gaming elements X Task motivation		1.299***	1.294***	0.599***
		(0.45)	(0.45)	(0.16)
Interaction: Piece rate X Task motivation			-0.083	-0.747
			(0.50)	(0.61)
Interaction: Gaming X Piece rate X Task motivation				1.581***
				(0.60)
Gender	1.127	0.944	0.949	0.915
1 = Female	(0.70)	(0.71)	(0.73)	(0.71)
Age (in years)	0.026	0.018	0.019	0.017
	(0.05)	(0.05)	(0.05)	(0.05)
Risk preferences	0.365	0.413	0.413	0.449*
5 = Risk-seeking	(0.29)	(0.28)	(0.28)	(0.27)
Competitiveness	0.410	0.311	0.309	0.268
5 = Highly competitive	(0.63)	(0.58)	(0.58)	(0.58)
State of personal finances	0.523*	0.474*	0.475*	0.436*
5 = Very good	(0.31)	(0.27)	(0.27)	(0.26)
Gaming frequency	0.293	0.268	0.271	0.260
6 = More than twice a day	(0.22)	(0.23)	(0.24)	(0.24)
Ability in task	0.027	-0.005	-0.007	0.003
	(0.15)	(0.15)	(0.15)	(0.15)
Constant	-6.496**	-4.045	-4.172*	-4.979*
	(3.05)	(2.71)	(2.38)	(2.67)
Observations	189	189	189	189
Log likelihood	-411.4	-409.5	-409.5	-408.8
Left censors	0	0	0	0
Right censors	54	54	54	54

Notes: Tobit regressions with censoring at 13 (maximum number of times subjects could continue the task). The dependent variable is the number of rounds subjects continued the coding task. * 10%, ** 5%, *** 1% significance level. Robust standard errors with sessions-level clusters in parentheses.

Overall, gamification has heterogeneous effects on effort. Initiatives to insert gaming elements into employment or other relationships should take into account that they are less effective among less motivated individuals, and when extrinsic incentives are high.

One additional point to note is that task motivation is measured through the exit survey questions regarding task enjoyment. One might be concerned that subject evaluations of task enjoyment were influenced by the treatments to which they were exposed. We address this concern in two ways. First, our measures of task enjoyment are based on questions that clearly and repeatedly refer to the decoding of words, cuing respondents to ignore other elements of the experimental experience, including the nature of the feedback they received.

Second, we can reject the hypothesis that the treatments systematically affect task motivation. We regress subjects' task motivation on all treatments, comparing subject task motivation in each treatment group with the motivation of subjects in the baseline. Table 6 reports the results. In only one of the 7 treatment arms (piece rate plus combined gaming elements) is task motivation significantly greater than in the baseline ($p < .10$).

Table 6: Treatment effects on Task Motivation

Dependent variable: Intrinsic (Task) Motivation	
Treatment: Gaming elements (all)	0.269
	(0.21)
Treatment: Points	0.096
	(0.15)
Treatment: Badges	0.290
	(0.18)
Treatment: Leaderboards	0.038
	(0.16)
Treatment: Stories	0.085
	(0.18)
Treatment: Piece rate	0.210
	(0.24)
Treatment: Piece rate + gaming elements	0.317*
	(0.16)
Constant	2.894***
	(0.15)
Observations	362
R-Squared	0.019
P	0.001

Finally, to further demonstrate that the effects of task motivation are not a spurious consequence of exposure to the gaming treatments, we estimate a new measure of task motivation, using the residual of the regression reported in Table 6. This residual is orthogonal to, and therefore independent of, the treatments. We substitute this measure for the raw measure of task enjoyment that is employed in the specifications in Table 5. The results regarding the effects of task motivation on the impact of gamification are unchanged (see Appendix Table A.1).

Which gaming elements matter?

The four gaming elements exhibit substantive differences that might affect their influence on individuals' behavior. The point system we use, common in gaming, attempts to offset the increasing cost of effort by allowing points to accumulate faster as individuals decode more words in a round. Badges introduce goal-setting, a fundamental feature of traditional performance feedback, combined with symbolic rewards that vary with higher levels of achievement. Leaderboards introduce relative performance information, another feature of traditional performance feedback, but the information is anonymous: individuals know only their place on the leaderboard. Narrative context introduces a second output of subject effort: not only does effort lead to the decoding of more words, it also completes a story. However, only the decoded words are the focus of the contractual relationship. The second output – the story – has no "commercial" value, and should therefore motivate effort only to the degree to which subjects enjoy completing it.

Just as we expect successful games to incorporate multiple gaming elements, we expect the combined effect of all gaming elements to have the largest effect on subject effort, compared to the baseline of basic performance feedback. In addition, not all gaming elements are alike in their capacity to elicit additional effort, above and beyond that triggered by basic performance feedback. Points, for example, can be viewed as a repackaging of the number of words that subjects decode, information that they receive in the baseline treatment. We might therefore expect narrative context (which introduces an entirely new output), leaderboards (which introduce relative performance), and badges (which introduce goal-setting and symbolic rewards) to have a greater impact on effort than points.

These considerations are relevant, too, when thinking about the interactive effects of task motivation and gaming elements. First, since the effects of task motivation and gaming elements are expected to be multiplicative, neither should have a significant effect when the other is low, precisely what we observe when examining the combined effect of gaming elements. Second, to the extent that a gaming element shifts attention away from the production of decoded words towards other objectives, such as the completion of a story, motivation to decode words should matter less.

We estimated the individual effects of each of the four gaming elements in the setting with low extrinsic motivation (flat salary). We can therefore revisit two conclusions from the earlier analysis: that the effect of gaming elements has a significant effect on effort when extrinsic motivation is low (flat salary); and that task motivation significantly increases the effect of gaming elements. Is it only the combination of gaming elements that drives these effects, or do individual elements also significantly affect behavior?

Figure 5 displays the average effort levels provided by subjects for each gaming element implemented. The gaming element that manipulates feedback the least (points) also has no significant effect on effort (p=0.42). Goal-setting and symbolic rewards (badges) have a larger effect, but still not significant (p=0.15). Leaderboards, which do little more than introduce relative performance measurement, as in the traditional feedback literature, have a significant effect (p<0.10). However, the gaming element with the largest individual effect on effort is the one least connected to traditional performance feedback, narrative context (p<0.10).

Subjects repeat the baseline task (with a simple counter on the number of words decoded correctly) 0.88 times on average. In the Points treatment (with 50 points per word decoded correctly and a multiplier starting at 1 and incremented by 0.1 for every word decoded correctly in a row), subject repeat the task 1.12 times, which is higher but not significantly different from the baseline (p=0.42). When subjects are provided with pre-specified goals, and receive badges for achieving goals, effort increases to 1.36, which is higher, but not significantly different from the baseline (p=0.15). Adding leaderboards (subjects are ranked against a leaderboard using data

from pilot sessions) increases effort to 1.38 repetitions, which is a significant increase over the baseline (p<0.06). However, introducing context (decoded words complete a short story) has the largest increase in effort, to 1.70 repetitions, almost double the number of rounds played by subjects in the baseline (p<0.06).

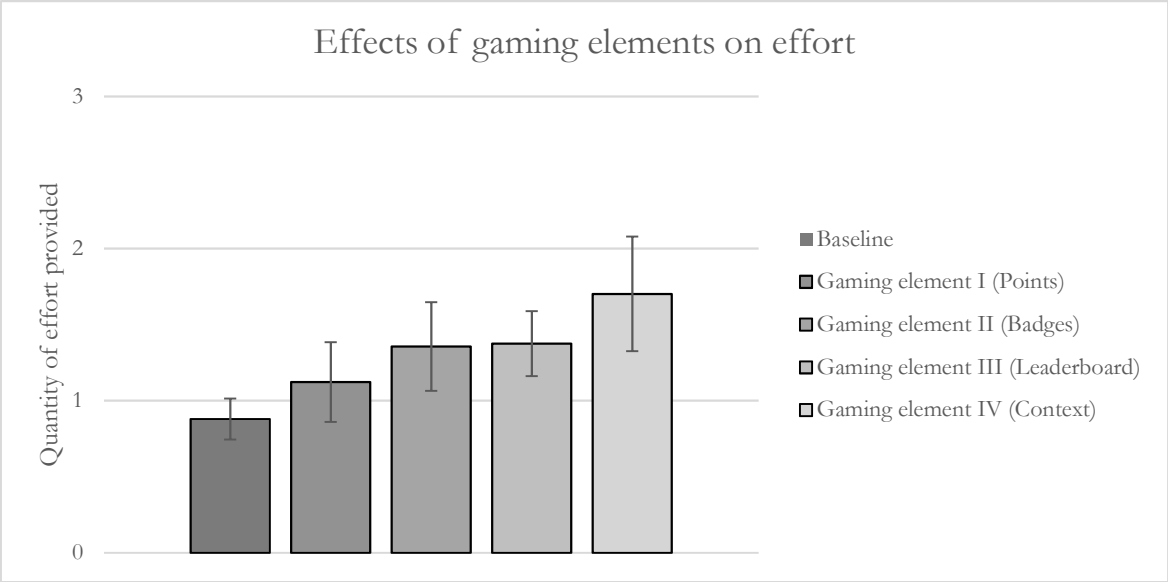


Figure 5: Effects of each gaming element on effort

Figure 5 does not address the effects of task motivation, nor does it control for subject characteristics. Table 7 presents results of tobit specifications that responds to both issues. All models include treatment dummies, our measure of intrinsic motivation, and their interaction. In addition, the models include controls for gender, age, risk and competitive preferences, personal finances, gaming frequency, and task ability.

In model 1, we compare effort exerted in the Points treatment, relative to the baseline. Model 2 focuses on Badges, models 3 and 4 on Leaderboards, and model 5 on Context. As mentioned, each of the treatments implement one of the gaming elements independently, and so each treatment is compared to the baseline. Model 4 (Leaderboards) adds an interaction term for treatment and competitive preferences, as we would expect competitive individuals to engage in greater effort in the leaderboard treatment.

The results in Table 7 are consistent with the comparisons in the figure: neither points nor badges, individually, have a significant effect on effort, but both Leaderboards and narrative context do. We find strong evidence for the importance of task motivation across all models: subjects who enjoy decoding words spend more time on this task. However, as we expect, where the gaming element has no significant effect on effort, task motivation does not help: for Points and Badges, there are no significant differences in effort compared to the baseline group, either in the main treatment effect (p=0.53 and p=0.77 respectively) or in the motivation and treatment interaction term (p=0.58 and p=0.70 respectively).

Model 3 estimates the effort effects of competition with peers using a leaderboard. Subjects compete with others (from previous sessions) for a high rank on the leaderboard (based on the number of total words decoded per round). Model 4 adds one additional interaction term between the treatment and preferences for competition. Individuals who exhibit low competitiveness and are not task motivated exert significantly *less* effort in the Leaderboard treatment, as many as 3 fewer rounds, three times larger than the average number of rounds in the baseline (one). The possibility that gaming elements could reduce effort is acknowledged in

the gamification literature (Hanus and Fox, 2015; Domínguez et al., 2013; Fitz-Walter, Tjondronegoro, and Wyeth, 2011; Montola et al., 2009) and finds some evidence here, among unmotivated and competition-averse individuals. Subjects who are more competitive exert significantly more effort in the Leaderboard treatment relative to baseline ($p<0.01$). So also do more task-motivated subjects ($p<0.10$). As before, selection matters: competition works well for those who have strong preferences for competition, but not for others (Niederle and Vesterlund, 2007; Kleinjans, 2009; Garrett et al. 2013; among others).

Model 5 focuses on our final gaming element: Narrative context. Regardless of intrinsic motivation, subjects exert significantly more effort when exposed to the context treatment ($p<0.05$), and the magnitude of the effect is large (the unmotivated exposed to the context treatment play almost two additional rounds compared to the unmotivated in the baseline, who play less than one). However, consistent with our conjecture that the treatment motivates subjects by allowing them to produce an output that is unrelated to the number of words that they decode, the interaction with task motivation is insignificant.

Table 7: Gaming elements and effort

Dependent Variable: Quantity of effort provided (Rounds continued)					
	I	II	III	IV	V
	Gaming element I (Points)	Gaming element II (Badges)	Gaming element III (Leaderboards)	Gaming element III (Leaderboards)	Gaming element IV (Context)
Treatment	0.621	0.121	-0.493	-3.053***	1.980**
	(0.98)	(0.41)	(0.56)	(0.76)	(0.84)
Task motivation	0.385***	0.329***	0.288***	0.294***	0.338**
	(0.12)	(0.11)	(0.10)	(0.10)	(0.16)
Task motivation and Treatment interaction	-0.126	0.069	0.344**	0.323*	-0.389
	(0.23)	(0.18)	(0.14)	(0.16)	(0.26)
Competitiveness	0.154	0.239	0.001	-0.255	0.303
5 = Highly competitive	(0.17)	(0.34)	(0.28)	(0.23)	(0.47)
Competitiveness and Treatment interaction				0.788***	
				(0.25)	
Gender	0.089	-0.189	-0.213	-0.300	-0.510
1 = Female	(0.36)	(0.29)	(0.65)	(0.63)	(0.51)
Age (in years)	0.039***	0.031***	0.040***	0.041***	-0.002
	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)
Risk preferences	-0.183	-0.281	-0.096	-0.113	-0.287
5 = Risk-seeking	(0.14)	(0.19)	(0.06)	(0.08)	(0.22)
State of personal finances	-0.068	0.212	0.263	0.302	0.095
5 = Very good	(0.17)	(0.18)	(0.20)	(0.18)	(0.09)
Gaming frequency	0.039	-0.159	0.068	0.043	-0.393
6 = More than twice a day	(0.08)	(0.11)	(0.16)	(0.16)	(0.24)
Ability in task	0.007	-0.033	0.009	0.001	0.031
	(0.03)	(0.05)	(0.04)	(0.04)	(0.02)
Constant	-1.017	-0.456	-1.458	-0.517	0.624
	(0.93)	(0.99)	(1.32)	(0.98)	(2.05)
Observations	84	88	83	83	90
Log likelihood	-138.8	-156.3	-117.5	-116.0	-186.6

Notes: Tobit regressions with censoring at 13 (maximum number of times subjects could continue the task). The dependent variable is the number of rounds subjects continued the coding task. * 10%, ** 5%, *** 1% significance level. Robust standard errors with sessions-level clusters in parentheses.

The largest effects of gamification on effort, relative to basic performance feedback, arise when simple feedback is supplemented with multiple gaming elements. The elements that individually have the largest effects on behavior are linked to social competition (Leaderboards) and context (Stories). In looking at the combined effect of gaming elements, selection mattered: the task motivated exerted substantially more effort. Selection also matters strongly for Leaderboards (gaming elements that exploit social competition): task-motivated subjects with high competitive preferences responded much more to the treatment, while the unmotivated without competitive preferences worked less. In contrast, narrative context increases subject welfare even if the decoding of words has no intrinsic attraction for them and even if they do not have competitive preferences. Consistent with this, selection plays no role in the effects of context: the interaction of task motivation and context is insignificant.

Conclusions

In this paper, we present a simple experiment designed to address ambiguities in the literature on the effects of feedback (gamification) by shedding light on the motivational mechanisms underlying the process. The experiment starts with a simple real effort task with low feedback (without gaming elements) and low extrinsic incentives. We then manipulate the sophistication of feedback, by adding gaming elements, and of extrinsic incentives, by introducing piece rate compensation. In line with theory and intuition, enhanced feedback increases effort, but only when extrinsic incentives are low. Furthermore, we shed light on the mechanism underlying the increase: those subjects with higher task motivation respond to feedback by increasing their effort. The particular gaming elements used to enhance feedback also matter. We find large effects from inducing competition between subjects (Leaderboards) and adding additional output (Context). Inducing competition operates through competitive preferences (competitive individuals respond by increasing effort) whereas Context appears to be independent of preferences.

The experiment reported in this paper scratches the surface of what enhanced feedback (through gamification) has to offer. We focus on the most common gaming elements, but there are many more to consider (for example, relationships, challenges, constraints, onboarding, scaffolding, progress, emotion, etc. – See Kumar and Herger, 2013 for a full discussion). In addition to this, we also focus on two distinct contract types, common in the literature, but polar opposites with respect to their effects on extrinsic motivation. Future work would consider additional contract types that contain different mixes of fixed and variable components (as in Pokorny, 2008).

Nevertheless, our results have important implications for managers seeking to improve efficiency in operations. Particularly, by demonstrating the importance of task motivation and feedback, along with extrinsic incentives, the results can be used by managers to tailor their selection policies as well as consider task-related enhancements, particularly in situations where effort is not easily observable or contractible. Of course, one major caveat here is that our study uses perfectly observable and unidimensional effort to study these effects. Future work would likely consider adding mission-based motives to the tasks, and multiple tasks to understand the effects under more contextual settings.

References

Ariely, D., Bracha, A. and Meier, S., (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. The American economic review, 99(1), pp.544-555.

Ariely, D., Kamenica, E. and Prelec, D., (2008). Man's search for meaning: The case of Legos. Journal of Economic Behavior & Organization, 67(3), pp.671-677.

Ashraf, N., Bandiera, O. and Jack, B.K., (2014). No margin, no mission? A field experiment on incentives for public service delivery. Journal of Public Economics, 120, pp.1-17.

Azmat, G. and Iriberri, N., 2010. The importance of relative performance feedback information: Evidence from a natural experiment using high school students. Journal of Public Economics, 94(7), pp.435-452.

Bandiera, O., Barankay, I. and Rasul, I., 2007. Incentives for managers and inequality among workers: Evidence from a firm-level experiment. The Quarterly Journal of Economics, 122(2), pp.729-773.

Banuri, S. and Keefer, P. (2016). "Pro-social motivation, effort and the call to public service." European Economic Review 83: 139-164.

Banuri, S., Keefer, P., and Damien de Walque (2017). " Love the job... or the patient? Task vs. mission-based motivations in healthcare." Mimeo.

Barankay, I. "Rank incentives: Evidence from a randomized workplace experiment." Unpublished manuscript (2012).

Blanes i Vidal, J. and Nossol, M., (2011). Tournaments without prizes: Evidence from personnel records. Management Science, 57(10), pp.1721-1736.

Bloom, N. and Van Reenen, J., (2007). Measuring and explaining management practices across firms and countries. The Quarterly Journal of Economics, 122(4), pp.1351-1408.

Bock, O., Nicklisch, A. and Baetge, I., (2012). hroot: Hamburg registration and organization online tool. H-Lab Working Paper (1).

Bradler, C., Dur, R., Neckermann, S., & Non, A. (2016). Employee recognition and performance: A field experiment. Management Science, 62(11), 3085-3099.

Carpenter, J. and Myers, C.K., (2010). Why volunteer? Evidence on the role of altruism, image, and incentives. Journal of Public Economics, 94(11), pp.911-920.

Casas-Arce, P. & Martínez-Jerez, F. A. (2009). Relative performance compensation, contests, and dynamic incentives. Management Science, 55, 1306-1320.

Deci, E. L., Koestner, R. and Ryan, R.M., (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation, American Psychological Association.

Deci, E.L. and Ryan, R.M., (1985). The general causality orientations scale: Self-determination in personality. Journal of research in personality, 19(2), pp.109-134.

Deterding, S., et al. (2011). From game design elements to gamefulness: defining gamification. Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments, ACM.

Domínguez, A., Saenz-de-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. J. (2013). Gamifying learning experiences: Practical implications and outcomes. Computers and Education, 63, 380–392

Dubey, P. and Geanakoplos, J., (2010). Grading exams: 100, 99, 98,… or A, B, C?. Games and Economic Behavior, 69(1), pp.72-94.

Dubey, P. K. & Geanakoplos, J. (2005). Grading in games of status: Marking exams and setting wages. Cowles Foundation Discussion Paper No. 1544, Yale University.

Ederer, F., (2010). Feedback and motivation in dynamic tournaments. Journal of Economics & Management Strategy, 19(3), pp.733-769.

Eriksson, T., Poulsen, A. and Villeval, M.C., 2009. Feedback and incentives: Experimental evidence. Labour Economics, 16(6), pp.679-688.

Erkal, N., Gangadharan, L. and Nikiforakis, N., (2011). Relative earnings and giving in a real-effort experiment. The American Economic Review, 101(7), pp.3330-3348.

Falk, A., Gächter, S. and Kovács, J., (1999). Intrinsic motivation and extrinsic incentives in a repeated game with incomplete contracts. Journal of Economic Psychology, 20(3), pp.251-284.

Fehr, E. and Falk, A., (1999). Wage rigidity in a competitive incomplete contract market. Journal of political Economy, 107(1), pp.106-134.

Fischbacher, U. (2007). "z-Tree: Zurich toolbox for ready-made economic experiments." Experimental economics 10(2): 171-178.

Fitz-Walter, Z., Tjondronegoro, D., & Wyeth, P. (2011). Orientation passport: Using gamification to engage university students. In Proceedings of the 23rd Australian computer-human interaction conference (pp. 122–125). ACM.

Frey, Bruno S., and Felix Oberholzer-Gee. "The cost of price incentives: An empirical analysis of motivation crowding-out." The American Economic Review 87.4 (1997): 746-755.

Friedrichsen, J., and Engelmann, D. (2017). Who Cares about Social Image? (No. 1634). DIW Berlin, German Institute for Economic Research.

Garratt, Rodney J., Catherine Weinberger, and Nick Johnson. (2013). The state street mile: Age and gender differences in competition aversion in the field. Economic Inquiry 51, no. 1: 806-815.

Gosling, S. D., Rentfrow, P.J. and Swann, W.B., (2003). A very brief measure of the Big-Five personality domains. Journal of Research in personality 37(6): 504-528.

Graziano, W. G., et al. (2007). "Agreeableness, empathy, and helping: a person× situation perspective." Journal of personality and social psychology 93(4): 583-599.

Hamari, J. (2013). "Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service." Electronic commerce research and applications 12(4): 236-245.

Hamari, J., et al. (2014). Does gamification work?--a literature review of empirical studies on gamification. System Sciences (HICSS), 2014 47th Hawaii International Conference on, IEEE.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. Computers & Education, 80, 152-161.

Harris, P. B., & Houston, J. M. (2010). A reliability analysis of the revised competitiveness index. Psychological Reports, 106(3), 870-874.

Hattie, J. and Timperley, H., (2007). The power of feedback. Review of educational research, 77(1), pp.81-112.

Herger, M. and J. Kumar (2013). "Gamification at Work—." Designing engaging business Software. Aarhus: Interaction Design Foundation.

Herzberg, F. I. (1966). Work and the nature of man. Oxford, World Publications

Houston, J. M., Harris, P. B., McIntire, S., & Francis, D. (2002) Revising the Competitiveness Index. Psychological Reports, 90, 31-34.

Kleinjans, Kristin J. "Do gender differences in preferences for competition matter for occupational expectations?." Journal of Economic Psychology 30, no. 5 (2009): 701-710.

Kluger, A.N. and DeNisi, A., 1996. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory.

Kosfeld, Michael and Susanne Neckermann (2011). "Getting More Work for Nothing? Symbolic Awards and Worker Performance." American Economic Journal: Microeconomics 3 (August): 86-99.

Kosfeld, M., Neckermann, S. and Yang, X., (2014). Knowing that you matter, matters! The interplay of meaning, monetary incentives, and worker recognition.

Kumar, J., and Herger, M. (2013). "Gamification at work: Designing engaging business software." International Conference of Design, User Experience, and Usability. Springer, Berlin, Heidelberg.

List, J.A. and Rasul, I., (2011). Field experiments in labor economics. Handbook of labor economics, 4, pp.103-228.

Lizzeri, A., Meyer, M.A. and Persico, N., (2002). The incentive effects of interim performance evaluations. University of Pennsylvania, Center for Analytic Research in Economics and the Social Sciences.

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

Mekler, E. D., et al. (2013). Do points, levels and leaderboards harm intrinsic motivation? An empirical analysis of common gamification elements. Proceedings of the First International Conference on gameful design, research, and applications, ACM.

Montola, M., Nummenmaa, T., Lucerano, A., Boberg, M., & Korhonen, H. (2009). Applying game achievement systems to enhance user experience in a photo sharing service. In Proceedings of the 13th international academic mindtrek conference: Everyday life in the Ubiquitous Era (pp. 94–97). Tampere, Finland, September 30–October 2, 2009.

Neitzel, J. and Sääksvuori, L., (2013). Normative Conflict and Cooperation in Sequential Social Dilemmas. Verein für Socialpolitik/German Economic Association.

Niederle, Muriel, and Lise Vesterlund. "Do women shy away from competition? Do men compete too much?." The Quarterly Journal of Economics 122, no. 3 (2007): 1067-1101.

Pokorny, K., (2008). Pay—but do not pay too much: An experimental study on the impact of incentives. Journal of Economic Behavior & Organization, 66(2), pp.251-264.

Porter, L. W. and E. E. Lawler (1968). Managerial attitudes and performance, Homewood, IL: Irwin-Dorsey.

Ryan, R. M. (1982). "Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory." Journal of personality and social psychology 43(3): 450-461.

Ryan, R. M. and E. L. Deci (2000). "Intrinsic and extrinsic motivations: Classic definitions and new directions." Contemporary educational psychology 25(1): 54-67.

Smither, R. D., & Houston, J. M. (1992) The nature of competitiveness: construction and validation of the Competitiveness Index. Educational and Psychological Measurement, 52, 407-418.

Staw, B. M. (1977). "Motivation in organizations: Toward synthesis and redirection." New directions in organizational behavior 1: 54-95.

Appendix A: Additional tables for review

Table A.1: Gaming elements and effort

Dependent Variable: Quantity of effort provided (Rounds continued)				
	I	II	III	IV
Treatment: Gaming elements	0.944***	0.952***	0.954***	0.940***
	(0.25)	(0.30)	(0.30)	(0.25)
Treatment: Piece rate	11.00***	11.00***	11.00***	10.99***
	(0.59)	(0.57)	(0.57)	(0.56)
Interaction: Gaming elements X Piece rate	-1.248	-1.215	-1.218	-1.188
	(1.50)	(1.44)	(1.42)	(1.40)
Task motivation	0.412*	-0.118	-0.077	0.254*
	(0.23)	(0.31)	(0.31)	(0.13)
Interaction: Gaming elements X Task motivation		1.299***	1.294***	0.599***
		(0.45)	(0.45)	(0.16)
Interaction: Piece rate X Task motivation			-0.083	-0.747
			(0.50)	(0.61)
Interaction: Gaming X Piece rate X Task motivation				1.581***
				(0.60)
Gender	1.127	0.944	0.949	0.915
1 = Female	(0.70)	(0.71)	(0.73)	(0.71)
Age (in years)	0.026	0.018	0.019	0.017
	(0.05)	(0.05)	(0.05)	(0.05)
Risk preferences	0.365	0.413	0.413	0.449*
5 = Risk-seeking	(0.29)	(0.28)	(0.28)	(0.27)
Competitiveness	0.410	0.311	0.309	0.268
5 = Highly competitive	(0.63)	(0.58)	(0.58)	(0.58)
State of personal finances	0.523*	0.474*	0.475*	0.436*
5 = Very good	(0.31)	(0.27)	(0.27)	(0.26)
Gaming frequency	0.293	0.268	0.271	0.260
6 = More than twice a day	(0.22)	(0.23)	(0.24)	(0.24)
Ability in task	0.027	-0.005	-0.007	0.003
	(0.15)	(0.15)	(0.15)	(0.15)
Constant	-5.305*	-4.386	-4.395	-4.244
	(2.96)	(2.71)	(2.69)	(2.76)
Observations	189	189	189	189
Log likelihood	-411.4	-409.5	-409.5	-408.8
Left censors	0	0	0	0
Right censors	54	54	54	54

Notes: Tobit regressions with censoring at 13 (maximum number of times subjects could continue the task). The dependent variable is the number of rounds subjects continued the coding task. * 10%, ** 5%, *** 1% significance level. Robust standard errors with sessions-level clusters in parentheses.

Table A.2: Gaming elements and effort (total words coded)

Dependent Variable: Total words coded (Effort)				
	I	II	III	IV
Treatment: Gaming elements	9.449***	7.542**	9.449***	7.203*
	(2.05)	(2.90)	(1.99)	(3.73)
Treatment: Piece rate			98.63***	95.37***
			(2.65)	(3.78)
Interaction: Gaming elements X Piece rate			-11.220	-11.170
			(6.44)	(8.52)
Gender		9.553		5.892
1 = Female		(5.74)		(4.76)
Age (in years)		0.268		0.206
		(0.37)		(0.50)
Risk preferences		3.273		4.414
5 = Risk-seeking		(2.63)		(3.35)
Competitiveness		-3.491		0.940
5 = Highly competitive		(2.13)		(5.62)
State of personal finances		4.139**		5.266
5 = Very good		(1.55)		(3.37)
Gaming frequency		2.411		2.229
6 = More than twice a day		(1.32)		(1.98)
Ability in task		1.962*		3.532**
		(0.81)		(1.40)
Constant	16.51***	-28.240	16.51***	-60.03*
	(1.59)	(22.14)	(1.54)	(31.94)
Observations	94	94	189	189
Log likelihood	-406.8	-396.4	-932.8	-921.8
P	0.004	.	0.000	0.000
R-squared	0.062	0.247	0.656	0.694

Notes: OLS regressions with robust standard errors clustered at session-level (standard errors in parentheses). The dependent variable is the total number of words subjects decoded. * 10%, ** 5%, *** 1% significance level.

Table A.3: Gaming elements and effort (total words coded)

Dependent Variable: Quantity of effort provided (Rounds continued)				
	I	II	III	IV
Treatment: Gaming elements	6.266*	-36.18***	-36.15***	-9.824
	(3.40)	(10.75)	(11.15)	(6.22)
Treatment: Piece rate	94.81***	96.34***	97.16***	119.7***
	(3.63)	(3.44)	(14.40)	(10.73)
Interaction: Gaming elements X Piece rate	-10.690	-12.470	-12.410	-68.71***
	(8.46)	(8.06)	(8.77)	(16.10)
Task motivation	3.519	-1.994	-1.846	2.308*
	(2.92)	(2.83)	(3.73)	(1.23)
Interaction: Gaming elements X Task motivation		13.92***	13.90***	5.173**
		(3.77)	(3.97)	(1.98)
Interaction: Piece rate X Task motivation			-0.274	-7.838*
			(5.08)	(4.06)
Interaction: Gaming X Piece rate X Task motivation				18.10***
				(3.98)
Gender	5.346	3.401	3.413	3.088
1 = Female	(4.92)	(5.49)	(5.56)	(5.56)
Age (in years)	0.166	0.069	0.072	0.036
	(0.49)	(0.49)	(0.48)	(0.49)
Risk preferences	4.550	5.031	5.032	5.385*
5 = Risk-seeking	(3.13)	(2.88)	(2.90)	(2.79)
Competitiveness	0.627	-0.472	-0.473	-0.976
5 = Highly competitive	(5.42)	(5.08)	(5.11)	(5.14)
State of personal finances	4.510	3.920	3.925	3.428
5 = Very good	(2.79)	(2.51)	(2.45)	(2.28)
Gaming frequency	2.007	1.757	1.763	1.673
6 = More than twice a day	(2.08)	(2.25)	(2.27)	(2.27)
Ability in task	3.349**	3.014**	3.007**	3.113**
	(1.43)	(1.27)	(1.31)	(1.33)
Constant	-64.01*	-37.520	-37.990	-47.69*
	(32.43)	(25.92)	(21.61)	(26.42)
Observations	189	189	189	189
Log likelihood	-921.0	-917.7	-917.7	-916.3
P	.	.	.	.
R-squared	0.697	0.707	0.707	0.711

Notes: OLS regressions with robust standard errors clustered at session-level (standard errors in parentheses). The dependent variable is the total number of words subjects decoded. * 10%, ** 5%, *** 1% significance level.

Appendix B: Instructions and screenshots

Table B.1: Instructions for the effort task

Instructions page 1

FINAL PHASE

Thank you for completing Phase 1. This is the second and final phase, after which you will be asked to complete a short survey, and then will be paid your earnings for the session.

In this phase, you will engage in the decoding task again. However, the task is a little different this time.
You will be paid 400 tokens for participating in this phase. This amount is unconditional. This means that you will be paid 400 tokens regardless of your decisions in the final phase.

Once this phase is complete, you will be asked to complete a short survey, and then will be paid and be free to leave.

Please click "continue" to continue the instructions.

Instructions page 2

FINAL PHASE

For the final phase, you will engage in the decoding task again. However, there are a number of differences.

The major difference between this phase and previous ones is that the words are not random letters, but actual words.

The second difference is that you will now engage in the decoding task for 2 minutes (120 seconds) per round.

After the first round, you can choose to end the phase. Once you have completed the first round, you will be given feedback on your performance, and then asked whether you would like to stop the phase or to continue for another round.

Please click "continue" to continue the instructions.

Instructions page 3

FINAL TASK

If you choose to continue, you will be asked to engage in the decoding task for another 2 minutes. You can choose to continue the phase for as many rounds as you like.

If you choose to end the task, you will be asked to start the survey, at the end of which you will be paid for the session and then will be free to leave.

It is important to note that you will be paid 400 tokens no matter how long you engage in this phase, or how many words you decode correctly. You are free to end the phase at the end of any round after the first. As soon as you end the phase, you will be given the survey to complete and then paid and free to leave.

Please click "continue" to continue the instructions.

Instructions page 4

FINAL TASK

If you choose to continue, you will be asked to engage in the decoding task for another 2 minutes. You can choose to continue the phase for as many rounds as you like.

If you choose to end the task, you will be asked to start the survey, at the end of which you will be paid for the session and then will be free to leave.

It is important to note that you will be paid 400 tokens no matter how long you engage in this phase, or how many words you decode correctly. You are free to end the phase at the end of any round after the first. As soon as you end the phase, you will be given the survey to complete and then paid and free to leave.

Please click "continue" to continue the instructions.

Instructions – Context treatment

FINAL TASK

Each round contains a short story with certain key words missing. You task is to decode the words to complete the story.

Each round contains a different short story, which have been taken from different cultures. As you decode the words, the program will list them at the bottom of the screen. At the end of each round, you will get the story with the words you have decoded filled in.

Note that the amount of the story that is revealed depends on the number of words that you decode correctly. The more words you decode, the more of the story you get at the end of the round.

Please click "continue" to continue the instructions.

Table B.2: Instructions for the effort task

Instructions – Points treatment

FINAL TASK

In this round there are a number of ways you will be able to track your performance.

There is a points system that is used to keep track of your performance in each round. For each word you decode correctly, you will earn 50 points.

In addition to this, there is a multiplier that starts at 1, and increases by 0.1 for each word you decode correctly in a row. If you make a mistake, the multiplier will reset to 1. Therefore, the more accurate you are, the higher your points will be.

Please click "continue" to continue the instructions.

Instructions – Leaderboard treatment

FINAL TASK

Other students at UEA have participated in similar sessions earlier. They have been ranked according to their highest score in this phase.

Based on the number of words you decode correctly in this phase, you will be assigned a rank and place on the leaderboard. Note that only the highest score you achieve will be used to rank you. No identifying information will be used: You will be anonymously added to the leaderboard at the end of this session.

Also note that others in the session will be added to the leaderboard as well, but only after this session. Therefore, the ranking information you receive will not include other players in the session today.

Please click "continue" to continue the instructions.

Instructions – Quiz – Gamification treatment

1. This phase will continue for four rounds. ○ True ○ False

2. I can choose to end the phase at the end of each round. ○ True ○ False

3. There is a maximum of 20 rounds in this phase. ○ True ○ False

4. Ending the phase means that I will be asked to complete the final survey and then will be free to leave. ○ True ○ False

5. I can continue this phase for as many rounds as I wish. ○ True ○ False

6. I will be paid based on the number of words decoded correctly. ○ True ○ False

7. Each round contains a short story with key words that I have to decode. ○ True ○ False

8. For each word I decode correctly, I will earn points. ○ True ○ False

9. The points multiplier depends on how many words I decode correctly in a row. ○ True ○ False

10. My highest score will be anonymously added to the leaderboard. ○ True ○ False

11. If I code 19 words or more, I will earn the titanium medal. ○ True ○ False

Table B.3: Effort screenshots

Baseline (flat salary) treatment

Gamification (flat salary) treatment

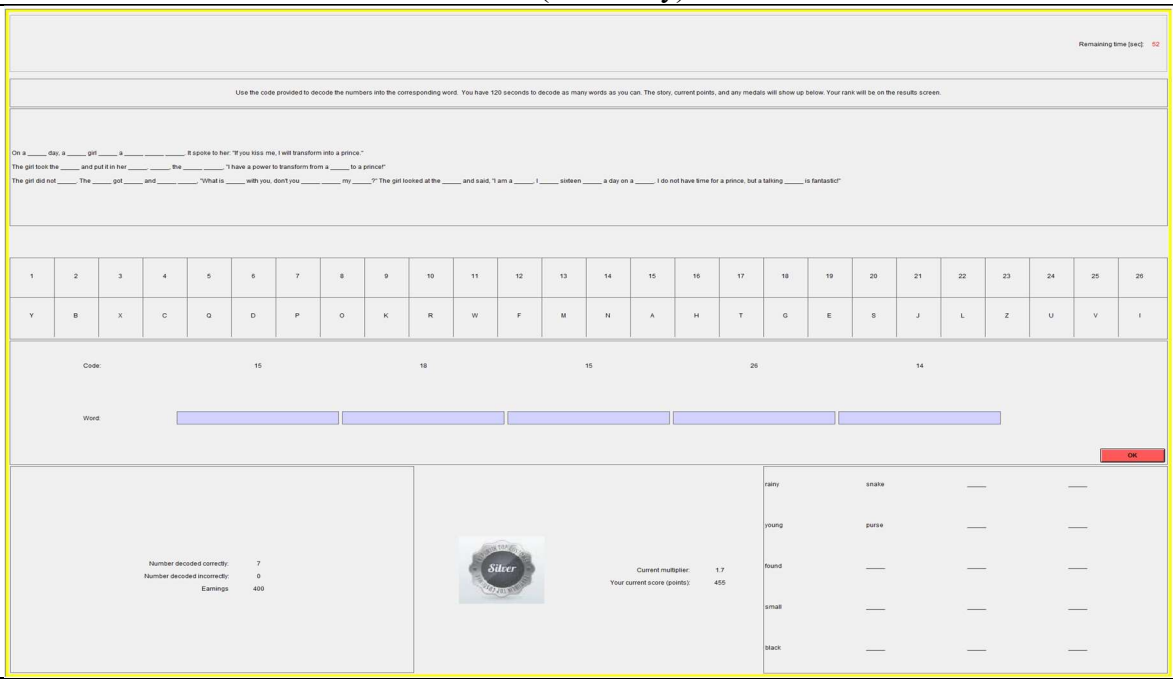
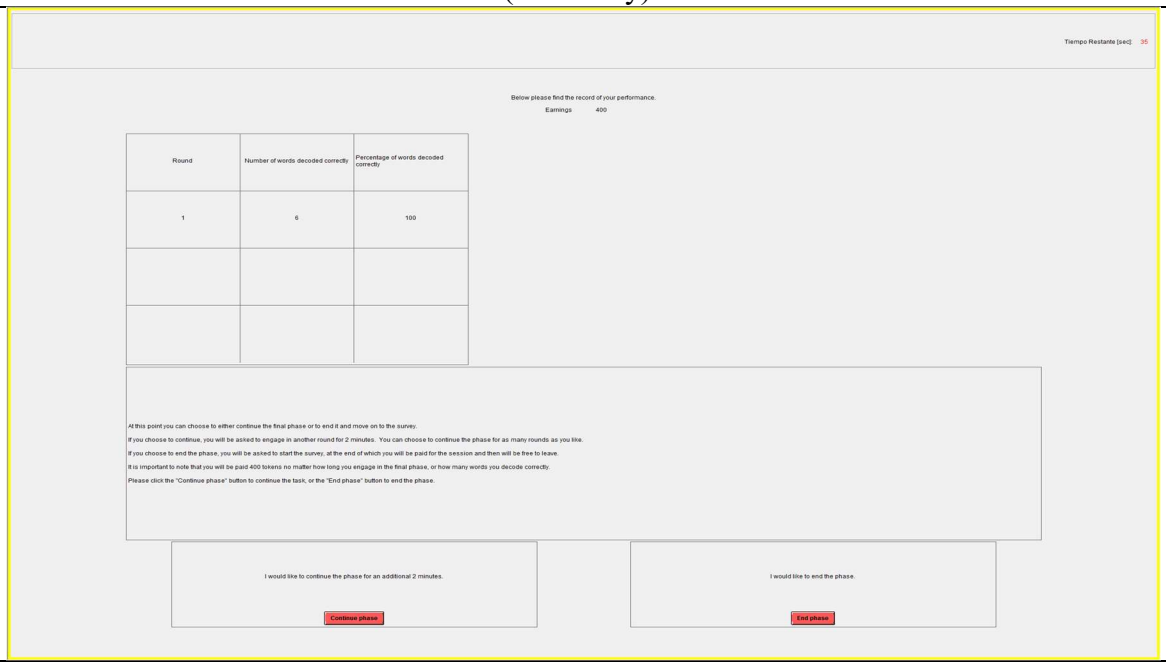
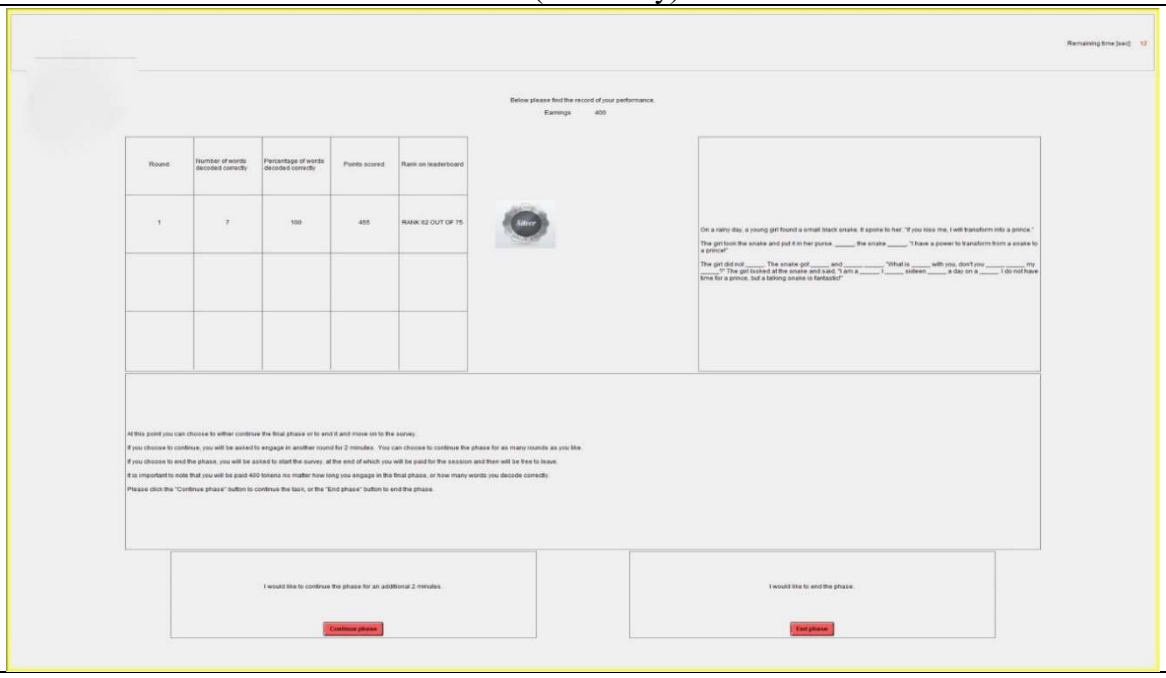


Table B.4: Decision and feedback screenshots

Baseline (flat salary) treatment


Gamification (flat salary) treatment


Appendix C: Intrinsic Motivation Inventory survey questions

Intrinsic Motivation Inventory – Task Evaluation

* Note: highlighted questions were used to measure intrinsic motivations (task evaluation/enjoyment)

1="Strongly disagree"; 2="Disagree"; 3="Neutral"; 4="Agree"; 5="Strongly agree"

1. While I was working on the decoding task I was thinking about how much I enjoyed it.
2. I did not feel at all nervous about doing the decoding task.
3. I think I am pretty good at the decoding task.
4. I found the decoding task very interesting.
5. I felt tense while doing the decoding task.
6. I think I did pretty well compared to other students.
7. Doing the decoding task was fun.
8. I felt relaxed while doing the decoding task.
9. I enjoyed doing the decoding task very much.
10. I am satisfied with my performance at the decoding task.
11. I was anxious while doing the decoding task.
12. I thought the task was very boring. (r)
13. I felt pretty skilled at this task.
14. I thought the decoding task was very interesting.
15. I felt pressured while doing the decoding task.
16. I would describe the task as very enjoyable.
17. After working at this task for a while, I felt pretty competent.