Author:
Wang, Guanlin

Title:
Incentive systems and control in a gamified era

Incentive Systems and Control in a Gamified Era

Guanlin Wang

A dissertation submitted to the University of Bristol in accordance with the requirements for the award of the degree of PhD Accounting and Finance in the Faculty of Social Sciences and Law

Bristol Business School

Department of Accounting and Finance

Sep 2023

Word Count: 41,987

Abstract

Incentives and control systems are one of the most important research areas for management accounting. Mobilising classic psychological theories, this thesis sheds light on the possibility and potential of gamification as an approach to designing incentives and control systems. On the one hand, this thesis raises the attention of management accounting scholars to further explore the motivational effects of gamification. On the other hand, it also sparks designers of gamification in organisations to consider their current use of gamification to incentivise their employees or users. In the first chapter, a netnographic study is conducted in an online on-demand delivery platform to explain how the gamified incentive system motivates delivery riders to perform and how the gamified incentive system helps create a sense of community among riders, as an incentive factor, to further motivate and retain riders on this platform. In the second chapter, in the same field setting, a survey study is conducted to examine the moderation effect of social comparison orientation and the mediation effect of occupational self-efficacy from the relative performance information in riders' gamified leaderboards. In the third chapter, another netnographic study is conducted in an online learning platform to explain how the platform uses gamification to change users' behaviour through the internalisation of external gamified elements and how these internalisations further complement the intrinsic motivation to satisfy the users' three basic psychological needs autonomy, competence, and relatedness.

Dedication and Acknowledgements

First of all, I want to express my deepest thanks to my supervisors Professor Chris Chapman and Dr. Mariano Scapin. Their invaluable guidance and support are significant for me to pursue my interests in conducting these studies. What I learned from them will keep influencing my whole academic career.

I am grateful to the Department of Accounting and Finance at the Business School of the University of Bristol for their generous financial support to enable my PhD life in Bristol and to attend various workshops and conferences during these four years.

Special thanks to Dr. Emma Carroll, for her great efforts in reviewing my chapters, as well as my PhD cohorts and other scholars for their insightful comments.

I appreciate the participation of all interviewees in my field studies and the data collection work of the survey company.

Last but not least, heartfelt thanks to my grandparents, parents Xinben Wang and Shumei Yang, twin brother Zhonglin Wang, and other cherished relatives and friends for their unwavering support throughout my life.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ███████████... DATE: …………..22/09/2023.............................

Table of Contents

5

List of tables and figures

Tables

Chapter 1:

Chapter 2:

Chapter 3:

Figures

Chapter 1:

Chapter 2:

Chapter 3:

1. Introduction

1.1 Background and Motivation

Incentives and control systems are one of the most important research areas for management accounting. In Otley's (1999) management control systems research framework, the fourth question that needs to be answered by any management control system is what rewards are given to managers or employees for achieving certain tasks. Merchant and Van der Stede (2017) argue that incentive systems are significant for organisations to motivate employees to achieve their targets. This is achieved by tying rewards and penalties to performance evaluation. A proper incentive system needs to provide benefits related to management control: first, incentive systems should provide informational benefits, namely, guidance to employees on how to work for the company. Second, it should provide motivational benefits, to push employees to work hard to achieve or exceed certain targets. Lastly, benefits in terms of attraction and retention of personnel are also crucial as an incentive system plays an important role in selecting suitable talent for the right position. These three benefits are fundamental for any organisation trying to stay alive and prosper, particularly in the post-pandemic period when many taken-for-granted priorities and ways of working are being reconsidered.

In this thesis, I focus on a novel way of designing incentive and control systems, namely, gamification. Gamification is defined as the use of game design elements in non-game contexts (Deterding et al., 2011). Some important aspects of this definition need to be highlighted (Sailer et al., 2017). Firstly, gamification is different from game. Gamification uses important building blocks from a complete game such as badges, points, and leaderboards in different real-world contexts for non-entertainment purposes such as retaining customers of certain platforms, motivating gig economy workers, or helping with hybrid education etc. The users of a gamified application retain the flexibility to engage with certain elements they feel intriguing. However, according to Deterding et al. (2011), the boundary between gamification and game can be blurry. This boundary is personal, social and more subjective. Sometimes how the group of people use the gamified elements might change them to a complete game. Meanwhile, gamification is also different from play. These game elements still have certain rules to engage with. Users have to follow these rules to use certain functions. As a result, in some circumstances, gamification might not necessarily bring strong playfulness or job-based emotion to its' users such as gig economy workers or blue-collar workers in

factories.

Secondly, "design" in this definition means that gamification does not rely on game technologies such as game engines and controllers but focuses on the game logic. This is important for gamification to be used in different contexts without the limitations of game-based technology, which is the third important aspect of this definition. Gamification provides a broad suitability to use in different contexts as a control practice. Sometimes, it blurs the boundaries between work and life. According to Dale (2014), from traditional conglomerates such as Nike and Volkswagen to new platform companies such as TripAdvisor and Foursquare, different types of organisations realised the usefulness of gamification nowadays. Relying on the use of big data and digital technology, platform companies are especially inclined to use gamification in their control systems. According to Chapman, Chua and Fiedler (2021), Foursquare uses gamification as a control practice to incentivise its users' behaviours to check-in the places they went and share their experiences in the community. Jeacle and Carter (2011) also show that TripAdvisor uses symbolic designs such as leaderboards and ratings to build trust from travellers around the world.

Compared with traditional management control systems, gamification, as a novel control practice, to some extent, seduces users to behave in a certain way through these deliberate designs. Meanwhile, these gamified design elements include more than one traditional incentive mechanism such as goal setting and comparison in their systems, which gave management accounting researchers to investigate the complexity of the configuration of different incentives. Moreover, not only does gamification provide control and motivational effects from these traditional mechanisms, but also it brings individuals' emotions and community effects into play (Chapman et al., 2021). Sometimes this community effect might not be the intention of the original designs from gamification designers. However, it emerges from the gamification because of the flexibility and giving space retained by users. In turn, it might further contribute to the original control purposes of the gamified control system.

Nonetheless, there is limited evidence in the management accounting literature (Chapman, Chua and Fiedler, 2021; Kelly, Valtchanov and Webb, 2021) explicitly looking into this topic. This drives my motivation to explore gamification in this thesis. During the pandemic, national governments implemented strict lockdowns. However,

several gig workers, such as food-delivery platforms' riders continued to perform their work despite the unknown risks associated with the pandemic. This brought to my attention how these riders are motivated and incentivised by these platforms. In line with this, some anecdotal evidence and recent studies show that riders are experiencing harsh control systems driven by big data and algorithms of platforms (Sun, 2019; Sun and Chen, 2021, Yang, 2021). According to Schwellnus et al. (2019), about 1 to 3 percent of total employment in developed economies can be classified as a gig economy platform employment, and this proportion continues to increase. As one type of gig economy platform, on-demand delivery platforms make a significant contribution to this growth. Since some food delivery platforms have been quite developed for many years and delivery riders become a popular gig economy job in China, I decided to delve into one of the most successful Chinese on-demand delivery platforms.

To have a deep understanding of this setting, I searched this platform on different Chinese social media for relevant information and found that an incentive system with gamified elements such as gamified levels, leaderboards and challenges played an important role in incentivising riders. Therefore, in order to explore the increasing role of gamification in the gig economy platform, I conducted two studies in one of the biggest on-demand food delivery platforms in China to discover how this gamified incentive system of delivery riders operates and affects employees' performance. One is a field study paper, and the other is a survey paper. This is the focus of the first two chapters of my dissertation.

In the third chapter, I use an online learning platform setting to explore how gamification influences users' extrinsic and intrinsic motivations. The reason why I chose this setting is that the implementation of gamification in recent years significantly affected my behaviour as a user of such platforms. This is also relevant to incentive and control systems in a workplace setting to understand how employees' behaviour is affected by gamified designs.

1.2 Netnography and Survey Methods

Two research methods are used in this thesis, netnography and survey methods. Netnography is defined as *"a detailed and specific approach to conducting qualitative research using social media as the basis of its data"* (Kozinets, 2019, p.5). In the first chapter, I use an investigative approach to search and understand short videos and

relevant information about the incentive system used by riders of an online delivery platform in popular Chinese social media. Considering the ethical issues of directly using these social media data, I recruit delivery rider participants from their social media accounts and conduct online semi-structured interviews with these riders around China to collect the main part of the data used in this paper. Video calls for most of the interviews provide me with an adequate level of understanding of riders' working environments, as well as a first-hand experience of the personal circumstances of the riders being interviewed.

In the third chapter, as a loyal user of this platform, I take an immersive approach of netnography to reflect on my own users' behaviours by taking some immersion notes. Meanwhile, I conduct online and in-person interviews with users of this platform. The deep level of engagement in the platform is quite helpful in getting access to users and having conversations with participants. Based on the consent from participants, I also collect data regarding performance learning to triangulate my findings from interviews.

These two chapters further explore the possibilities of netnography (Kozinets, 2019) in management accounting literature (Jeacle and Carter, 2011; Bialecki, O'Leary and Smith, 2017; Van den Bussche and Dambrin, 2020; Chapman, Chua and Fiedler, 2021; Yu, 2021) by exploring and immersing in the setting, recruiting interview participants and triangulating the findings from social media or public platforms. With the popularity of social media and public information on these social media, accounting scholars can make good use of these data in future research.

In the second chapter, I use a survey questionnaire to collect data from delivery riders. Following the predictive validity framework by Libby (2002), I develop a theoretical model and the associate questionnaire. To improve the order of the questions, and to prevent the potential demanding effect, I did a pilot test with some riders before formally collecting the data. Considering the opinions and comments from riders, I revised some questions and items and collected the survey data using the revised questionnaire. I used structural equation modelling to test my hypotheses.

1.3 Theoretical Context

In this thesis, I focus on different psychological theories across the three chapters. In the next paragraphs, I provide a brief introduction to these different theories.

First, I introduce social comparison theory, which informs all three chapters. Then, I introduce other main theories that relate to individual papers: sense of community, goal-setting theory, self-efficacy theory and self-determination theory.

Social comparison theory (Festinger, 1954) argues that people compare with others, especially with a similar individual to know themselves better and form a self-evaluation. There are two main effects of social comparison. First, there is a contrast effect where people feel good when they outperform others, which mainly serves people's self-evaluation and self-enhancement needs; second, there is an assimilative effect, where people develop similarity with outperformers to improve themselves, which serves people's self-improvement needs (Buunk and Gibbons, 2007). Smith (2000) also suggests that the different types of social comparison generate different emotions for people who engage in a social comparison.

As a popular theory used in relative performance information (thereafter RPI) literature, in a series of lab experiments, Tafkov (2013) finds that the provision of public RPI in a single-task context generates a positive effect on performance. Hannan, Towry and Zhang (2013) confirm this finding in a multi-task context. In a different field setting, Carroll and Marginson (2021) show how managers use RPI to conduct social comparisons and have motivational effects. Using a field experiment, Kohler et al. (2023) suggest that detailed RPI might have a risk to direct employees' attention to smaller benefits at the expense of the overall local performance.

In this thesis, social comparison theory serves as one building block in the first chapter to explain how gamified leaderboards influence riders' behaviour and performance. As the main theory for the second chapter, the individual orientation of social comparison as a personality trait is measured and tested for its moderation effect on riders' performance. Social comparison theory is also used in the third chapter to illustrate users' group behaviours on the online learning platform.

In terms of other theories applied in the first chapter, McMillan and Chavis (1986, p9) define a sense of community as *"a feeling that members have of belonging, a feeling that members matter to one another and to the group, and a shared faith that members' needs will be met through their commitment to be together"*. This definition includes four parts of a sense of community: Membership, Influence, Integration and

fulfilment of needs, and Shared emotional connection.

For membership, it means that to have a sense of community people need to be a member of certain groups and have feelings of belonging. They will have emotional safety to express their vulnerability in their group. For influence, it is a bidirectional process that, as a member of a group, people are influenced by other members of the group and, at the same time, they also influence other members. Although members understand the differences of each individual, there will the group conformity to strengthen the cohesiveness of the group. For Integration and fulfilment of needs, people need to fulfil their needs or get rewards from groups to maintain the positive feelings of being together. A group with a strong sense of community needs to fit every member and let them feel the shared value. For shared emotional connection, it suggested that the members of the community need to be involved in some high-quality interactions with other members such as experiencing a success or crisis together, which will further enhance the bond among members emotionally. (McMillian and Chavis, 1986)

Previous accounting studies refer to the sense of community to explain the connection among the users of online platform communities (McDaid, Boedker and Free, 2019; Van den Bussche and Morales, 2019; Van den Bussche and Dambrin, 2020; Chapman, Chua and Fiedler, 2021). Similarly, Baxter et al. (2019) explore the sense of community among football fans in a Swedish football club. I extend these works elaborating on the definition of sense of community and frame sense of community as an important incentive factor for delivery riders of online platforms to motivate them to keep working in this industry.

Serving as the other building block in the first chapter, I introduce goal-setting theory to explain how specific goals play a role in the riders' incentive system. Latham and Locke (1991) argue that people's goals are a primary determinant of their effort. Providing a goal is an important mechanism to incentivise people to work. However, some factors are likely to mediate the relationship between goals and performance (Birnberg, Luft and Shields, 2006). First, the goal difficulty has a positive relationship with performance until individuals reach their ability limit and the commitment to a goal is exhausted. Second, to have a stable performance, the goal should be specific, which can provide a clear direction for people's efforts. Lastly, providing performance

14

feedback during the process is an important element to let people assess the effort or resources they need to achieve the goals.

In Bonner and Sprinkle's (2002) framework, assigned goals are one of the significant environmental moderators between people's effort and performance. Webb, Williamson and Zhang (2013) also find that overall productivity is higher when target-based pay is used in their experiment. In a recent study, Berger, Guo, and Presslee (2023) find that the motivational benefits from the goal-based prosocial reward compared with cash rewards. As the other important mechanism for delivery riders' gamified incentive system, goal-setting theory is used to discuss the effects of gamified challenges and gamified rider-level design.

In the second chapter, related to social comparison theory, I also use self-efficacy theory to motivate my third hypothesis. Self-efficacy is defined as people's belief in their capability to finish certain tasks and achieve certain objectives (Bandura, 1997). As a cognitive factor affecting people's effort and performance, self-efficacy derives mainly from three channels: performance accomplishment, verbal persuasion and vicarious learning. I explored vicarious learning in the second chapter of my thesis. Related to our RPI from riders' leaderboards, riders could increase their self-efficacy in their delivery work by observing their peers' successful performance information and learning from outperformers' experiences.

Finally, the main theory for the third chapter is self-determination theory. Self-determination theory explores people's basic psychological needs and motivations (Deci and Ryan, 2000). Similarly, to the people's physical needs for safety and health, three psychological needs should also be satisfied. They are autonomy (people can self-regulate and choose their behaviours), competence (people can feel they are capable of what they choose to do), and relatedness (people feel that they are connected and accepted by others). Deci and Ryan (2000) also argue that individuals have two types of motivations: intrinsic motivation (defined as the motivation purely from the enjoyment of doing certain activities instead of external rewards or controls) and extrinsic motivation (arising from external consequences, rules, or rewards).

In previous accounting literature, Pfister and Lukka (2018) use self-determination theory to discuss how employees internalise stretched goals as an internal part of their values and beliefs through the design of management control systems. Also,

in a field study, Bouten and Hoozee (2022) show that employees internalise environmental objectives through the design of need-supportive controls. Some survey studies also use self-determination theory as a theoretical foundation. In a recent study, Van der Hauwaert et al. (2022) find a mediation effect of autonomous work motivation between enabling performance measurement systems and managerial performance by satisfying three aforementioned basic psychological needs of managers. Based on this previous literature, the third paper further mobilises self-determination theory to analyse how external gamified motivations are internalised and enhance the intrinsic motivation of learning in an online learning platform.

1.4 Main Findings

The first chapter analyses how a gamified incentive system motivates delivery riders' work. Using goal-setting theory and social comparison theory, the study finds that the gamified challenges with bonuses motivate riders to keep chasing assigned goals by the platform and setting their own goals to achieve. Meanwhile, the gamified leaderboard triggered social comparisons between riders and generated motivational effects. More importantly, based on this gamified incentive system, riders generate a sense of community among riders, which help riders digest their negative emotions from this delivery work and motivate them to keep working as a rider by getting support from other members of their community.

The first chapter contributes to previous platform literature analysing how gamified incentive systems help riders generate SoC, which influences riders' behaviour and performance (Kornberger, Pflueger and Mouritsen, 2017; Van den Bussche and Dambrin, 2020; Chapman, Chua and Fiedler, 2021; McDaid, Andon, and Free, 2023). This chapter also calls for experimental and field research on the motivational effect of SoC as a significant incentive factor. This chapter also responds to Repenning, Löhlein and Schäffer's (2021) call for accounting research on the role of emotions in the digital platform settings to illustrate the change of emotions of riders on this food delivery platform. Moreover, adding to previous goal-setting and RPI literature (Eyring, Ferguson and Koppers, 2021; Hannan, Krishnan and Newman, 2008; Hartmann and Schreck, 2016; Tafkov, 2013), I analyse goal-setting and social comparison mechanisms in a gamified incentive system to motivate gig economy workers and respond to the questions by Carroll and Marginson (2021) on how the

effect of social comparison is affected by the number of comparers in gamified leaderboards.

The second chapter of my dissertation zooms into the gamified leaderboard function of riders' gamified system to discover the effect of this real-time relative performance information (RRPI). I introduce a personality construct, the social comparison orientation (thereafter SCO) (Gibbons and Buunk, 1999) to find the moderating effect of SCO on the direct effect of RRPI on riders' performance. Moreover, this chapter also examines the mediation effect of occupational self-efficacy (thereafter OSE) between RRPI and riders' performance. The results from the survey data of delivery riders show that there is a positive association between the perceived usefulness of RRPI and the use frequency of RRPI and riders' performance. For the moderation effect of SCO, I find that from the use frequency of RRPI measure, riders with High SCO deliver more orders than riders with low SCO. I also document that there is a significant positive mediation effect of OSE between the perceived usefulness of RRPI and monthly orders delivered. In the additional analysis, I find that riders' familiarity with their work also influences riders' attitude to RRPI and use frequency of RRPI.

This paper contributes to RPI literature related to personality traits (Wang, 2017; Thomas and Thornock, 2017), by introducing the moderation effect of SCO. This personality of social comparison is, to the best of my knowledge, not addressed by previous literature. I find evidence that RPI produces different effects on people with heterogeneous SCO. This result informs organisations in terms of reconsidering using RPI more carefully as feedback or incentives, tailored to the characteristics of employees. Moreover, the study finds a mediation effect of OSE between RPI and riders' performance, contributing to performance information and incentive literature (Gist and Mitchell, 1992; Bonner and Sprinkle, 2002; Hall, 2008; Burney and Widener, 2013). This is important and relevant to organisations to make their employees feel capable of doing their work at their position by bringing RPI. Moreover, I also give one explanation of why a small proportion of people access the RPI that is provided to them (Kohler et al, 2023, Eyring and Narayanan, 2018). The familiarity of work influences workers' motivation to access RPI and also the attitudes to the usefulness of RPI.

The third chapter analyses a gamified system for users in an online learning

platform. I conducted 39 semi-structured interviews with users around China to discuss their user behaviours and attitudes to the gamified system. Using self-determination theory, I find that users engaged in the platform with the intrinsic motivation of learning different courses and reading books on the platform. With the implementation of the gamified system in the application, users' behaviour started to change based on these gamified elements. These external gamified elements such as the check-in button, gamified challenges and badges, and a learning credit system are internalised to be part of users' own lives and value to strengthen the original intrinsic motivation of learning on this platform. These internalisations of gamified elements further satisfy users' basic psychological needs of competency, autonomy, and relatedness.

This paper extends the literature on gamification (Chapman, Chua and Fiedler, 2021; Kelly, Valtchanov and Webb, 2021) and responds to Kelly, Valtchanov and Webb's (2021) call for using interviews to understand the issues related to gamification more in depth. Meanwhile, I contribute to previous novel incentives literature (Vogelsang, 2022; Kelly, Valtchanov and Webb, 2021, Kelly, Presslee and Webb, 2017; Presslee, Vance and Webb, 2013) by explaining how gamified designs as an incentive system are internalised by users to become part of their value. This internalisation of gamified designs can be generalised to the workplace. Furthermore, for designers of gamified systems, the paper suggests that some vague gamified performance measure criteria could prevent certain negative gaming effects from users. Instead of chasing one particular performance measure to boost their learning credits, users need to improve five different performance measures together. Finally, the designers also need to consider the existence of different types of users in a gamified system and use different gamified mechanics to motivate their behaviours (Cardador, Northcraft and Whicker, 2017; Kelly, Valchanov and Webb, 2021).

Paper 1

Play a working game: A field study of delivery riders' sense of community and motivation from a gamified incentive system

Play a working game: A field study of delivery riders' sense of community and motivation from a gamified incentive system

Abstract

Gamified incentive systems are a defining feature of platform organisations. Such systems involve many mechanics used to monitor participants' performance, such as leaderboards, points and levels as well as real-time performance information; these mechanics aim to create a gamified context that can motivate and direct the activities of employees, customers, or users. We conducted semi-structured interviews with delivery riders and a station manager drawn from a large platform organisation in China to collect their experiences of how the gamified incentive system employed by this organisation shapes riders' motivation, performance, and peer relationships. Drawing on goal-setting theory and social comparison theory, our study investigates how this gamified system helps create a sense of community that further incentivises rider performance by mitigating the negative emotions arising from gamified incentives, leading to a positive impact on their performance by creating a context for social comparison. We introduce the notion of a sense of community to the extant literature as a significant incentivising factor and seek to encourage further research into how a sense of community complements individual-level analysis of employees' motivation.

2.1 Introduction

Every day in China, millions of food delivery riders assigned by platform companies drive through cities to deliver food and grocery orders. Antonopoulou and Begkos (2020) argued that "the diffusion of digital platforms has revolutionised training modes of business and generated new professions". In light of these tremendous changes in how people work, as management accounting researchers, we are interested in what motivates these delivery riders to join and work diligently as part of these platform companies.

Merchant and Van der Stede (2017) argued for the significance of incentive systems with regard to organisations' ability to motivate employees by linking rewards and penalties to their performance. Three management control benefits result from an appropriate incentive system: informational benefits guide employees' effort in a

certain direction, motivational benefits encourage employees to work more diligently, and sorting benefits enable companies to attract and retain talent from job markets. By conducting a series of theoretical analyses, Bonner and Sprinkle (2002) introduced one of the most famous frameworks to explain how monetary incentives influence the direction, duration, and intensity of individuals' effort as well as individuals' strategy development.

Although traditional monetary incentives are relevant in this context, some organisations, especially platform companies, have started to use innovative methods to design their incentive systems (Dale, 2014). Gamification is a typical example of such a method. According to Deterding *et al.* (2011, p. 10), gamification is defined as "the use of game design elements in non-game context". Gamified incentive systems utilise the motivational power of games to incentivise individuals to engage and perform certain tasks (Richter, Raban and Rafaeli, 2015). One important characteristic of the game context is its complexity and the corresponding combination of different incentive mechanics (Crawford, 1997). Unlike traditional monetary incentives or other single-source incentives, gamified mechanics motivate players through different channels. The first such channel involves giving players different goals to achieve in the game alongside real-time performance information, such as a progression bar and rewards when a goal is achieved (Latham and Locke, 1991). Simultaneously, this approach provides public relative performance information (RPI) through leaderboards to allow players to compare themselves with others to self-evaluate their performance, self-enhance their confidence, and motivate them to engage in subsequent self-improvement (Festinger, 1954).

Moreover, financial and non-financial rewards are integrated with the two mechanisms mentioned above. According to Zichermann and Cunningham (2011), a good gamified incentive system can combine intrinsic motivation and extrinsic rewards properly to influence individuals' behaviour and performance. On the one hand, some gamified mechanics, such as specific challenges, gamified badges and levels, are directly or indirectly linked to extrinsic rewards, such as individuals' bonuses and wages. On the other hand, as mentioned by Deci and Ryan (2004), relatedness, as a basic need that motivates individuals intrinsically, is based on the connections between individuals and others with regard to certain behaviours. Some symbolic mechanics,

such as gamified leaderboards, facilitate social engagement among individuals within a work setting (Miri and Macke, 2021; Jacob *et al.*, 2022; Oprescu, Jones and Katsikitis, 2014). These gamified mechanics highlight the most important factor that we investigate in this paper: a sense of community (SoC). McMillan and Chavis (1986, p. 9) defined an SoC as "a feeling that members have of belonging, a feeling that members matter to one another and to the group, and a shared faith that members' needs will be met through their commitment to be together". This definition relies mainly on four essential perspectives: membership, influence, the integration and fulfilment of needs, and shared emotional connections.

Similar to the group emotional attachment among football fans that was discussed by Baxter *et al.* (2019), in a setting featuring gamification, we can explore not only the effects of different incentives but also the community that emerges among members. Different gamified mechanics, such as gamified goals, leaderboards, and levels, provide common symbols that help individuals unite to create communities within the boundaries of certain geographical spaces or social media groups. The SoC exhibited by these members provides emotional safety and a sense of belonging that enables them to express their vulnerability safely (McMillan, 2011), thus mitigating the negative effects of gamified challenges and self-established goals. In addition, the motivational effect of the RPI provided by the gamified leaderboards is also influenced by the SoC exhibited by members.

Therefore, we conduct a field study in a Chinese food delivery platform setting with the goal of explaining how different gamified mechanics motivate riders and explicating the role played by an SoC in this gamified context. We initially use the netnography method (Kozinets, 2019) to explore our research setting by reference to hundreds of short videos posted on the social media accounts of riders on Douyin (the Chinese name for TikTok). We then conduct 36 online semi-structured interviews with delivery riders and one station manager to collect our primary data. Finally, the riders' videos are also used to justify our findings based on the interviews.

We find that from individual perspectives, delivery riders are assigned gamified challenges by the platform system and set their own goals to achieve, while the real-time performance information provided allows riders to be aware of their progress towards certain goals. Meanwhile, riders' leaderboards indeed have motivational and

learning effects on riders' performance. Based on this gamified incentive system, riders unite to create either local or distant online communities, thus generating an SoC among members, which further motivates and encourages riders to continue working by mitigating the negative emotional influence and strengthening the motivational effect of gamified incentive mechanics.

Our study first contributes to the extant platform literature by utilising SoC theory (McMillan and Chavis, 1986) to show how SoC influences riders' work behaviour and performance. Unlike other platform literature (Chapman, Chua and Fiedler, 2021; Van den Bussche and Dambrin, 2020; Kornberger, Pflueger and Mouritsen, 2017; McDaid, Andon, and Free, 2023), we adopt a positivistic perspective to extend our understanding of the platform that enables these underrepresented gig economy workers to make a living (Goodman and Halsley, 2007). Instead of operating as individual delivery riders incentivised by the gamified incentive system, riders actually develop an SoC by working and living together with their peers while addressing the difficulties and strict controls associated with the platform. We hope that the notion of SoC can be an important factor in this context, and we aim to inspire experimental and empirical researchers to test its motivational effects in further detail. Moreover, given that emotions play a key role in SoC, we also respond to Repenning, Löhlein and Schäffer's (2021) call for accounting research on mixed emotions in digitalised platform settings. We show how the dynamic changes in riders' emotions resulting from SoC further influence riders' behaviour.

Second, we contribute to the goal-setting literature and the RPI literature (Eyring, Ferguson and Koppers, 2021; Hannan, Krishnan and Newman, 2008; Hartmann and Schreck, 2016; Tafkov, 2013) in a gamification setting. We provide field evidence regarding how this gamified incentive system influences riders' behaviour and performance through mechanisms related to goal-setting and social comparison. The complexity of this gamified incentive system allows us to respond to Hartmann and Schreck's (2016) call for further investigation of the motivational effects of both absolute goals and RPI. We explain how riders react when they face different mechanics with regard to both their absolute goals and RPI. We also answer the question raised by Carroll and Marginson (2021) regarding the possibility that the number of comparers in gamified leaderboards influences the effect of social comparison.

Finally, our paper also draws attention to the potential contributions that can be made in this context by adopting netnography to explore our setting, recruit our interview participants, and triangulate our findings based on social media. Although this approach has been mostly used in other research areas such as sociology, or marketing (Kozinets, 2019), we are trying to add to the hitherto quite limited literature that has sought to explore the methodological possibilities of netnography (Jeacle and Carter, 2011; Bialecki, O'Leary and Smith, 2017; Van den Bussche and Dambrin, 2020; Chapman, Chua and Fiedler, 2021; Yu, 2021). Given the many new kinds of data that are hidden in different social media and public websites, the potential to explore and take advantage of these data in management accounting research is tremendous.

In the following section, we introduce the main theories used in this paper as well as the relevant literature. The third section explicates the netnography methodology used in this paper. Section four presents our field findings regarding how this gamified system combines with an SoC to influence riders' motivation and behaviour. Based on these field findings, in the final section, we conclude with a discussion of how an SoC is developed from the gamified incentive system and in turn increases riders' motivation.

2.2 Theories and Research Questions

2.2.1 Goal Setting

According to Latham and Locke (1991), goal-setting theory posits that "the simplest and most direct motivational explanation of why some people perform better than others is because they have different performance goals". Two important goal attributes mediate the relationship between goal and performance (Latham and Locke, 1991; Birnberg, Luft and Shields, 2006): specificity and difficulty. Birnberg, Luft and Shields (2006) argued that the specific goal offers a clear direction for individual effort and reduces the ambiguity of goal setting. Regarding the difficulty of the goal in question, many experimental studies in the field of psychology have found a linear relationship between goal difficulty and performance. However, to achieve this positive effect, the difficulty should be within the individual's ability to achieve as well as within the individual's range of goal commitment (Birnberg, Luft and Shields, 2006; Merchant and Manzoni, 1989). Webb, Williamson and Zhang (2013) found that participants in the group that were associated with more challenging goals exhibited higher productivity, thus highlighting the motivational effect of more difficult goals.

In terms of the ability of goals to regulate individuals' performance, three main direct causal mechanisms have been proposed: goal direction, goal intensity and goal duration. First, a specific goal directs individuals towards goal-relevant activities, thus offering the information advantages of a good incentive system with regard to guiding employees in the correct direction (Merchant and Van der Stede, 2017). Second, goal intensity pertains to how much physical or mental effort an individual wants to invest in the achievement of certain targets. An important perspective on goal intensity that has frequently been adopted is that of goal commitment. This goal commitment can be enhanced by monetary incentives, ultimately motivating individuals to invest more effort and exhibit higher performance (Locke *et al.*, 1981; Bonner *et al.*, 2000; Webb, Williamson and Zhang, 2013). The third mechanism is goal persistence, which refers to how long we can exert effort towards our directed goal. Goal persistence can be viewed as a combination of the first two mechanisms (Locke *et al.*, 1981). From a temporal perspective, goal persistence mainly focuses on the long-term goal that an individual seeks to achieve.

In light of these three mechanisms underlying goal setting, gamification, as a novel incentive method, plays an important role in motivating individuals' behaviours. According to Krath, Schürmann and Von Korflesch (2021), two of the most significant characteristics of gamification are the presentation of clear and transparent goals and the relevance of those goals as well as the ability to help players set their own goals. Some gamification mechanics, such as assigned goals, badges, points, levels, and progress bars, are based on goal-setting theory. In some games, to level up, collect certain badges, etc., participants are given specific goals to achieve. In terms of difficulty, some games adjust the difficulty based on the real-time performance of players to match their current ability (Hamari, Hassan and Dias, 2018). The goal should be challenging but within the capability of players. A player can lose confidence and leave the game if it is too difficult. If the goals are too easy to achieve, the player quickly becomes bored with the game and similarly leaves.

The main aim of this approach is to encourage players to continue playing the game. This approach is in line with the two main mechanisms underlying goal setting: goal intensity and goal commitment (Latham and Locke, 1991). With appropriate specificity and difficulty, the game design enhances players' goal commitment. In

addition, the game's progression bar design continually provides players with real-time performance, thus allowing them to evaluate their current performance and distance from certain goals. Given the assigned goals and real-time performance information, players also try to set their own goals to achieve in the game. In our settings, these gamified mechanics in the riders' incentive systems aim to encourage riders to continue working to achieve specific goals to obtain higher bonuses; furthermore, they can generate their own delivery goals every day.

2.2.2 Social Comparison

Social Comparison Theory (SCT), which was developed by Festinger (1954), argues that to know ourselves and develop a self-evaluation, we need to not only obtain objective information but also to compare ourselves with others, especially with those who are similar to us in terms of the ability and effort we invest into the accomplishment of a specific task. Based on this classic theory, Buunk and Gibbons (2007) reviewed various evolutions of SCT. One of the most important such evolutions interprets social comparison as social cognition. When individuals compare themselves with others and view themselves as similar to or different from others, social comparison has different impacts on them. Two main effects thus emerge: the assimilation effect and the contrast effect (Carroll and Marginson, 2021; Esteves, Valogianni and Greenhill, 2021). According to Greenberg, Ashton-James and Ashkanasy (2007), the contrast effect generates a positive emotion when the individual compares himself or herself with others who are worse off and a negative emotion when those others are better off than the individual in question. Lyubomirsky and Ross (1997) identified this effect as "the hedonic consequences of social comparison".

Alternatively, the assimilation effect mainly occurs when individuals seek to establish similarities with comparison targets who are better off than they are. The assimilation effect often impacts individuals who have self-improvement motives (Buunk and Gibbons, 2007), i.e., individuals who want to enhance their current status through upward comparison. Therefore, the upward assimilation effect generates more positive emotions, such as inspiration, optimism, and admiration (Smith, 2000). However, negative emotions, such as pity, occasionally result from a downward assimilative comparison. When the individual compares himself or herself with a worse-off target in an assimilative way, this comparison leads to a sense of similarity in terms of fate (Smith, 2000; Buunk, Kuyper and Van Der Zee, 2005).

Most previous management accounting literature has used SCT to examine the effect of RPI on performance by employing an experimental approach (Frederickson, 1992; Tafkov, 2013; Hannan, Towry and Zhang, 2013; Berger, Fiolleau and MacTavish, 2019; Schedlinsky, Schmidt and Wöhrmann, 2020). This literature does offer us a good understanding of the motivational effect of RPI as well as some countereffects in different experimental contexts. However, field research (Carroll and Marginson, 2021) has rarely been conducted to explain how individuals use public RPI to make social comparisons and elicit the resulting motivational effects. Greenberg, Ashton-James and Ashkanasy (2007) noted that organisational settings offer us valuable opportunities to study social comparison processes. Garcia, Tor and Schiff (2013) also proposed a model to explain the differences between the individual factors and situational factors that influence social comparison. The situational factors that these authors mentioned include the number of competitors, incentive structures, and proximity to a standard. In our setting, the gamified leaderboard for riders is a significant factor with regard to riders' social comparison behaviours. We thus focus on the gamified leaderboards used by riders in this gamified incentive system to explore the effect of RPI.

2.2.3 Sense of Community

Recently, the important idea of community has been used in the platform organisation literature (Kornberger, Pflueger and Mouritsen, 2017; Van den Bussche and Dambrin, 2020; Van den Bussche and Morales, 2019; Chapman, Chua and Fiedler, 2021) to explain how accounting plays an important role in establishing relations among individuals with similar interests and norms. In line with this literature, in our delivery platform, the individual incentive system powered by different gamified mechanics such as assigned goals and leaderboards not only motivates players to continue playing the game but also provides a common language and shared experience that enables riders to unite and establish different local or distant online rider communities.

According to Oprescu, Jones and Katsikitis (2014), gamification can help solve some contemporary problems in the workplace, such as a reduced SoC. For a community, an SoC is a key motivator that encourages members to remain in the community, which is the main theory we employed in this paper. McMillan and Chavis (1986, p. 9) first defined an SoC as "a feeling that members have of belonging, a feeling that members matter to one another and to the group, and a shared faith that members' needs will be met through their commitment to be together". These authors specified

this definition by reference to four different aspects of an SoC: membership, influence, the integration and fulfilment of needs, and shared emotional connection.

First, to establish an SoC, an individual must be involved and become a member of a group, leading to a feeling of belonging (Esteves, Valogianni and Greenhill, 2021; Backman and Secord, 1959; Van den Bussche and Morales, 2019). Normally, this group has a common symbol system that is used to unite members with certain characteristics and establish boundaries for the group. The common symbol system exhibited by a group can take the form of a way of dressing, a language, rites of passage, etc. (McMillan and Chavis, 1986). As members are contained within certain boundaries, they obtain emotional safety and identification from other members in the group and improve their resilience when facing negative events (Benson and Whitson, 2022), which causes them to feel that they are accepted by the group and increases their willingness to sacrifice for the group. Each member's investment in the group makes membership more meaningful and causes members to cherish the place they have earned within the group.

Second, influence is important with regard to generating an SoC. According to McMillan and Chavis (1986), when members have a large influence on the community, they are more attracted to that community. However, they are simultaneously influenced by other members of the group. Members appreciate individual differences, but they also face the pressure of conformity, which is the result of consensual validation as a member of the group (Deutsch and Gerard, 1955; Kidd and Campbell, 1955; Chen, Nichol and Zhou, 2017). Conformity strengthens the bonds among members and improves cohesiveness among members (Shah, 1998; Dholakia, Bagozzi and Pearo, 2004). Therefore, this influence is bidirectional, referring to both the influence of a member on the community and the influence of the community on a member.

The integration and fulfilment of needs is the third component of an SoC, which is also called reinforcement. McMillan and Chavis (1986) argued that reward derived from the individual-group association is required for any group to maintain and enhance a positive sense of togetherness among members. Some factors serve as reinforcers of communities, such as the status of group members, the competence of group members and the success of the group (McMillan, 2011). Group members must feel proud of

being a member of this community and learn or benefit from other members. Therefore, a group with a strong SoC considers the shared value of members and ensures that all group members cooperate to meet others' needs while simultaneously meeting their own needs.

The final element of an SoC is a shared emotional connection. This element is closely linked to the first three elements. It further emphasises the shared traditions, history and experience of members (McMillan, 2011). To enhance this shared emotional connection, members need to share their times with others in the community to facilitate high-quality interactions with other members (Dholakia, Bagozzi and Pearo, 2004) while participating in certain important events together, such as successes or crises. As concluded by McMillan and Chavis (1986, p. 14), "Strong communities are those that offer members positive ways to interact, important events to share and ways to resolve them positively, opportunities to honour members, opportunities to invest in the community, and opportunities to experience a spiritual bond among members".

Based on these four perspectives, an SoC can be seen to exert a strong influence on the individual's behaviour, especially in the context of the emergencies faced by many platform organisations. Some research on peer-to-peer platforms such as Airbnb, Instagram, and eBay has explicitly mentioned the notion of SoC. Antonopoulou and Begkos (2020) found that experiencing many likes, followers, and reposts by other successful accounts can cause users on Instagram to feel a sense of recognition and an SoC. In the context of Airbnb, McDaid, Boedker and Free (2019) found that reciprocity, as an important characteristic in a gift economy, generates a sense of connectedness and community among Airbnb users, which has an inflationary impact on the reviews of Airbnb. Van den Bussche and Dambrin (2020) also confirmed that Airbnb users are more willing to engage with people who are similar to themselves. A platform such as Airbnb functions as a mirror to define the "narcissistic entrepreneurs of the self" and consolidate the community of users based on their evaluations. Van den Bussche and Morales (2019) concluded that the evaluation or review systems designed by the sharing platform help create a community among users.

In the context of Foursquare, Chapman, Chua and Fiedler (2021) argued that gamification encourages users to use the application in return for experiencing an SoC. According to Richter, Raban and Rafaeli (2015), the shared interests of players in and

around the gaming environment help these players form communities. Many gamified mechanics contribute to an SoC among players. Gamified points, levels, and leaderboards can depict the status of members in the community, which serves as a significant reinforcer of an SoC for some members. These mechanics can also be viewed as common symbols that delineate the boundaries of the communities. Only players who participate in the game understand what these gamified mechanics mean to them, which triggers interactions among members. Through communication and the sharing of experiences and information, players of the game can meet their personal needs and improve their mutual benefits.

In our setting, we explore how the gamified incentive system helps foster an SoC among riders, which in turn influences riders' delivery jobs. Therefore, we frame our research question as follows: ***How does an SoC among riders emerge from the gamified incentive system and further motivate riders' work?***

2.3 Research method – Netnography

2.3.1 Investigative Approach: Watching Videos and Collecting Public Online Data
Due to the evolution of social media, netnography, which is defined as "a detailed and specific approach to conducting qualitative research using social media as the basis of its datasets" (Kozinets, 2019, p. 5), has started to appear in the accounting literature (Jeacle and Carter, 2011; Bialecki, O'Leary and Smith, 2017; Van den Bussche and Dambrin, 2020; Chapman, Chua and Fiedler, 2021). Following the investigative data collection approach of netnography, we first searched and examined relevant incentive system information broadly across some popular Chinese social media platforms, such as Douyin (the Chinese version of Tiktok) and Xiaohongshu, as an initial exploration of this specific field. Although we selected and saved some useful field data, the consent gap, which refers to the fact that "the consumer who created the data do not necessarily intend or welcome the data's use in research representations" (Kozinets, 2019, p. 175), raised concerns regarding the usability of the social media data produced by these riders. To prevent ethical challenges to our research regarding the use of these blurred public data, we decided to use these data only indirectly to help explain and triangulate the findings generated from our interview data.

Furthermore, the platform's public subscription social media account data provided articles published by the platform from 2021 to 2022, which aimed to share

different riders' and station managers' stories with the public. We collected these data from this public account using an iterative approach. We checked the account regularly to determine whether new articles had been uploaded, and then we downloaded and saved every article about riders and station managers in an individual Word document. Our main reason for saving these articles immediately was the impermanence of these online articles. Some articles might not be available in the future, or potential technical issues might impede access to this platform's public subscription social media account. By the end of January 2022, we had collected 25 articles about riders and 3 articles about station managers from this public account. Some of these articles pertained to the learning and growth of these riders from new riders to veterans, which explicitly highlighted their closeness with other riders and their sense of belonging within the nearby rider community. Based on this part of the public data, we improved our understanding of some riders' attitudes towards their job and their peers. These public data also contributed to our interactive approach to data collection.

2.3.2 Interactive Approach: Online Interviews

The interactive approach in netnography is a way to coproduce data rather than merely observing and recording existing data (Kozinets, 2019, p. 245). The online interview was the main part of our approach, enabling us to collect empirical data in the field (Ahrens and Chapman, 2006). Since our interviewees were Chinese riders, we used WeChat as a synchronous video teleconferencing program to conduct all the interviews.[1] Since we collected informed consent forms, our data collection process was in line with the ethical standards of netnography[2] (Kozinets, 2019, p. 179). Following the standard framework for netnographic research with regard to the recruitment and compensation of participants (Kozinets, 2019, p. 274), every participant received a flat payment of 90 yuan (approximately £10) in compensation for their time at the end of the interview[3]. We explicitly mentioned in the consent form that participants would still receive the 90-yuan payment if they chose to withdraw during

---

[1] Concerned about the security of WeChat, we explicitly included a reference in the consent form and verbally informed participants of the potential confidentiality risks resulting from the use of WeChat due to the fact that WeChat complies with all requests from state authorities for information about data collected by this application.

[2] This field research and data collection process was approved by the ethical committee of Bristol Business School.

[3] Only one participant was concerning about the money transaction and refused to accept compensation.

the interview to protect their participation rights.

To obtain access to the riders, we first directly contacted riders with public social media accounts on platforms such as Douyin (the Chinese name of Tiktok) to ask whether they wanted to participate in this research. To verify their identities as current riders on the platform, we contacted only rider vloggers who had continued to publish rider job-related videos on their accounts in the past three months and who had worn a certain rider uniform. If they were interested in participating, we added their WeChat account, formally sought their consent with a participation consent form and conducted an online interview with them. Then, the snowballing strategy was used to recruit more participants by consulting participants who had already participated in the interview. Obtaining access to strangers online in this manner is difficult. According to Baxter and Chua (1998), demonstrating our trustworthiness as a researcher by developing rapport between interviewer and interviewee is significant in the access stage. To address these riders' concerns, we emphasised the confidentiality of their interview data and ensured that the data were anonymised and that the participants were not identifiable based on what they said.

Using some tactics drawn from Marginson (2004), we added some detailed prompt questions to inspire riders to discuss their delivery jobs in further detail. Given that some of the participants were not talkative and were not accustomed to answering open questions, providing some prompt questions and asking for some examples were especially helpful. We asked whether they wanted to be given interview transcripts to double-check their responses. Most riders were not interested in the transcripts. After each interview, following Salmons (2015), we conducted a preliminary data review to make notes on some key points and refine the questions or our approach to participants in subsequent interviews.

We conducted three rounds of online interviews. Eisenhardt (1989) mentioned that one strength of field study is its iterative approach to data collection. In the first exploratory phase, which took place in November 2020, we conducted 11 one-on-one semi-structured online interviews with current delivery riders to obtain a general understanding of riders' delivery work and the gamified incentive system. We asked them some general questions, such as "Could you please introduce your current incentive systems to me?", "In this system, which part do you pay most attention to and

why?", and "How do you think about the gamified system?".

In the second phase, which took place in April 2021, we conducted an additional 5 semi-structured online interviews: 4 follow-up interviews with previous participants to discuss changes in their incentive system and their experience with those gamified mechanics and 1 interview conducted with a station manager who was responsible for supervising and managing riders in his station[4]. By interviewing a station manager, we improved our understanding of riders' experience with current incentive systems, as station managers help riders with their work and communicate with them daily.

At the stage of exploring this setting, the videos posted on riders' social media accounts reveal that riders have good relationships with their colleagues. During the first two rounds of interviews with riders, this community-related topic was often mentioned. This situation offers us an incentive to investigate this community. Based on an abductive approach (Ahrens and Chapman, 2006), in our third phase, which took place in April 2022, we updated our interview protocols and conducted 20 interviews to improve our understanding of the gamified incentive systems and the SoC among riders. Some new questions were added, such as "Could you please share an unforgettable or recent event between you and your nearby rider friends?", "What are your feelings about being a member of a delivery rider group? Why?". In total, we conducted 36 interviews with 29 participants across these three rounds. All the interviews were recorded and manually transcribed in Chinese. One author, who is a native speaker, coded all the Chinese transcripts and chose some representative quotations to translate into English for the empirical parts of the essay. He first used professional translation software for the first-round translation and then made minor adjustments based on the Chinese language context to ensure that these quotations were more natural and authentic. A summary of the interview details is shown in Table[5].

---

[4] As station managers are scrutinised and controlled by third-party companies and the platform, most potential participants were afraid of participating in the interview without permission from their companies and thus rejected our invitations. Accordingly, the sample contains only one station manager participant. However, this factor did not influence our research results as we focused on the motivation and sociality of riders.

[5] Each column indicates the name of the participants and one phase of the interview. Each row shows the length of every interview. Given that most interviews were conducted during the riders' workday, some riders were quite busy delivering orders on the appointed days. In addition, according to the different personalities of riders, some participants were willing to provide more information regarding the job and their lives, while others were reluctant to do so. Both factors led to variations in the length of interviews, which ranged from approximately 25 min to approximately 75 min.

2.3.3 Data Analysis

Iterative data analysis features three stages. We mainly used an abductive approach to establish a dialogue between the theoretical framework and fieldwork (Alvesson and Kärreman, 2007; Ahrens and Chapman, 2006). First, during the transcription process, we highlighted some important sentences from each interview that were relevant to specific incentive functions and commented on these sentences, linking them to specific incentive practices of management accounting and psychological theories. This process provided us with a first impression of how riders use and feel about the current incentive system. Based on the first phase of interviews, we realised that these riders exhibit an SoC. Following the suggestions of Marginson (2004), during the next two phases of the interview, we attempted to improve our understanding of this SoC in the field.

Second, after transcribing and commenting on all the interviews, we started to code interview transcripts manually using Nvivo software. At this point, to understand the events that occurred in the field more deeply, we used the inductive approach to code each document sentence by sentence to categorise all the sentences into different nodes. A total of 12 main codes were included in our coding, such as GAMIFICATION, GOALSETTING INCENTIVE, PERFORMANCE INFORMATION, and RIDER SOCIETY. Under each main code, different subcodes are used to classify different specific topics. For example, the GAMIFICATION code includes the subcodes GAMIFIED LEVEL, GAMIFIED BADGES, GAMIFIED LEADERBOARD, etc. Figure 1 shows the codes and subcodes in Nvivo. Given that we coded every transcript inductively, after each phase of the interviews, we adjusted and added some new codes to our analysis and recoded previous interview transcripts. These 12 main codes remain in the final Nvivo documents. The SoC-related subcodes are contained in the RIDER SOCIETY code. We included sentences that were relevant to different codes or subcodes in all the relevant codes or subcodes to show that such sentences conveyed more than one important piece of information. Through this second round of coding, we obtained a clearer picture of how the riders responded to this gamified incentive system and how their performance was influenced by this system. Moreover, the SoC among riders was analysed in detail from four perspectives based on the original theory.

[Figure 1 about here]

Finally, in the third stage, to improve the triangulation, we verified our analysis by reference to short videos recently posted by delivery riders to ensure that our analysis and discussion reflected their work and feelings authentically. Moreover, we deepened our discussion by integrating our analysis with the extant literature, which offered us the opportunity to elucidate the distinctions among these settings and enhance the generalisability of our case to other incentive settings. While analysing these field transcripts and online articles, we relied on reflexive pragmatism (Alvesson, 2003) to account for the original framing of the information that our participants and the online articles wanted to relate. Especially for the online articles posted on the platform's public social media account, the platform framed these articles favourably to convey a good and attractive image of this on-demand delivery rider job to the public and new riders. Combining primary field data drawn from the online interviews with participants, we hope to develop an overall picture of the incentive system and the SoC among riders.

2.4 Working with a Gamified Incentive System

We begin by introducing this platform setting and gamified incentive system in Section 4.1. We then show how different gamified mechanics influence riders' performance in Section 4.2. Finally, we present our main findings on the SoC among riders in Section 4.3.

2.4.1 Platform Settings and the Gamified Incentive System

The research site is a leading e-commerce platform for life services in China. This platform focuses on mass-market, essential, and high-frequency services, which are closely linked to consumers' daily lives. Its services include food delivery, in-store dining, hotel booking, and movie tickets. In our paper, we mainly focus on the food delivery aspect. Approximately 4 million delivery riders worked or are currently working for this platform in the food delivery area (The Platform Prospectus, 2018).

There are two main categories of riders: full-time riders (FTs) and part-time riders (PTs)[6]. The differences between these riders are shown in Table 2. The platform outsources its delivery licenses to different city partner companies to hire FTs.

---

[6] The term part-time (PT) is merely a name for the kind of crowdsourcing riders in question. This name does not indicate that these riders have another full-time job. All the participants we interviewed relied solely on this delivery job as their main source of income.

Following the operating standards established by the platform, the city partner companies sign employment agreements with FTs. To ensure the quality of the services and reduce the safety risks faced by riders, the platform provides some standardised training processes and emergency response processes that city partners use to train FTs. All these trained FTs work at different stations in the city and serve a particular area[7]. PTs must only download the mobile application and complete the online training sessions using the mobile application before starting to work. Although the platform still assigns certain city partner companies to manage PTs, this form of management is quite loose and mostly performed using only the mobile application.

[Table 2 about here]

Most importantly, the platform develops a mobile application for delivery riders, which is driven by big data and an artificial intelligence algorithm; this application is currently used by all riders to support and complete their deliveries for this platform.[8] The mobile application includes different gamified mechanics. First, these mobile applications feature a gamified level system for riders. In the gamified level system, riders have different levels ranging from Bronze to King (for FTs) or from Fresher to Supreme God of War (for PTs) based on the points they earn through their delivery performance. Every successfully delivered order earns the rider 1 or 2 points. However, riders are penalised in terms of points if they make mistakes such as receiving a bad review or failing to complete a delivery on time. As riders accumulate these points, they advance to higher levels, through which they, especially PTs, can obtain privileges that lower-level riders do not have, such as collecting more orders simultaneously and receiving extra money for each order. To maintain their level every week, they must also attain a certain number of points. Otherwise, they are downgraded and lose their previous privileges. Based on their delivery achievements, they also receive symbolic gamified badges.

Second, the application also features a gamified leaderboard that provides real-time rankings of riders in the same station for FTs or riders in the same city for PTs. Each rider can notice their real-time rank and the performance of the top 50 riders.

---

[7] Each station features approximately 50 to 200 riders depending on the scale of the cities and areas served.

[8] The two kinds of riders use two different mobile applications, but most of the functions are identical.

Although different performance measurement leaderboards are used, such as leaderboards for order numbers and driving distance, the most important and popular leaderboard that riders publish on their social media accounts is the order number leaderboard. Figure 3 shows an example of one PT's current rank, which is 849 due to delivering 26 orders in the city, whereas the top rider delivered 82 orders.

[Figure 2 about here]

Finally, riders can always check real-time performance information ranging from daily performance to monthly cumulative performance. This performance information is updated in a timely manner as riders work. Some gamified mechanics, such as the progress bar, reflect the differences between the current number of orders delivered and the assigned goals associated with different levels of bonuses or indicate how many points the rider needs to upgrade or retain their rider level. Therefore, these gamified mechanics, such as points, levels, badges, leaderboards, and real-time performance information, motivate us to choose this incentive system as an appropriate and relevant field case (Eisenhardt, 1989).

2.4.2 Motivation through the Gamified Incentive System
2.4.2.1 Gamified Challenges and Rider Levels

When we asked riders about the incentive system, most riders first mentioned that the most important incentive is the ability to receive more orders to deliver at a higher order price. To earn more income, the most direct path is to deliver more orders successfully. Given that the number of orders is one of the most important key performance indicators, this gamified system includes some gamified challenges for riders to achieve based on the number of orders they deliver. Bonus challenges for riders were mentioned by many riders when we asked. PTs are often presented with daily, weekly or peak-hour bonus challenges, which require them to deliver a certain number of orders during a limited number of hours in a single day or week to complete these challenges and obtain the corresponding bonus; some monthly bonus challenges are also offered in peak seasons[9], i.e., summer and winter, which are called "summer challenge" and "winter challenge" by the platform.

---

[9] This delivery rider job is impacted by seasonal effects; namely, in summer and winter, more orders are available to deliver as customers are more willing to order food online than to eat in restaurants due to the presence of hot or freezing weather. In contrast, in spring and autumn, the total number of orders tends to decrease.

These specific challenges incentivise PTs to continue working during peak times until they reach the target or fail to achieve the challenge in a certain period. For most of these challenges, the platform usually sets three difficulty-based levels of challenge for riders to achieve and obtain bonuses. Normally, the lowest level is easy for all riders to complete, and the medium level can still be reached by half of the riders; however, it is difficult for most riders to complete the most difficult challenge:

> *"I think it is hard because the big data is really scary. It can calculate the limits of every rider and try to stretch you. It makes you believe that you can finish it, but when you actually start to deliver, you can't make it at all. It's really scary!" (Rider 13, PT)*

Some riders experience emotions of anger and disappointment because they feel that they cannot succeed in these challenges proposed by the gamified system in light of big data. For this delivery job, effort and ability are both important to achieve the goal[10]. When riders fail to complete these challenges, they start to feel disappointed in themselves and complain about the system. For the purpose of ensuring that riders continue to strive to succeed at such challenges, real-time performance information is important to guide their behaviour (Locke and Latham, 2002). To incentivise riders to participate in these bonus challenges, a progression bar is associated with each challenge that informs riders of their own position and the number of orders that must be completed to obtain a certain level of bonus. Based on this real-time performance information, riders can adjust their effort and strategy to ensure that they can complete the challenge by the deadline.

Similar to New York cab drivers (Camerer *et al.*, 1997), most PTs tend to set their own daily goals regarding the number of orders they can deliver or how much money they must make. However, due to the pandemic and seasonal effects, the total number of orders can decrease considerably and become more unstable every day, which makes it difficult for riders to complete the goals they have previously established. Their specific self-established goals thus become vague:

> *This season is particularly bad, and I have no target at present. I just planned*

---

[10] The effort involved in this context pertains to the fact that riders require more time to perform deliveries every day, while ability is more closely related to long-term familiarity with the location, problem-solving skills in response to accidents and communication skills with customers and restaurants.

to make good use of the running time. (Rider 2, PT)

Riders expressed anxiety about their incomes in those off-peak seasons. When riders feel that they are not capable of reaching these goals in their working time, their specific goals change to the vague goal of "do my best" (Latham and Locke, 1991). They can only try to plan their time and routes more efficiently to deliver as many orders as they can. Similar to PTs, FTs are also assigned gamified challenges to complete. An additional absolute goal for FTs is to deliver a certain number of orders, which results in changes to the price per order. Because of this semifixed order price, FTs have specific goals that enable them to earn more from each order. In addition to these changes in order price, most stations offer some monthly bonuses for riders who perform well[11].

Another specific goal for riders is to advance or maintain their gamified rider level. The gamified level system is also based on the number of orders that the rider delivers. When riders collect a certain number of points, they advance to the next level. Higher-level riders can collect more orders in the system on their current delivery route if they feel that they are capable of completing all deliveries on time. As the number of orders delivered determines their income, upgrading and retaining certain gamified levels becomes a certain goal for riders to achieve[12]. However, some riders mentioned that the motivational effect of this system becomes weaker when riders are not capable of fulfilling all the orders on time. The marginal return from upgrading to a higher level thus diminishes.

When PTs cannot benefit from this order privilege, they have less motivation to maintain their rider level. During the third phase of interviews, this gamified level system underwent a significant update to address this issue. The platform linked monetary income to the two highest rider levels, offering an extra 0.3 or 0.7 yuan[13] per order to these riders at these two gamified levels. In addition to the order number and price privilege, the system also prioritises high-level riders when it assigns orders. For

---

[11] As different stations in different cities are operated by different third-party companies, they have the autonomy to implement different station rules regarding monetary incentives based on data collected from riders' mobile applications. Therefore, not every station offers monthly bonuses.

[12] This situation might give rise to a concern with fairness. However, new riders are not able to deliver larger number of orders on time simultaneously. The gamified level system also offers protection for new riders with regard to overcollecting orders that they cannot complete.

[13] Normally, the price per order ranges from approximately 4 to 7 yuan depending on the distance and weight of the orders.

FTs, this gamified level system is more symbolic. Experienced riders can show off their level to new riders to obtain a sense of achievement. As FTs serve the same station in the same area, no matter how long every rider has been working as a rider, the system assigns orders to them at similar rates.

2.4.2.2 Gamified Leaderboards

Social comparison theory (Festinger, 1954) claims that people tend to compare their performance with relevant peers. Since riders have access to real-time leaderboards through their mobile applications, thus enabling them to access RPI at any time, their different attitudes towards these leaderboards lead to different effects and emotions for different riders. Most experienced PTs mentioned that they rarely check the leaderboards:

> *This (leaderboard) does not affect me. Probably, at the beginning of my career, if someone delivered more orders, I would try to learn some tactics from them. When I clearly knew the limit of how many orders I could deliver per day, I am not interested in that. Therefore, I do not pay attention to this leaderboard. (Rider 2, PT)*

During the initial stage, these leaderboard rankings satisfied riders' self-improvement needs by enabling them to compare themselves with other outperformers. This upward assimilative comparison generates emotions related to appreciation and inspiration that encourage riders to pursue better performance. To set proper goals for themselves, riders use the leaderboards to identify their own position and learn from other riders by asking for tips from top riders who are located nearby. The real-time RPI of other riders has some motivational and learning effect initially but then gradually becomes less important as riders encounter their limitations in terms of ability and the effort they want to invest into this job. Another reason why riders ignore the leaderboard is that the top performers seem too good to be true. As noted by Newman and Tafkov (2014), when lower and middle performers receive public RPI that indicates that they are unlikely to reach the top, they start to decrease their efforts:

> *I really hate this leaderboard. I feel that the leaderboard shows some model riders in Beijing, who only care about work and money, as examples to push more riders to work harder without considering their health. (Rider 10, PT)*

Instead of participating in this tournament leaderboard game, which lacked any real

monetary incentives, some riders also take into account their work-life balance and their health. To some extent, this quotation confirms the motivational effect of leaderboards on some riders. To earn more money and thus become a top rider, some riders ignore their health and choose to continue participating in this tournament game despite insufficient rest.

Rider 3 transferred from FT to PT and highlighted another main reason for ignoring this leaderboard:

> *For PT, the leaderboard doesn't matter. As the number of riders in the city is too large and there are many brilliant riders, you can't compete with them. Furthermore, riders are located in different areas of this city. Some areas have more orders to deliver, but your area has fewer. It is not comparable at all. (Rider 3, FT to PT)*

Two important factors mentioned in this context were the number of comparers and the external context comparability. Tafkov (2013) mentioned the conditions for comparison. One such condition is the similarity of relevant performance characteristics. If a rider participating in the tournament game is not familiar with most of other riders in an incomparable context, the motivation to participate in this tournament game decreases. Therefore, the city leaderboards of PTs have a weak motivational effect on PTs. This finding also answered the question raised by Carroll and Marginson (2021) regarding the influence of the number of comparers on the effect of social comparison. When the leaderboards include many comparers who are not closely related to riders, riders do not pay attention to the RPI provided by leaderboards.

Interestingly, although veterans claimed that they rarely use this leaderboard function, they do check their rankings on days when their performance was fairly good. This behaviour mainly facilitates self-enhancement by informing riders of the fact that they can beat other riders based on comparisons between themselves and others. The social comparison facilitated by this leaderboard elicits emotions of joy and pride from riders based on self-focused perspectives (Smith, 2000). Although some riders ignore or even resist this leaderboard, it does motivate other riders. These leaderboards provide a tournament game that encourages some riders to continue chasing high performance or pushing themselves closer to top performers.

When FTs obtain public RPI, most FTs experience more emotions related to inspiration, treating the top rider in their station as a model from whom they can learn and whose performance they can seek to emulate. As the leaderboards for FTs show only the rankings of other riders in their station, the leaderboards facilitate comparison with similar peers and small numbers of comparers. Furthermore, the motivational effect of the leaderboard becomes more obvious when the station offers a performance bonus to the top rider. Newman and Tafkov (2014) found that public RPI has a positive effect on the performance of the top performers in a rewarding tournament. In some stations that did not offer a bonus to the top rider, the effect of the leaderboard gradually decreased. In contrast, riders in stations that offer a bonus to the top 5 riders, i.e., those who occupy the top 5 ranks (Garcia, Tor and Gonzalez, 2006), are more competitive at the end of the month in hopes of receiving the bonus. If this leaderboard function is not linked to such a monetary incentive, both PTs and FTs become bored with this tournament game since the ranking has only the symbolic meaning of enhancing the current top riders' sense of achievement.

2.4.3 Community under the Gamified Incentive System

In most interviews, riders noted that they had good relationships with their colleagues despite the fact that they worked as competitors within the same gamified leaderboard to some extent. A strong SoC is evident among riders. In this section, we show how four elements (membership, influence, the integration and fulfilment of needs, and shared emotional connection) interact to establish a strong SoC among riders.

2.4.3.1 Membership – Feelings of Emotional Safety among Riders

According to McMillan and Chavis (1986), membership is a feeling that one belongs to a particular group. Membership requires certain boundaries to enable members to experience the emotional safety necessary to express their needs and expose their feelings. The common symbol system establishes certain boundaries that enable members to unite. In our setting, the same gamified incentive systems offered such common symbols to these riders, thus enabling them to establish boundaries for the group. Only riders who are currently participating in this delivery game understand what their peers are currently experiencing. Many different situations and obstacles can influence their deliveries:

> *"For example, like what happened this morning, a restaurant was too slow to cook my orders, or a big community didn't allow me to drive in …, all these*

made me frustrated during the day because they all wasted much of my time (that could have been used to deliver more orders)" (Rider 15, PT)

As a relatively individualised job, riders need a group to be able to speak honestly and exhibit vulnerability safely (McMillan, 2011). Normally, based on their geographical location, PTs are always located close to a small group containing approximately 10 to 15 members. Some of these groups created a WeChat group to communicate with other members:

"We have our communication channels and pour out our unhappiness into the group chat, or sometimes when I meet other friends, I will talk with them about the situations and customers I met; otherwise, I will be very depressed if I hold it in my heart for a long time". (Rider 18, PT)

Although gamified challenges with bonuses do offer some direction and motivate riders to work diligently, these goals also trigger some negative emotions such as anger, disappointment and anxiety during the challenges, thus demotivating riders. These negative emotions and experiences, as emotional hooks, unite group members. In rider groups, riders can express this vulnerability safely and feel a sense of belonging and acceptance. Riders must overcome their frustration and disappointment to improve their delivery efficiency at a later point. As group members normally deliver orders in similar areas, when they are free at an off-peak time, they have a fixed place to gather to chat with each other about their performance on the day and how many orders they need to deliver to complete challenges, complain about some ridiculous orders and goals assigned by the platform, and comfort each other, thus relieving their anger and frustration.

Beyond the constraints of geographical location, to establish more connections, some riders operate public social media accounts. They use short videos to share their experiences or tactics with other riders and to give guidance to potential new riders regarding how to join the platform and deliver orders. They try to create a wide community through social media, involving riders in different cities. This behaviour even crosses different delivery platforms. Given the similar gamified functions offered by different platform mobile applications, in an online broad community, a common symbol among the members is their identity as riders:

43

"Every time they come to my livestream, I will tell them no matter if you are from the XXX, XXX or XXX platform, we only have one nickname, that is, "the delivery rider". We are all in this industry together." (Station Manager)

This quotation from a station manager reflects a strong sense of belonging or identification; the manager had previously served as a rider for two years. Although these online communities do not feature strong in-person connections, they also provide riders around China with a safe place in which they can discuss and complain about events that occurred during their deliveries. Some live-streamed discussions we observed pertained to the price and number of orders in a different city, the achievability of the gamified challenge goals and the difficulties riders encountered during the delivery process. Livestreaming riders chat with their followers and respond to their questions and emotions.

In these two different ways, riders become members of a local or online group, and they obtain emotional safety and a sense of belonging from their groups. However, according to McMillan (1996), membership can be only the initial spark for an SoC. To ensure that this spark becomes a fire, the notion of influence plays a significant role in the SoC.

2.4.3.2 Influence – Conforming to More Efficient Delivery Practices

Influence, as the second element of an SoC, is a bidirectional process; thus, while riders can influence the group, they are simultaneously influenced by the group. In a group, members have an innate need to ensure that what they are experiencing, seeing and feeling is the same as what is experienced by other members; this need is known as "consensual validation" (McMillan and Chavis, 1986). Daily chats with other riders facilitate this process of consensual validation, allowing riders to reach an agreement regarding when the off-peak season is and when it is unnecessary to set high self-goals to stretch themselves. This consensual validation eliminates riders' anxiety due to their failure to achieve their previously self-established goals. Riders are aware of this situation of low demand for delivery services and adjust their expectations regarding their performance jointly, allowing themselves to have vague self-established goals in off-peak seasons.

In addition to adjusting riders' expectations, during riders' breaks, members congregate to share some real-time delivery information, such as which roads are closed

or which buildings do not allow riders to go upstairs. This sharing can subsequently help improve group members' efficiency. Although in principle, the platform does not allow riders to deliver others' orders, members of the same group try to allocate their labour resources more efficiently. When they collect or deliver their orders, they check with nearby members regarding whether they have orders to collect or deliver in the same building, a task with which they can help. More interestingly, some exclusive tips must be kept within the small group so that only group members can benefit from such tips:

> *We also have selfish considerations. For example, a place where it is easy to deliver is known by many riders. Therefore, we don't want more riders to know these tips. (Rider 23, PT)*

The sharing of these tips and real-time information helps generate group conformity (Deutsch and Gerard, 1955) and promotes more efficient delivery behaviour. Every rider contributes some information, tips or labour resources, thus improving the delivery efficiency and ultimately the performance of all members of the group.

The gamified leaderboards also facilitate influence among members. One previously mentioned reason for the tendency of riders, especially PTs, to ignore the leaderboards is the incomparability of these leaderboards, as too many riders from different districts are included in the same city leaderboard. However, although riders might not be interested in the performance of the top 10 performers in the city, as members of a small community, the ranking information of their friends provided by the leaderboards does indeed facilitate social comparisons between PTs and their close friends, who normally have similar delivery routes and who wait for orders and rest together in the same place. Not only do the rankings of group members shown by the leaderboards trigger social comparison, they also create the opportunity to engage in benign competition among members:

> *Between us, we have a little bit of a feeling of a small competition, that is, I am not convinced that you are better than me. You are not convinced that I am better than you. There are a few friends like this around me. (Rider 27, PT)*

Such benign competition among members within the group is based on their rankings

according to the gamified leaderboards (Shah, 1998; Mussweiler and Rüter, 2003). As riders' most familiar peers, nearby group members represent their first choice of social referents or standards for comparison. The RPI provided by leaderboards offers a specific and fair performance measure that enables group members to determine who wins the game. Although such a game features winners and losers, group members all improve their performance and earn more money through these benign competitions. Instead of generating shame-related emotions through social comparison (Hannan *et al.*, 2013), riders experienced more joy-based emotions due to this win-win benign competition within their groups. This influence among members enhances the function of the leaderboards by providing a context in which PTs can use the ranking RPI provided by the gamified leaderboards. In other words, this factor improves the comparability and usefulness of the gamified leaderboards among PTs, who normally feel that the leaderboard is useless.

The same claim is true of FTs. As FTs share the same leaderboards with their station peers, this community at the station level ensures that the leaderboards are more comparable among group members. After being promoted from a rider to a station manager, the station manager related the following information:

> ***"This leaderboard has a huge motivational effect. When I was a rider, I still remembered that the first day I delivered, I could not find my name on the leaderboard, and I only delivered 20 orders. Gradually, I started to see my name on the leaderboard, and I became the top rider once in our station". (Station Manager)***

In this context, influence also pertained to the fact that new riders learn from old riders and conform to their more effective behaviours. Initially, new FTs learn different skills from veterans. The more experience they obtain, the more they can narrow the gap between themselves and old riders, sometimes even surpassing veterans' performance. All this improvement is reflected in the leaderboards. Moreover, the top performance exhibited by veterans in the station leaderboards also provides certainty and inspiration to new riders, informing them that they can also exhibit such performances if they continue to work diligently in their station.

In this section, with the help of the gamified leaderboards, we show how groups influence riders' behaviour towards the development of group conformity; we also

explore how riders contribute to group norms and collective wisdom, thus enhancing the closeness experienced by members.

2.4.3.3 The Integration and Fulfilment of Needs – Satisfying the Needs of Riders

Although members within certain boundaries influence each other to sustain an SoC, reinforcing the relationship among members requires another element, namely, the integration and fulfilment of needs. For any group to maintain a positive sense of togetherness, members must receive rewards from this group. People are attracted by the group that can offer the most rewards (McMillan and Chavis, 1986). The case is the same for delivery riders. Regarding the positive influence of group conformity mentioned above, only when riders feel that they can benefit from their group can they be more aligned with group members.

The WeChat group not only provides a safe environment in which riders can express their unhappiness; it is also a place in which members can ask for help during delivery. Every day, riders face different kinds of circumstances, such as accidents during deliveries or power failures with regard to their electric bicycles. When their friends in the group are in need, other members of the group come to help them even at the expense of their delivery jobs:

> *One day at 11 pm in the evening, I delivered a distant order to a place 8 kilometres away. After I delivered the order, my electric bicycle ran out of power. I asked for help in our group. One of my friends, who was 8 kilometres away from where I was, drove to bring me back home after he finished his delivery order. (Rider 27, PT)*

Most riders mentioned similar events and highlighted this reciprocal relationship among members. In line with certain kinds of group norms or shared values mentioned by Gouldner (1960), riders believe that helping others at this time will result in reciprocal behaviour from other members when they need help in the future. As McMillan (1996, p. 322) noted, "A community cannot survive unless members make fair trades with one another". When riders fulfil the needs of other members of the group, to some extent, they can also fulfil their own needs because everyone will encounter some circumstances on the road that require help from others.

In the broad community based on the internet, rider influencers also try to use their short videos and livestreaming events to fulfil the needs of other riders and

potential new entrants to this job. Using their social media accounts, they are happy to help new riders with any issues they encounter during their work:

"Through these short videos, these followers found me to ask some related questions to participate in this job. I helped many new entrants start their rider career journey". (Rider 3, PT)

Simultaneously, these rider influencers are becoming more famous on social media and normally generate some extra income from the Douyin platform by making videos and even inserting advertisements into their new short videos. This situation once again reflects the fulfilment of needs in the community since these influencers do not simply make those videos for fun or due to pure altruism; rather, they receive some monetary compensation by fulfilling other riders' needs online.

Moreover, one reinforcer of the community pertains to the competencies of other members (McMillan, 2011). The gamified leaderboards help highlight the competence of the top riders in the riders' community. High rankings on the leaderboards provide top riders with opportunities to screenshot their leaderboard rankings and share them with other group members to highlight their achievements. By fulfilling top riders' self-evaluation and self-enhancement needs, this gamified leaderboard reinforcer enhances the emotions of joy and pride resulting from social comparison. Recognition from members in riders' communities thus enhances the motivational effect of the leaderboards. In some short videos, when new riders find out there are Top 10 riders in their small community, they also feel honoured to be members of this community because they observe the competence of their peers and feel more certain that being a member of this small community is a good decision.

2.4.3.4 Shared Emotional Connection – Sharing Time with Other Riders

As a final element of SoC, a shared history is the basis on which members can establish a close and shared emotional connection. This fact does not entail that members must participate in the history to be shared, but they should identify with it (McMillan and Chavis, 1986). In terms of such a shared history, the reasons why members taken on this job are important. First, it is difficult for most riders to find better-paying jobs because of their low levels of education and low-skilled background. Second, some riders face certain financial pressures, such as paying off a mortgage or debt, and they thus choose this job to earn quick money. Third, some riders have a

background of failed start-ups or restaurants and became riders to accumulate initial funding to restart their businesses at a later point. These three reasons offer a common basis on which most riders can understand other members' previous backgrounds.

After taking on this job, in which context they are controlled by the same gamified incentive system, all riders experienced similar feelings of failing to achieve their gamified challenges, being insulted by rude customers and being penalised by the platform. All these shared experiences strengthen riders' shared emotional connections within this community (Esteves, Valogianni and Greenhill, 2021; Backman and Secord, 1959; Van den Bussche and Morales, 2019). To enhance this shared emotional connection further, high-quality interaction among members is quite significant. The quality of the interaction depends on how dramatic the event in which members jointly participate is (McMillan, 1996). Most riders can recall some unforgettable events in which they helped each other during their delivery job, such as the points that we mentioned in the previous section. Some riders even noted that members experienced the pandemic crisis together:

> *"In Shanghai, we are still in lockdown at home. At present, there are a few riders who can come out to deliver orders. Then, they delivered some groceries to us who are still in self-isolation". (Rider 19, FT)*

During the most severe period of the Shanghai pandemic, which was associated with a shortage of groceries in every household, her friends at the station continued to share and deliver groceries to her, which struck her as a very touching and unforgettable moment.

Most riders also mentioned that they organised some joint activities, such as playing cards, sharing a meal or taking a short day trip with other members of their group to whom they were close; these activities also represented good opportunities to rest and discuss their overall job experiences:

> *"Because there are many places to visit in Wuhan, for example, we will take a break to see cherry blossoms together". (Rider 26, PT)*

> *"Some time ago, the spring weather was particularly good. We went to a park to barbecue, and in some other groups, they said they like fishing, so they made an appointment to go fishing". (Rider 27, PT)*

49

During these events, riders established high-quality connections with other members and became more familiar with other members. In addition to sharing their recent job experiences, riders also discussed their personal lives and recent headline news. As a result of some video interviews, we found that some of these riders also lived together as flatmates:

> *"It is good to live with other riders. If anyone has some issues or needs any help, we can help each other out. During the traditional festival, we also have a meal together". (Rider 2, PT)*

As the final element of SoC, the shared emotional connections became much closer as a result of every high-quality interaction, which further strengthened the SoC within the group.

In conclusion, as mentioned by McMillan (1996), four elements of SoC lead to a self-reinforcing circle that encourages riders to remain in this community and continue working as delivery riders. Based on our follow-up investigation in 2023, in the context of a job featuring high turnover rates, at least half of our participants were still working as delivery riders in their cities, although the gamified incentive systems continued to be updated and became more controlling and restrictive over that three-year period.

2.5 Discussion

Based on our field findings, Figure 4 illustrates the process by which this platform company uses a gamified incentive system to motivate riders through different channels, thus helping generate an SoC among riders.

[Figure 4 about here]

Confirming the extant MA literature related to goal setting (Birnberg, Luft and Shields, 2006; Webb, Williamson and Zhang, 2013; Bonner and Sprinkle, 2002), two important gamified mechanics (gamified challenges with bonuses and gamified rider level) provide specific goals for riders to motivate them during their work. Although specific goals are provided to guide riders, the player's possession of high goal intensity and persistence (or the lack thereof) is key to the motivational effect on performance. The need to encourage riders to continue engaging in this delivery game is satisfied by immediate feedback from gamification (Krath, Schürmann and Von Korflesch, 2021). Based on the real-time performance information provided by the gamified progression

bar, riders can always check their progress towards a certain goal. This situation enhances riders' goal commitment, which motivates them to continue to pursue their goals instead of giving up. In this context, we are aware of the combined effect of monetary incentives and gamified challenge goals. As previous research has suggested (Bonner *et al.*, 2000; Bonner and Sprinkle, 2002; Webb, Williamson and Zhang, 2013), a connection between monetary bonuses and gamified challenges plays an important role in motivating riders' performance.

The other motivation mechanism of this gamified incentive system pertains to social comparison (Festinger, 1954). We extend our understanding of the effect of RPI on performance as a dynamic process. At the beginning of riders' careers, when riders check the leaderboard, they tend to make upward assimilative social comparisons (Smith, 2000) to self-evaluate their performance and find a benchmark or model from whom they can learn. At this stage, the leaderboards have both a motivational effect and a learning effect on riders' performance. As time progresses and riders become aware of their boundaries in terms of ability and effort, the RPI provided by leaderboards starts to lose its original value as an incentive. Riders rarely check the leaderboards. Instead, they focus more on their own absolute goals and performance. This finding responds to the call by Hartmann and Schreck (2016) for further investigation of how rankings influence individuals' behaviour in cases of gamified absolute goals. Moreover, our finding regarding the difference between PT and FT leaderboards answers Carroll and Marginson's (2021) question concerning the effect of the number of comparers on the individuals' behaviours and further confirms the conditions of social comparison mentioned by Tafkov (2013). Future studies can explore the effects of these attitude changes on RPI resulting from familiarity with a certain job and other situational factors in further detail.

Nevertheless, the gamified incentive system simultaneously has a demotivational effect and leads to mixed negative emotions. On the one hand, when gamified goals are not achieved, this failure significantly influences riders' self-efficacy in their delivery job, which causes frustration, disappointment and anxiety emotions in riders. On the other hand, gamified leaderboards also elicit hopelessness, discomfort and even anger emotions when riders feel that the top level of performance is too high to emulate. To mitigate the negative effects of these gamified incentives, an SoC

51

(McMillan and Chavis, 1986) plays an important role in enhancing the function of this gamified incentive system.

Based on certain geographic locations, riders can establish small communities with nearby peers. Riders in the community can feel a sense of emotional safety due to the presence of other members that enables them to express their anger-based and fear-based emotions (Repenning, Löhlein and Schäffer, 2021) and share their happiness (Liu *et al.*, 2016). We extend the findings of Benson and Whitson (2022) indicating that an SoC improves people's resilience in the face of job difficulties from undergraduate students to riders; this impact can enhance riders' goal commitment to gamified challenges. When riders deal with their mixed negative emotions by connecting with other community members, they feel relieved and recover their confidence with regard to their ability to achieve the new gamified challenges or self-established goals.

Within such a small community, riders tend to develop shared norms of behaviour that are followed by community members (Hamari and Koivisto, 2015). As members of the group, riders must conform to these norms to meet the expectations of others (Deutsch and Gerard, 1955; Kidd and Campbell, 1955; Bicchieri, 2005; Shang, Abernethy and Hung, 2020). As Van den Bussche and Morales (2019, p. 355) noted, "Not everyone will follow the norms, but users who feel they belong to the community almost always will". According to McMillan and Chavis (1986), this group conformity further enhances the cohesiveness of the group of riders and thus helps improve all members' performance. In this context, the gamified leaderboards also serve as reinforcers of communities by highlighting the competence of top riders for other group members. Future research can examine the moderating effect of an SoC on the relationship between RPI and performance in further detail. Moreover, based on the fourth element reported by McMillan and Chavis (1986), these high-quality interactions based on unforgettable events and activities create more shared emotional connections among riders, thus further enhancing the SoC within the small group.

Although a small group based on geographic location exhibits a stronger SoC, with the development of current social network technology (Hamari, Hassan and Dias, 2018), this SoC does not stop at the level of small local groups. Riders also establish broader communities online with distant riders as well as workers from other sharing economy platforms (Van den Bussche and Morales, 2019; Van den Bussche and

Dambrin, 2020). By filming short videos or livestreaming, riders who participate in the same or similar gamified incentive systems reunite in online communities, thus acquiring emotional safety and feelings of belonging. Future research can examine the differences between the local communities and online communities formulated by different evaluation or review measures in further detail based on the vibrant development of sharing economy platforms (Kornberger, Pflueger and Mouritsen, 2017; Van den Bussche and Morales, 2019).

2.6 Concluding Remarks

In summary, we analyse how different gamified mechanics motivate riders to work through two channels, i.e., goal-setting and social comparison. Furthermore, we explain how riders under this gamified incentive system unite to generate an SoC, which further incentivises riders to continue playing this game by working as riders. According to Deterding (2019, p. 131), gamification can be framed in two ways in the management context. On the one hand, "the rhetoric of choice architecture casts humans as rational actors and games as perfect information and incentive dispensers, giving managers fine-grained control over people's behaviour". On the other hand, "the rhetoric of humanistic design casts humans as growth-oriented and games as environments optimally designed to afford positive, meaningful experiences". In line with the rhetoric of choice architecture, the gamification of delivery riders' incentive system encourages riders to continue working by offering gamified challenges, rider levels, and leaderboards. Although less playfulness and joy result from gamification itself, the complexity of the gamified incentive system offered us a good opportunity to explore the motivational and learning effects of different incentives and the SoC among these gig economy workers.

In light of the increasing use of gamification in the workplace, our field study hopes to draw the attention of management accounting researchers to this novel kind of incentive system. In addition, the SoC derived from the gamified incentive system is also an important moderating factor in the behaviour of riders with regard to acquiring emotional safety, conforming to certain group norms and interacting with group members. We highlight this factor to encourage future experimental or survey researchers to test the motivational effect of the SoC in further detail. Although it is difficult to simulate this long-term dynamic SoC in the context of a single experiment,

we hope that the four different ingredients of SoC (membership, influence, the integration and fulfilment of needs, and shared emotional connection) can inspire some future intriguing experimental and survey studies.

For practitioners, our study helps platform organisations understand gig economy workers' attitudes towards these gamified incentives and the corresponding changes in their behaviours and emotions due to gamification, which can inspire platform organisations to design more efficient and effective gamified incentive systems based on their specific organisational situations. Although some large platforms have already used gamification to motivate workers individually and are worried about the unionisation of individual workers, we find that the SoC resulting from individual gamified incentive systems can enhance the impact of gamification if the platform offers workers more space to use gamified mechanics in their own way. The current use of gamification in factories by some large platforms might sacrifice these benefits by isolating the workers and encouraging them to focus on their own work, resulting in negative emotional impacts on workers and damage to platforms' image of corporate social responsibility (Anderson, 2021).

While the main purpose of gamification in the workplace is always to ensure better control of worker behaviours, according to the rhetoric-humanistic design perspective endorsed by Deterding (2019), organisations should consider the design of gamification in further detail to elicit intrinsic motivation (competence, autonomy and relatedness) (Ryan and Deci, 2017) from every employee to create a more enjoyable and inclusive experience for them during the working process, which can hopefully improve the overall performance of organisations.

2.7 Reference

Ahrens, T. and Chapman, C. S. (2006) 'Doing qualitative field research in management accounting: Positioning data to contribute to theory', Accounting, Organizations and Society, 31(8), pp. 819-841.

Alvesson, M. (2003) 'Beyond neopositivists, romantics, and localists: A reflexive approach to interviews in organizational research', Academy of management review, 28(1), pp. 13-33.

Alvesson, M. and Kärreman, D. (2007) 'Constructing mystery: Empirical matters in theory development', Academy of Management Review, 32(4), pp. 1265-1281.

Anderson, G. (2021) 'Amazon Gamifies Warehouse Worker Tasks, Raising Experts' Concerns'. Available at: https://www.forbes.com/sites/retailwire/2021/03/23/amazon-gamifies-warehouse-worker-tasks-raising-experts-concerns/?sh=54f909103810 (Accessed 08/09/2023).

Antonopoulou, K. and Begkos, C. (2020) 'Strategizing for digital innovations: Value propositions for transcending market boundaries', Technological Forecasting and Social Change, 156, pp. 1-13.

Backman, C. W. and Secord, P. F. (1959) 'The effect of perceived liking on interpersonal attraction', Human Relations, 12(4), pp. 379-384.

Baxter, J., Carlsson-Wall, M., Chua, W. F. and Kraus, K. (2019) 'Accounting and passionate interests: the case of a Swedish football club', Accounting, Organizations and Society, 74, pp. 21-40.

Baxter, J. A. and Chua, W. F. (1998) 'Doing field research: practice and meta-theory in counterpoint', Journal of Management Accounting Research, 10, pp. 69-87.

Benson, O. M. and Whitson, M. L. (2022) 'The protective role of sense of community and access to resources on college student stress and COVID‐19‐related daily life disruptions', Journal of Community Psychology.

Berger, L., Fiolleau, K. and MacTavish, C. (2019) 'I Know Something You Don't Know: The Effect of Relative Performance Information and Individual Performance Incentives on Knowledge Sharing', Journal of Management Accounting

Research, 31(2), pp. 19-35.

Bialecki, M., O'Leary, S. and Smith, D. (2017) 'Judgement devices and the evaluation of singularities: The use of performance ratings and narrative information to guide film viewer choice', Management Accounting Research, 35, pp. 56-65.

Bicchieri, C. (2005) The grammar of society: The nature and dynamics of social norms. Cambridge University Press.

Birnberg, J. G., Luft, J. and Shields, M. D. (2006) 'Psychology Theory in Management Accounting Research', in Chapman, C.S., Hopwood, A.G. and Shields, M.D. (eds.) Handbooks of Management Accounting Research: Elsevier, pp. 113-135.

Bonner, S. E., Hastie, R., Sprinkle, G. B. and Young, S. M. (2000) 'A Review of the Effects of Financial Incentives on Performance in Laboratory Tasks: Implications for Management Accounting.', Journal of Management Accounting Research. 2000, 12, pp. 19.

Bonner, S. E. and Sprinkle, G. B. (2002) 'The effects of monetary incentives on effort and task performance: theories, evidence, and a framework for research', Accounting Organizations and Society, 27(4-5), pp. 303-345.

Buunk, A. P. and Gibbons, F. X. (2007) 'Social comparison: The end of a theory and the emergence of a field', Organizational Behavior and Human Decision Processes, 102(1), pp. 3-21.

Buunk, B. P., Kuyper, H. and Van Der Zee, Y. G. (2005) 'Affective response to social comparison in the classroom', Basic and applied social psychology, 27(3), pp. 229-237.

Camerer, C., Babcock, L., Loewenstein, G. and Thaler, R. (1997) 'Labor Supply of New York City Cabdrivers: One Day at a Time*', The Quarterly Journal of Economics, 112(2), pp. 407-441.

Carroll, E. and Marginson, D. (2021) 'Relative performance information and social comparisons: Exploring managers' cognitive, emotional and dysfunctional behavioral processes', Management Accounting Research, 53.

Chapman, C., Chua, W. F. and Fiedler, T. (2021) 'Seduction as control:

Gamification at Foursquare', Management Accounting Research, 53, pp. 100765.

Chen, C. X., Nichol, J. E. and Zhou, F. H. (2017) 'The effect of incentive framing and descriptive norms on internal whistleblowing', Contemporary Accounting Research, 34(4), pp. 1757-1778.

Deci, E. L. and Ryan, R. M. (2004) Handbook of self-determination research. University Rochester Press.

Deterding, S. (2019) 'Gamification in Management: Between Choice Architecture and Humanistic Design', Journal of Management Inquiry, 28(2), pp. 131-136.

Deterding, S., Dixon, D., Khaled, R. and Nacke, L. 'From game design elements to gamefulness: defining" gamification"'. 2011, 9-15.

Deutsch, M. and Gerard, H. B. (1955) 'A study of normative and informational social influences upon individual judgment', The journal of Abnormal and social psychology, 51(3), pp. 629.

Dholakia, U. M., Bagozzi, R. P. and Pearo, L. K. (2004) 'A social influence model of consumer participation in network-and small-group-based virtual communities', International journal of research in marketing, 21(3), pp. 241-263.

Eisenhardt, K. M. (1989) 'Building Theories from Case Study Research', The Academy of Management Review, 14(4), pp. 532.

Esteves, J., Valogianni, K. and Greenhill, A. (2021) 'Online social games: The effect of social comparison elements on continuance behaviour', Information & Management, 58(4), pp. 103452.

Eyring, H., Ferguson, P. J. and Koppers, S. (2021) 'Less Information, More Comparison, and Better Performance: Evidence from a Field Experiment', Journal of Accounting Research, 59(2), pp. 657-711.

Festinger, L. (1954) 'A Theory of Social Comparison Processes', Human Relations, 7(2), pp. 117-140.

Frederickson, J. R. (1992) 'Relative performance information: The effects of common uncertainty and contract type on agent effort', Accounting Review, pp. 647-

669.

Garcia, S. M., Tor, A. and Gonzalez, R. (2006) 'Ranks and rivals: A theory of competition', Personality and Social Psychology Bulletin, 32(7), pp. 970-982.

Garcia, S. M., Tor, A. and Schiff, T. M. (2013) 'The Psychology of Competition: A Social Comparison Perspective', Perspectives on Psychological Science, 8(6), pp. 634-650.

Goodman, P. and Halsley, E. (2007) 'Social comparison processes in an organizational context: New directions', Organizational Behavior and Human Decision Processes, 102(1), pp. 109-125.

Gouldner, A. W. (1960) 'The norm of reciprocity: A preliminary statement', American Sociological Review, pp. 161-178.

Greenberg, J., Ashton-James, C. E. and Ashkanasy, N. M. (2007) 'Social comparison processes in organizations', Organizational Behavior and Human Decision Processes, 102(1), pp. 22-41.

Hamari, J., Hassan, L. and Dias, A. (2018) 'Gamification, quantified-self or social networking? Matching users' goals with motivational technology', User Modeling and User-Adapted Interaction, 28(1), pp. 35-74.

Hannan, R. L., Krishnan, R. and Newman, A. H. (2008) 'The effects of disseminating relative performance feedback in tournament and individual performance compensation plans', Accounting Review, 83(4), pp. 893-913.

Hannan, R. L., McPhee, G. P., Newman, A. H. and Tafkov, I. D. (2013) 'The Effect of Relative Performance Information on Performance and Effort Allocation in a Multi-Task Environment', Accounting Review, 88(2), pp. 553-575.

Hannan, R. L., Towry, K. L. and Zhang, Y. (2013) 'Turning Up the Volume: An Experimental Investigation of the Role of Mutual Monitoring in Tournaments', Contemporary Accounting Research, 30(4), pp. 1401-1426.

Hartmann, F. and Schreck, P. (2016) 'Rankings, Performance, and Sabotage: The Moderating Effects of Target Setting', European Accounting Review, 27(2), pp. 363-382.

Jacob, A., Faatz, A., Knüppe, L. and Teuteberg, F. (2022) 'The Impact of Gamification on Macro-and Micro-level Social structures—The Case of an Industrial Organization', International Journal of Human–Computer Interaction, 38(7), pp. 614-630.

Jeacle, I. and Carter, C. (2011) 'In TripAdvisor we trust: Rankings, calculative regimes and abstract systems', Accounting, Organizations and Society, 36(4-5), pp. 293-309.

Kidd, J. S. and Campbell, D. T. (1955) 'Conformity to groups as a function of group success', The Journal of Abnormal and Social Psychology, 51(3), pp. 390.

Kornberger, M., Pflueger, D. and Mouritsen, J. (2017) 'Evaluative infrastructures: Accounting for platform organization.', Accounting, 60, pp. 79.

Kozinets, R. V. (2019) Netnography: The essential guide to qualitative social media research. Sage.

Krath, J., Schürmann, L. and Von Korflesch, H. F. O. (2021) 'Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning', Computers in Human Behavior, 125, pp. 106963.

Latham, G. P. and Locke, E. A. (1991) 'Self-regulation through goal setting', Organizational behavior and human decision processes, 50(2), pp. 212-247.

Liu, J., Li, C., Carcioppolo, N. and North, M. (2016) 'Do Our Facebook Friends Make Us Feel Worse? A Study of Social Comparison and Emotion', Human Communication Research, 42(4), pp. 619-640.

Locke, E. A. and Latham, G. P. (2002) 'Building a practically useful theory of goal setting and task motivation. A 35-year odyssey', Am Psychol, 57(9), pp. 705-17.

Locke, E. A., Shaw, K. N., Saari, L. M. and Latham, G. P. (1981) 'Goal setting and task performance: 1969–1980', Psychological Bulletin, 90(1), pp. 125.

Lyubomirsky, S. and Ross, L. (1997) 'Hedonic consequences of social comparison: a contrast of happy and unhappy people', Journal of personality and social psychology, 73(6), pp. 1141.

Marginson, D. E. W. (2004) 'Chapter 19 - The Case Study, The Interview and The Issues: A Personal Reflection', in Humphrey, C. and Lee, B. (eds.) The Real Life Guide to Accounting Research. Oxford: Elsevier, pp. 325-337.

McDaid, E., Andon, P. and Free, C. (2023) 'Algorithmic management and the politics of demand: Control and resistance at Uber', Accounting, Organizations and Society, pp. 101465.

McDaid, E., Boedker, C. and Free, C. (2019) 'Close encounters and the illusion of accountability in the sharing economy', Accounting, Auditing & Accountability Journal, 32(5), pp. 1437-1466.

McMillan, D. W. (1996) 'Sense of community', Journal of community psychology, 24(4), pp. 315-325.

McMillan, D. W. 2011. Sense of community, a theory not a value: A response to Nowell and Boyd. Wiley Online Library.

McMillan, D. W. and Chavis, D. M. (1986) 'Sense of community: A definition and theory', Journal of Community Psychology, 14(1), pp. 6-23.

Merchant, K. and Van der Stede, W. (2017) Management Control Systems 4th Edition : Performance Measurement, Evaluation and Incentives. Harlow, United Kingdom, UNITED KINGDOM: Pearson Education, Limited.

Merchant, K. A. and Manzoni, J.-F. (1989) 'The Achievability of Budget Targets in Profit Centers: A Field Study', The Accounting Review, 64(3), pp. 539-558.

Miri, D. H. and Macke, J. (2021) 'Gamification, motivation, and engagement at work: a qualitative multiple case study', European Business Review.

Mussweiler, T. and Rüter, K. (2003) 'What friends are for! The use of routine standards in social comparison', Journal of Personality and Social Psychology, 85(3), pp. 467.

Newman, A. H. and Tafkov, I. D. (2014) 'Relative performance information in tournaments with different prize structures', Accounting, Organizations and Society, 39(5), pp. 348-361.

Oprescu, F., Jones, C. and Katsikitis, M. (2014) 'I PLAY AT WORK—ten

principles for transforming work processes through gamification', Frontiers in Psychology, 5, pp. 14.

Repenning, N., Löhlein, L. and Schäffer, U. (2021) 'Emotions in Accounting: A Review to Bridge the Paradigmatic Divide', European Accounting Review, 31(1), pp. 241-267.

Richter, G., Raban, D. R. and Rafaeli, S. (2015) 'Studying gamification: The effect of rewards and incentives on motivation', Gamification in education and business: Springer, pp. 21-46.

Ryan, R. M. and Deci, E. L. (2017) Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Publications.

Salmons, J. E. (2015) Doing qualitative research online. Sage.

Schedlinsky, I., Schmidt, M. and Wöhrmann, A. (2020) 'Interaction of information and control systems: How the perception of behavior control affects the motivational effect of relative performance information', Accounting, Organizations and Society, 86, pp. 101171.

Shah, P. P. (1998) 'Who are employees' social referents? Using a network perspective to determine referent others', Academy of Management Journal, 41(3), pp. 249-268.

Shang, R., Abernethy, M. A. and Hung, C.-Y. (2020) 'Group identity, performance transparency, and employee performance', The Accounting Review, 95(5), pp. 373-397.

Smith, R. H. (2000) 'Assimilative and contrastive emotional reactions to upward and downward social comparisons', Handbook of social comparison: Springer, pp. 173-200.

Tafkov, I. D. (2013) 'Private and Public Relative Performance Information under Different Compensation Contracts', Accounting Review, 88(1), pp. 327-350.

Van den Bussche, P. and Dambrin, C. (2020) 'Peer-to-peer evaluations as narcissistic devices: fabricating an entrepreneurial community', Accounting, Auditing & Accountability Journal.

Van den Bussche, P. and Morales, J. (2019) 'Performing (in) the community: accounting, biopower and the sharing economy', Handbook of the Sharing Economy: Edward Elgar Publishing.

Webb, R. A., Williamson, M. G. and Zhang, Y. (2013) 'Productivity-Target Difficulty, Target-Based Pay, and Outside-the-Box Thinking', The Accounting Review, 88(4), pp. 1433-1457.

Yu, A. (2021) 'Accountability as mourning: Accounting for death in the time of COVID-19', Accounting, Organizations and Society, 90, pp. 101198.

Zichermann, G. and Cunningham, C. (2011) Gamification by design: Implementing game mechanics in web and mobile apps. " O'Reilly Media, Inc.".

Tables and Figures

Table 1 Rider Interview Details

Interview Details			
Name	Phase 1 (November 2020)	Phase 2 (April 2021)	Phase 3 (April 2022)
Rider 1	34min 30sec	39min	
Rider 2	40min 00sec	1 hour	43min 20sec
Rider 3	48min 20sec	45min	53min 38sec
Rider 4	44min 25sec		30min 48sec
Rider 5	24min 49sec	44min	
Rider 6	24min 13sec		
Rider 7	37min 40sec		
Rider 8	46min 23sec		
Rider 9	30min 15sec		
Rider 10	38min 00sec		
Rider 11	29min 30sec		
Station Manager 1		58min 50sec	
Rider 12			66min 50sec
Rider 13			35min 44sec
Rider 14			35min 54sec
Rider 15			42min 30sec
Rider 16			39min 30sec
Rider 17			58min
Rider 18			54min 05sec
Rider 19			50min 50sec
Rider 20			45min 21sec
Rider 21			41min 01sec
Rider 22			57min 40sec
Rider 23			42min 10sec
Rider 24			75min 07sec
Rider 25			31min 44sec
Rider 26			44min 51sec
Rider 27			65min 51sec
Rider 28			40min 29sec

Table 2 Summary of two kinds of riders

Category	PT	FT
Working Time Per Day	No requirement	8 hours fixed working time
Wage Payment	Daily by the mobile application	Monthly by stations
Price Per Order	Different depending on the distance and time	Fixed at a certain level and changing after reaching a certain level
Training process	Online training in the mobile application	Formal training in the station
The roles of Station Managers from City Partners	1. Only make announcements in the riders' mobile applications. 2. Prepare some gifts in the station for PT to collect during the traditional festival.	1. Hold morning meetings every day with FTs. 2. Help with riders' daily delivery work and responsible for riders' physical and mental health
Services Area	The Whole city or district	3 km or 6 km areas (depends on different area)
Order Assignment	1. Assigned by the system. 2. Pick up orders from the system by riders	Assigned by the system and the station manager can help to adjust some insensible assigned orders.
Right to Rejection Orders Assigned by the system	Y (But if reject too many times, the account will be temporarily blocked for a few hours to 1 day as a penalty)	Y (Depending on the delivery situation station managers will make some adjustments to help riders)
Penalty	Mild (around 5 yuan per bad review depending on the different cities, the rules are slightly different)	Harsh (around 50 yuan per bad review, depending on the different stations from different city partners, the rules are different)
Reasons to be a rider (PT or FT)	1. Low-educated and skilled 2. Temporary financial pressures (Mortgage) 3. Failure of running start-ups	
Reasons to choose a certain kind of rider	1. Introduction from friends and relatives 2. Pursuing formal training from stations 3. Personality (Strong or Weak Self-disciplinary) 4. Working habit (Prefer freedom of the working habit or Certain working time) 5. Wage concern (Prefer the quick payment and unlimited but unstable number of orders or the monthly payment and limited but stable number of orders)	

Figure 1 Codes in Nvivo for Paper 1
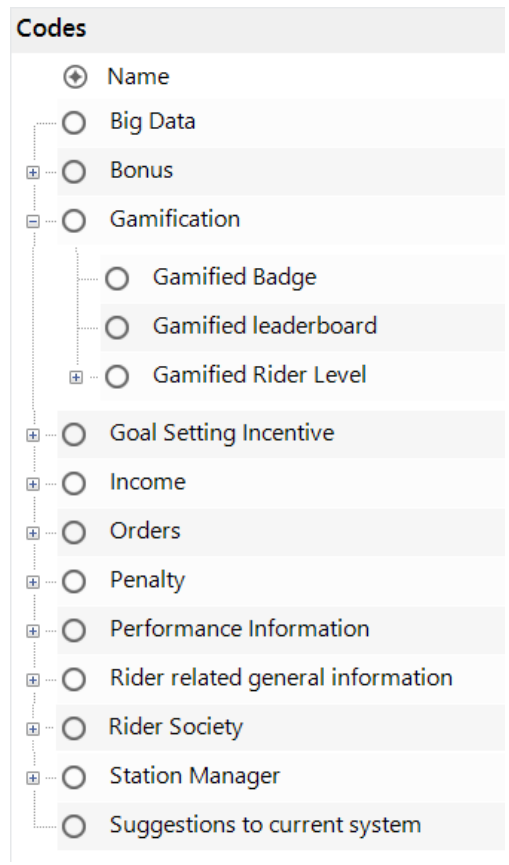
Figure 2 Rider Gamifed Leaderboards
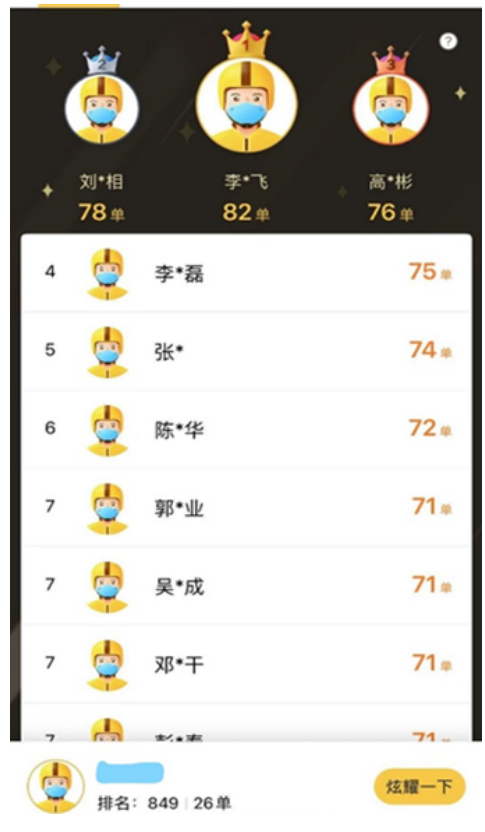
Figure 3 Discussion Framework
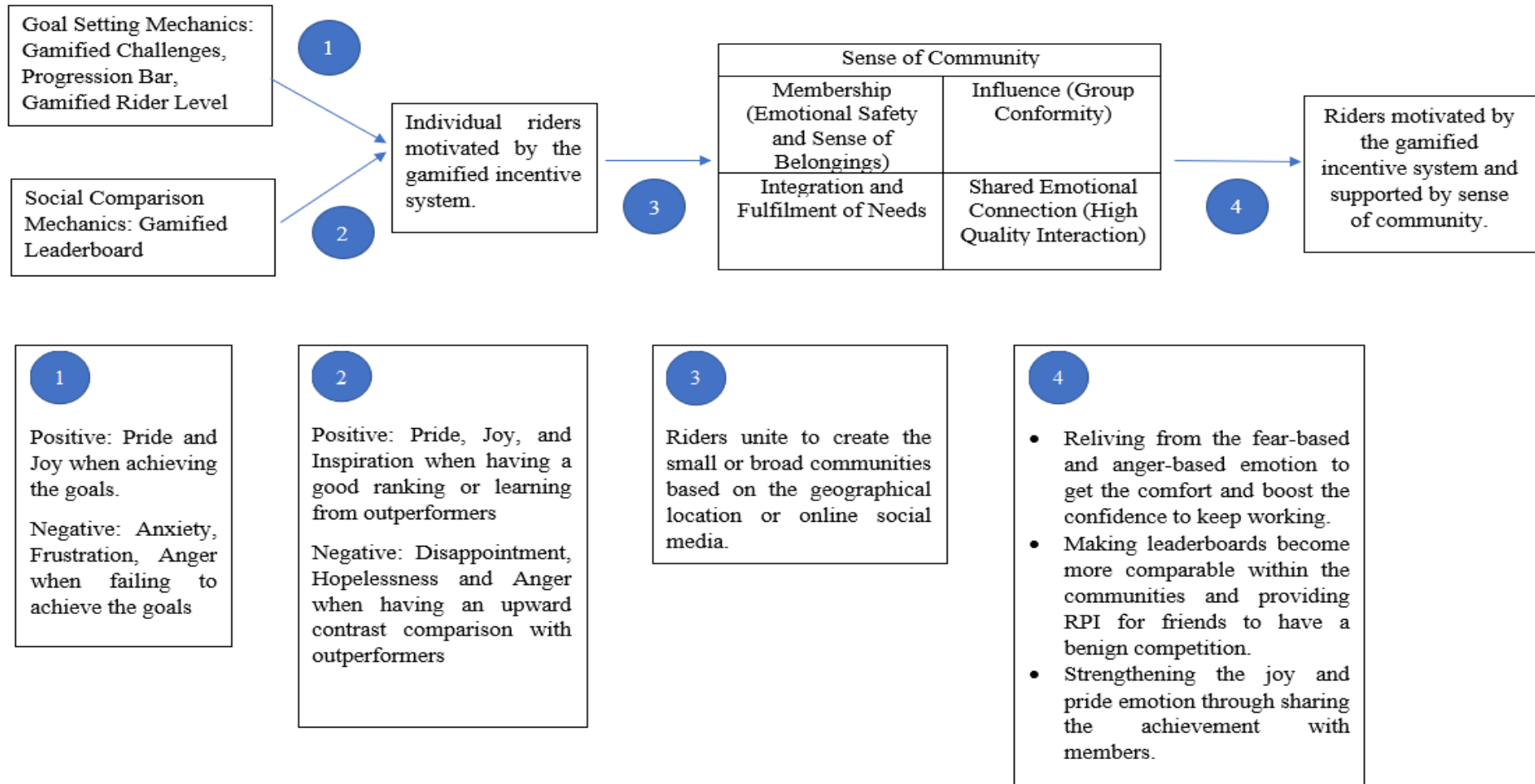
Paper 2

We are different: personality traits and the use of real-time relative performance information

We are different: personality traits and the use of real-time relative performance information

Abstract

Prior management accounting studies on relative performance information have mainly examined its direct effect on individuals' performance using experimental research. Motivational effects and learning effects are mainly mentioned and examined. Shedding light on the personality characteristics and personal beliefs, we examine the moderation effect of social comparison orientation (SCO) on the direct effect of perceived usefulness or use frequency of real-time relative performance information (RRPI) on performance, and the indirect mediation effect of occupational self-efficacy (OSE). The results from a survey in an on-demand food delivery rider setting using a gamified incentive system suggest that the higher usefulness of or higher use frequency of RRPI has a positive effect on riders' performance. For the moderation effect of SCO, from the use frequency perspective, we find that riders with high SCO delivered more orders in a specific month. As for the indirect effect of OSE, OSE has a significant and positive effect between the perceived usefulness of RPI and the monthly orders delivered. In an additional analysis, we find that the riders' familiarity to their work, which is measured by their gamified rider level in the system, influences their attitudes to RRPI and use frequency of RRPI.

3.1 Introduction

A central question in management accounting is how information about individual performance influences employee actions. This is the case as performance feedback potentially has learning and motivational effects on individuals' productivity (Bonner and Sprinkle, 2002). With the development of information technology, more frequent and detailed information about employees' performance is available. According to Cappelli and Tavis (2016), more than one-third of U.S. companies such as General Electric and Microsoft changed their traditional annual review to a more frequent digital real-time feedback system. However, Casas‑Arce, Lourenço and Martínez‑Jerez (2017) suggest that the use of real-time information about employee performance negatively influences managerial decision-making, as real-time information could potentially lead to focus on short-term factors.

Regarding the amount and type of performance information available, Eyring, Ferguson and Koppers (2021) show that the provision of relative performance information (thereafter

RPI) has a better motivational effect than the joint provision of absolute performance information (thereafter *API*) and RPI. Recently, in a field experiment setting, Kohler *et al.* (2023) find that detailed RPI may result in employees' neglect of low-return tasks and a reduction of overall local performance. In this paper, we extend this literature focusing on the role of employee personal traits and their effect on the role of real-time RPI feedback systems on employee productivity/performance.

Extant literature examines the effect of the provision of performance information and the interaction effects with different incentive systems (Sprinkle, 2000; Hannan, Krishnan and Newman, 2008; Tafkov, 2013; Newman and Tafkov, 2014; Eyring, Ferguson and Koppers, 2021). However, several authors argue that performance evaluation systems should consider employees' personality traits since they can influence the effect of performance feedback (Holderness, Olsen and Thornock, 2016; Wang, 2017; Buser, Gerhards and Van Der Weele, 2018). For example, Wang (2017) examines the moderation effect of the role of the *dark triad* of personality traits (Machiavellianism, narcissism, and psychopathy) on the provision of RPI.

To extend these findings, we first focus on the role of social comparison orientation (thereafter SCO)—a personality trait that captures "the extent to which people compare with others and in the way that comparison information is interpreted" (Buunk and Gibbons, 2007, p. 13)—to examine its moderation effects on workers' performance. Although it is an obvious characteristic of workers, SCO influences how workers react to RPI and influences their further behaviours and performance. Holderness, Olsen and Thornock (2016) argue that the designers of performance evaluation systems need to consider personality traits as they can influence the effect of performance feedback. However, the SCO of workers seems been overlooked in previous RPI literature.

Second, we use self-efficacy theory (Bandura, 1999) to further examine the mediation effect of workers' occupational self-efficacy between RPI and individuals' performance. According to self-efficacy theory, vicarious learning is one of the main mechanisms to boost people's self-efficacy. Through learning from others or observing others' performance, people will feel more capable of doing certain things. As public RPI gives workers information about what others achieved, this is an important source of vicarious learning to improve self-efficacy and thus performance. Based on this potential mediation effect from self-efficacy theory, using a domain-specific occupational self-efficacy (thereafter OSE) construct, we examine how RPI influences workers' performance through perceived OSE. Therefore, our paper focuses on whether SCO moderates the effect of RPI on individuals' performance and provides evidence on the mediating role of OSE in this relationship.

To address these questions, we use responses from a survey of riders of an on-demand food delivery platform setting in China. Not only motivated by our interest in psychological theories but this study is also motivated by our field study in the same setting (Chapter 2). In particular, this setting offers a rich environment of different gamified elements in a gamified incentive system. Some of the game elements presented on riders' mobile applications are related to social comparison and RPI, such as real-time rankings, leaderboards and gamified levels. Delivery riders receive real-time RPI in their daily and monthly leaderboards. The leaderboards show the top 50 riders' performance and the rider's own rank and performance, which is the number of orders they delivered. In that study, we find that riders have different attitudes to the social comparisons and the gamified leaderboards. To further understand the effect of RPI from gamified leaderboards, we further conduct this survey study. We use structural equation modelling to test our hypothesis, via AMOS 28 based on a survey of 1,000 current delivery riders serving in a platform company in China.

In line with previous RPI literature, our results show that the direct effect of the usefulness of RPI on monthly orders is statistically significant and positive. Similarly, we document a positive effect of RPI on drivers' service reviews. When we focus on the role of SCO for the perceived usefulness of RPI measure, we don't find a significant moderation effect on performance. In terms of OSE, we find that it has a positive indirect effect from the perceived usefulness of RPI to the number of monthly orders delivered via OSE and a negative indirect effect from the usefulness of RPI to the efficiency measure via OSE.

As for the use frequency of RPI, consistent with the usefulness of RPI, we document that the direct effect on the total number of monthly orders and the efficiency measure is positive and significant. We also find a negative and significant effect of the use frequency of RPI on the number of bad reviews. However, the effect on the total number of estimated average orders per month is negative but weakly significant. For the moderating role of SCO, there is a positive moderation effect of SCO on the number of monthly orders. In terms of the mediating role of OSE, we don't find a significant effect of OSE between the use frequency of RPI and different performance measures.

Additionally, we find that riders' perceived usefulness of RPI reduced with the upgrade of their gamified rider levels, which means they are more familiar with this delivery job. The use frequency of RPI is changed in an inverted-U shape. At the initial stage, riders tend to check their RPI more often. When they reach certain levels, their frequency of checking their RPI from leaderboards is reduced. This is relevant because when the riders get more familiar with their work they tend to access their RRPI less often and their attitude to RRPI is also changing.

We contribute to the literature and practice in several ways. First, following previous literature on RPI (Wang, 2017; Thomas and Thornock, 2017), our study extends our understanding of the effect of RPI on employee performance by focusing on personality characteristics, more specifically, SCO. Showing how this personal orientation influences the RPI effect, our study provides important insights into how organisations can use incentive systems more effectively, based on different employees' personalities to generate better incentives for individual employees. Second, we contribute to performance information and incentive systems literature (Gist and Mitchell, 1992; Bonner and Sprinkle, 2002; Hall, 2008; Burney and Widener, 2013) providing evidence of an important mediator--perceived OSE-- between the RPI and riders' performance. This is relevant as perceived OSE is important for employees to feel that they are capable of completing assigned tasks. By enhancing employees' self-efficacy in their jobs through providing RPI, organizations could foster employees' better performance.

Additionally, as previous research (Kohler *et al.*, 2023; Eyring and Narayanan, 2018) shows that only about 26-30% of people access the RPI that is provided to them, our study gives one explanation of this situation that the familiarity of current work will influence workers' the access and the attitude to RPI, especially on this repeatable and labour-intense work. Finally, according to Ittner (2014), some qualitative mechanisms such as interviews can enhance studies' causal inference. Motivating by another field study in the same setting (Chapter 2), based on our field interviews, it is helpful for us to make and enhance the potential association inferences originally from the psychology theory in this rider setting, though it is hard to claim the pure causal effect of our findings in a survey study.

In the next part, we develop our hypotheses based on psychological theories and previous literature. We then introduce our survey research method. The fourth part is for the main result and additional analysis. In the final part, we concluded this paper with potential contributions and future research avenues.

3.2 Literature and Hypothesis Development
3.2.1 Real-time Relative Performance Information and Performance
Information about self-performance is central to determining the level of exerted effort by individuals. Previous literature shows that there are two main uses of performance information. From a learning perspective, performance information assists individuals in revising their beliefs about their effort and performance, allowing more sensible decisions to improve productivity. From a motivational perspective, performance information is connected

to incentive contracts designed to foster better performance (Atkinson, Waterhouse and Wells, 1997). Both learning and motivation uses of performance information are linked to positive effects on an individual's performance (Ammons, 1956; Sprinkle, 2000; Bonner and Sprinkle, 2002).

Following the development of information technology, more detailed and frequent performance information can be provided to employees (Cappelli and Tavis, 2016). However, this increment in the detail and frequency of feedback is likely to produce counterproductive effects. Casas-Arce, Lourenço and Martínez-Jerez (2017) conduct a field experiment in a business outsourcing company and find evidence that employees in the detailed but infrequent (monthly) feedback group have better performance with respect to the more frequent (weekly) group. More frequent feedback causes employees to focus more on recent performance and ignore other performance perspectives. In a multi-step task decision-quality feedback setting, Thornock (2016) also finds an inverse U-shape relationship between feedback timeliness and performance. In particular, slightly delayed feedback (after participants implemented their decision) is more useful than immediate feedback before decision implementation. However, long-delayed feedback negatively affects learning from the feedback.

Extant literature shows the importance of individuals' own performance information and feedback on their performance. According to social comparison theory (Festinger, 1954), individuals tend to evaluate themselves by comparing with others, especially with relevant peers. Information on peers' performance (RPI) has a significant impact on individual efforts and performance in two ways. First, social comparison can trigger competition to affect individuals' behaviour (Festinger, 1954). As social comparison drives individual motives for self-enhancement, in order to get a better self-image individuals compete to outperform relevant peers (Garcia, Reese and Tor, 2020). Providing RPI to individuals can activate competition to motivate them to exert more effort to outperform their peers. Especially from public RPI, such as rankings, having a good rank in the leaderboard will satisfy individuals' vanity and improve their status among peers. Therefore, individuals are likely to exert more effort to improve their rankings. Previous research confirms this positive competition effect of RPI (Hannan, Krishnan and Newman, 2008; Tafkov, 2013; Hartmann and Schreck, 2016; Eyring, Ferguson and Koppers, 2021). Second, social comparison generates a learning effect. Individuals have self-improvement motives on top of the self-evaluation motives for competition (Buunk and Gibbons, 2007). Through social comparison with outperformers, individuals find inspiration in superior performance and learn from their peers to improve themselves. Therefore, individuals spend more effort narrowing the performance gaps with respect to outperformers. However,

social comparison can serve individuals' self-improvement conditional on the perceived controllability of their relative performance (Carroll and Marginson, 2021).

In line with the arguments of social comparison theory (Festinger, 1954), the gamified leaderboards in our setting are likely to generate a competitive motivational effect among riders. Delivery riders in the same station or those who usually deliver orders in the same area[14] compare with each other through the ranking on the leaderboards. These riders will spend more effort to deliver more orders to improve their rankings and status among closer peers, satisfying their self-enhancement needs. Simultaneously, leaderboards show riders how many orders the top 50 riders deliver daily or monthly in a given station. According to our field interviews, we found that through this public RPI in their leaderboards, riders can realise the differences in performance between top performers and the rest of the riders. Moreover, some riders can find out and learn from top peers in their station by asking for advice or tips to improve. Top riders' performance on the leaderboard also gives peers a potentially achievable goal to aim for. Combining the previous literature and our field interview findings, we use two measures to analyse the effects of RRPI on riders' performance, namely, the perceived usefulness of RRPI and the use frequency of RRPI.

> ***H1a: The perceived usefulness of RRPI has positive effects on riders' performance.***

> ***H1b: The use frequency of RRPI has positive effects on riders' performance.***

3.2.2 Social comparison orientation moderation effects

According to Buunk and Gibbons (2007), social comparison theory has evolved from different perspectives. One of the most important perspectives is individual differences in social comparison, suggesting that whether people tend to make comparisons with others may be a characteristic of their personality (Hemphill and Lehman, 1991). Although some individuals like to compare with peers to self-evaluate, enhance, or improve themselves, others might feel less inclined to do so. Therefore, it is important to acknowledge individuals' differences in their willingness to engage in social comparison (Hemphill and Lehman, 1991, p. 390). Although Festinger (1954) suggests that social comparison is innate, the extent and the frequency of comparisons with others varies from one individual to another (Buunk *et al.*, 2019).

---

[14] There are mainly two types of riders in our sample. Full-time riders (FT) deliver orders only in a 3km to 6km square root area around their station. Riders in a station normally know each other, and their leaderboard only ranks riders' performance in that station. The average number of riders in a station is lower than 100. Part-time riders (PT) deliver orders around the whole city. Their leaderboard includes thousands of part-time riders' performance in the city.

To assess this personality trait, Gibbons and Buunk (1999) develop scales of social comparison orientation (SCO), measuring the extent to which people compare with others and the way that comparison information is processed. Previous studies have shown that individuals with higher levels of SCO compare with others more often. For example, individuals high in SCO spend more time checking others' performance and have more interest in discussing results with others (Gibbons and Buunk, 1999). Similarly, in a healthcare context, nurses with high SCO make more upward and downward comparisons with their peers (Buunk *et al.*, 2003).

Extant literature shows that higher SCO is correlated with personality characteristics (Buunk and Gibbons, 2007; Buunk *et al.*, 2019). First, individuals with higher SCO are self-focused, as they tend to conform to and fulfil their social and personal expectations through social comparison. Moreover, these individuals also possess higher interpersonal orientation, as they have stronger interests in what other people feel and are more sensitive to how other people behave and the needs of others. As a result of this focus on others, they are more easily influenced by peers. Finally, individuals with high SCO often have higher levels of neuroticism (Gibbons and Buunk, 1999; Buunk *et al.*, 2019), leading to higher self-uncertainty and interest in reducing it. Through comparisons with others, individuals feel temporarily more certain about themselves, which helps to create a sense of security, serving their self-evaluation needs (Festinger, 1954).

In line with the previous arguments, Holderness, Olsen and Thornock (2016) argue that the designers of performance evaluation systems should consider personality traits as they can influence the effect of performance feedback. However, there is limited research considering the individual differences of participants. Buser, Gerhards and Van Der Weele (2018) find that differences in feedback responsiveness, as a personal trait, are important in explaining participants' beliefs and decisions. Thomas and Thornock (2017) document that the influence of RPI on freeriding behaviour is moderated by individuals' social value orientation. Wang (2017) finds that the provision of RPI can generate productive or counterproductive effects conditional on individuals' scores on the dark triad of personality traits (Machiavellianism, narcissism, and psychopathy). In line with these findings, we argue that the individual personal trait directly linked to social comparison, SCO, is important to moderate the effect between RPI and individuals' performance. In particular, we expect that a higher individual's SCO is likely to be associated with a more often use of RPI for social comparison purposes, and the more useful she will consider RPI.

In our setting, riders are provided with the same gamified leaderboards to check their own ranking and top 50 riders list, as well as the number of orders they have delivered. The

SCO of each rider should influence riders' interest in checking the leaderboard rankings, and their perceived usefulness of this RRPI. Riders with higher levels of SCO are likely to care more about their peers' performance. However, they might feel uncertain about the number of orders they can fulfil in future. Providing other riders' performance information should to some extent help riders to increase their feelings of certainty about expected targets, as this performance information reassures them about realistic expectations about targets and monetary rewards they can get in future. Therefore, we expect that the positive effect of RPI on performance will be stronger if riders have a higher SCO. Here is our second hypothesis:

> *H2a: The positive effect of the perceived usefulness of RRPI on performance is larger for the riders with high social comparison orientation.*
>
> *H2b: The positive effect of the use frequency of RRPI on performance is larger for the riders with high social comparison orientation.*

3.2.3 Self-efficacy Mediate effects

We next focus on the role of self-efficacy theory to explain the mediation effect of occupational self-efficacy (OSE) between RPI and riders' performance. Bandura (1997, p. 3) defines self-efficacy as the own belief in the individual capabilities to organise and execute alternative courses of action required to meet objectives. Therefore, self-efficacy explains how cognitive factors influence individuals' effort and performance. Individuals' perceived self-efficacy on tasks is linked to the amount of effort individuals want to expend on a particular task and how long they can persist when they face difficulties during the process (Bandura, 1978).

Individual self-efficacy can be mainly affected by three sources of information. Performance accomplishment is the first and most significant source of self-efficacy, as individuals obtain the most authentic experiential evidence about their ability to achieve their objective in a specific task which, in turn, increases their capabilities for subsequent tasks (Bandura, 1997). When individuals experience success, this increases their self-efficacy (Peterson and Arnn, 2005). Second, verbal persuasion or feedback received from others can sustain individuals' capabilities to believe that they can achieve the target. Finally, vicarious learning plays an important role in boosting an individual's self-efficacy. Individuals can increase their self-efficacy by observing their peers' success and learning from their experiences. When individuals observe their relevant peers achieving a particular performance level, they

persuade themselves to achieve at least some improvement in their performance as well (Bandura and Barab, 1973).

Relevant to social comparison (Festinger, 1954), Bandura and Jourden (1991) argue that social comparison operates as a primary factor in the self-appraisal of capabilities because there are no objective and non-social standards for most activities to gauge the level of ability. Individuals need to evaluate their capabilities by comparing them with others' performance. Bandura and Jourden (1991) find when the participants continued to close the gap between themselves and outperformed the group and eventually surpassed outperformers, their perceived self-efficacy increased to the highest level compared with other manipulated social comparison contexts. However, when participants are in a manipulated comparison context where they are like the counterparties at the beginning but surpassed by their peers, and the gap becomes wider and wider, participants' perceived self-efficacy keeps decreasing.

In this study, we focus on the role of vicarious learning on self-efficacy. In particular, RPI is one of the most important external information signals for individuals to evaluate their self-efficacy. Accordingly, individuals can find outperforming peers as models to learn from and improve their self-efficacy. In our setting, riders have real-time leaderboards showing the Top 50 riders' performance, allowing riders to find their role models. From our field interview findings, within stations, riders can observe their models' working behaviour to identify how to deal with different complex circumstances, as well as ask for tips and recommendations. Therefore, we expect that perceived self-efficacy positively influences individuals' performance through direct and indirect mechanisms. From a direct perspective, self-efficacy helps individuals self-regulate their effort duration, direction, intensity, and strategy development to finish tasks (Bonner and Sprinkle, 2002; Burney and Widener, 2013). Individuals with high self-efficacy believe that if they put more effort in the right direction, they can succeed. This belief also provides emotional and mental power to tackle negative emotions and overcome obstacles and difficulties during the process.

From an indirect perspective, self-efficacy influences individuals' performance through its relationship with goal setting (Bandura and Jourden, 1991). When selecting goals, individuals with high self-efficacy tend to choose more challenging objectives. However, individuals with low self-efficacy are less likely to choose goals beyond their perceived capability. High self-efficacy increases commitment to goals, as well as determination to achieve them. Similarly, high self-efficacy also influences individuals to set more challenging goals themselves. As Locke and Latham (1991) described, there is a high-performance cycle: when having high commitment, high self-efficacy, adequate feedback, and proper strategies,

individuals achieve high performance even on difficult tasks. When individuals are satisfied with the reward for their high performance, their commitment to the task and self-efficacy are likely to increase further.

Providing relevant outperforming models' performance information is a significant strategy to enhance self-efficacy (Gist and Mitchell, 1992). We expect that by providing RPI, riders' perceived OSE will increase. An increase in OSE is likely to motivate riders to achieve better performance by increasing their efforts in their jobs directly and influencing their commitments to goals indirectly. More specifically, riders with high self-efficacy should believe they can deliver more orders based on their previous performance and comparisons with peers, thus putting more effort into their jobs. At the same time, they are likely to choose to reach higher assigned goals (or set higher personal goals) and keep working persistently until they achieve those goals. These riders will feel more capable of dealing with the obstacles and difficulties with strategies they have acquired before (Stajkovic and Luthans, 1998; Gist and Mitchell, 1992; Burney and Widener, 2013). Therefore, we hypothesise that riders' OSE is a mediator between the perceived usefulness of RRPI or use frequency of RRPI and their performance[15]:

> ***H3a: Riders' occupational self-efficacy mediates the relationship between the perceived usefulness of RRPI and riders' performance.***
> ***H3b: Riders' occupational self-efficacy mediates the relationship between the use frequency of RRPI and riders' performance.***

3.3 Sampling and Methodology:

3.3.1 Riders' survey

Digital labour platforms have a significant influence on modern life. Culpepper and Thelen (2019) argue that through cultivating platform loyalty and providing better support, platform services become part of the infrastructure of their consumers. Not only do platform organisations change customers' lives, but they also create gig work opportunities. Schwellnus *et al.* (2019) estimate that the share of gig economy platform employment ranges between 1 and 3 per cent of total employment in developed economies, and this proportion is increasing fast. Within this gig economy employment, on-demand delivery platforms contribute significantly to this growth. On-demand delivery platforms only provide software and data

---

[15] Although there is potential positive effect of self-efficacy on the frequency use of RRPI, based on the field interviews, most of experienced participants actually don't check the leaderboard more often as they care more about their own performance information than RRPI on the leaderboards.

analytics, whilst outsourcing other aspects of their business (Srnicek, 2017). In the Western world, popular examples of such platforms are Deliveroo and Uber Eats. Delivery riders are considered independent contractors that use the mobile application provided by platform companies to supply services to customers.

Given the emergence of these on-demand delivery platforms, we choose the online food delivery riders setting in China to conduct our survey for several reasons. First, the mobile applications used by riders provide RRPI in a gamified way to motivate riders' delivery work, which allowed us to examine the direct effect of the perceived usefulness of RPI on performance for gig economy riders in platform organisations. Second, based on our interviews with delivery riders, different riders have different opinions and attitudes about these RRPI. Our field study further enhanced our hypotheses from the psychological theories that people's personality traits influenced the relationship between the perceived usefulness of performance information and performance.

We use online survey questionnaires to gather our data. We designed our questionnaire including constructs used in previous accounting and psychological studies, demographic questions, and objective performance measures questions. We rely on the *back translation method* (Brislin, 1980)[16] to translate the English version of the questionnaire into Chinese. To verify that the questions are understandable and answerable, we performed a pilot test with 20 riders and requested feedback on the design of the questions. These riders provided us with suggestions on how to frame some specific questions in the local language more properly, reducing ambiguity.

After finalising our pilot test, we hired a professional survey company to contact 1,000 riders who fit our requirements and collect data from these individuals. We chose a survey company that has extensive experience with local academic institutions and corporations, as well as being in the same province where the sample was collected. In terms of the survey's participants, we imposed the following conditions: first, riders should be working in two Chinese cities that have not been significantly influenced by COVID restrictions during our sample collection period. Second, these riders should be working mainly for a specific platform. Third, these riders' current main income sources should be from the on-demand delivery job. These three requirements were included as questions before the formal questionnaire to eliminate unqualified riders and to confirm that every response was from the target respondent.

---

[16] The back translation has three steps: firstly, we translate original questionnaires into Chinese. Secondly, we compare the translation versions and original versions of questionnaires. Thirdly, we reconcile the meaningful differences between two questionnaires.

During the data collection period, we stayed in touch with the survey company to check the sample results. We used information from field interviews with riders and some online field materials to help the staff from the survey company to eliminate some potential outliers, i.e. if a rider fills in unattainable performance data, that data point was eliminated. Simultaneously, the survey company also has some rules to control their data quality. For example, if participants finished the questionnaire below a certain time threshold, which means these participants do not fill the questionnaire carefully, these samples were treated as invalid samples. Finally, we also have restrictions in terms of attention to the answers, e.g., if participants choose the same Likert choices for all items across the different constructs or choose the Likert choices in a certain order, these samples are invalid.

Based on those requirements, we obtained a sample of 1,000 valid responses from two cities. Table 1 shows the sampling situations and response rate. The survey firm sent out a total of 7,882 links. 2,117 unqualified participants stopped filling out the questionnaires after the three qualifying questions, which means they are not from the two target cities we required, they are not mainly serving the targeted platform, or they have some other full-time jobs as their main income sources. 333 participants did not finish the questionnaires, and 233 invalid samples were cleaned based on invalid answers such as unattainable performance data or strange Likert choice sequences and quick responses which means finished the questionnaire in an unbelievably short time. After dealing with these different situations during the sampling, the final sample of 1,000 valid responses rate is 12.7%.

[Insert Table 1 about here]

From demographic perspectives, the full-time (part-time) riders are 75.3% (24.7%)[17]. In terms of gender, 92.6% are male and 7.4% are female. Although in recent years the proportion of female riders has increased, it is still a minor proportion with respect to the total population of riders. As for age, 6.5% of the surveyed riders are below 21 years old. Most riders are located in the 21 to 40-year-old bracket, about 81.4% of the sample. In terms of tenure, we find that more than half of the respondents (57.6%) have been working as delivery riders for about 11 to 20 months. Only 7.3% of riders have been working for over 31 months in our sample. Finally, 58.8% of riders are migrants from other cities, whereas the rest of the sample is from the two target cities. Table 2 includes the demographic information in detail. These demographic statistics reflect similar situations in our field interviews and online short videos

---

[17]Part-time is only the name of different kinds of full-time riders with a more flexible working model and time. It doesn't mean that these riders have other full-time jobs. On-demand delivery job is the only main source of income for all 1000 riders.

shared on riders' own social media.

[Insert Table 2 about here]

We also collected riders' gamified levels in our questionnaire[18], which reflects their familiarity with this delivery work. The distribution of levels in our survey is shown in Table 3. For full-time riders, only 8.2% of riders are at silver or bronze knight level; 25.2% of riders are at the highest possible level for full-time riders, Glory Knight. The largest distribution is for Pioneer Knight, which comprises about 27% of full-time riders. For part-time riders, more than 40% of riders are in the Strongest Hero level, whereas only 7% are Shining Star or below. 12.6% of part-time riders are in the Supreme God of War category in our sample, the highest possible level for this group.

[Insert Table 3 about here]

3.3.2 Variable Measurement

For the constructs we used, we followed the constructs used in the previous management accounting and psychology literature (Mahlendorf, Kleinschmit and Perego, 2014; Rigotti, Schyns and Mohr, 2008; Gibbons and Buunk, 1999). As our data is from single-source participants, we reviewed thoroughly the construct item design and question orders of the questionnaire to avoid the common method bias (Podsakoff *et al.*, 2003). We conducted a pilot test before collecting data formally to help us adjust some framing of the questions, which improves the face validity and understandability of the survey. Moreover, we use very general terms in our questionnaire ("Rider Information", "About You", "About your Work"), to make sure we haven't suggested any relationship between questions. We also informed riders that there were no right or wrong answers and that we would want to hear their real opinion about each question. Riders also knew that their answers were fully anonymous and only used for our research purposes. These steps are in line with the suggestions by Speklé and Widener (2018). Similarly, as our hypotheses include the test of some moderation effects, the existence of an interaction term also reduces some of the common method biases.

Finally, we ask for objective metrics in the survey to measure performance such as total monthly orders delivered, monthly drive distance, number of bad reviews, and total estimated

---

[18] FT gamified level order: Bronze Knight, Silver Knight, Golden Knight, Pioneer Knight, Elite Knight, Glory Knight. PT gamified level order: Novice Rider, Shining Star (1,2,3), Glory Elite (1,2,3), Strongest Hero (1,2,3), Supreme God of War. Consistent with our interviews, the riders' level distribution is right skewed as riders don't need a long time to reach the top level. For example, normally, if a part-time rider is hard-working for about 10 hours a day, they can be the Supreme God of War in 3 to 4 months.

average order per month. In the questionnaire, we prompt the riders to check their mobile application when filling in this information. After conducting Harman's one-factor test for common method issues, there is little evidence that our constructs suffer from this (Podsakoff and Organ, 1986). Although it is difficult to eliminate all sources of common biases in a survey study, we expect that these measures above should overcome major common method bias issues.

All the latent constructs we use in our research are well-established constructs from previous accounting or psychology literature. We follow the guidelines of management accounting research to perform a relevant reliability and validity check (Bedford and Speklé, 2018; Collier, 2020). To test the construct validity, we use factor loadings to calculate the composite reliability and Cronbach's alpha. Following Fornell and Larcker (1981), we calculate the AVE to test the convergent validity of each construct. By comparing the square root of AVE with the correlation between constructs, we test the discriminant validity of our constructs. We use AMOS 28 to do the confirmatory factor analysis for our measurement model. The result shows that the model is well-fit, with $\chi2$ /df = 5.819, CFI = 0.916, NFI = 0.901, TLI = 0.904, RMSEA = 0.069, and SRMR = 0.0517. All the factor loadings, reliability measures and AVE are shown in Table 4.

[Insert Table 4 about here]

Our main independent variables for the performance system in our model are *RRPI_USEFUL* and *RRPI_FREQ.* The usefulness of real-time relative performance information *(RRPI_USEFUL)* captures the perceived usefulness of RRPI provided by the gamified mobile application. We measure *RRPI_USEFUL* with 4 items adapted from Mahlendorf, Kleinschmit and Perego (2014). The construct is measured by using a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree). All the item factor loadings and AVE are larger than 0.5, which shows that the convergent validity is satisfied. Both Cronbach's alpha (0.874) and Composite Reliability (0.885) show strong reliability of the construct.

The Use Frequency of relative performance information *(RRPI_FREQ)* measures the use frequency of real-time relative performance by asking riders how frequently they check the gamified leaderboard ranking. The frequent scales are 1 = Every rest after peak-time period; 2= Daily; 3 = Weekly; 4=Monthly; 5 = Less than once every month. We reversely code the results before we do our analysis which makes the larger number represent the high frequency. We find that on average riders normally check their leaderboard rankings every day.

For personality beliefs and traits, we use occupational self-efficacy *(OSE)* and social comparison orientation *(SCO).* In terms of measuring self-efficacy, Bandura (1977) argues that

specifying the task being addressed is crucial to the assessment of self-efficacy. Although different self-efficacy constructs were developed for different tasks, these task-specific constructs make it hard to compare individuals in different organisations or different departments of the same organisation. To address this limitation, Rigotti, Schyns and Mohr (2008) develop the domain-specific *OSE* 6-item construct for organisational studies, as a proxy of the competence that a person feels concerning the ability to successfully fulfil tasks involved in her organizational role. This instrument has been tested in different language and cultural contexts (Rigotti, Schyns and Mohr, 2008; Peng and Zhang, 2021). We use this domain-specific self-efficacy construct in our model to examine the riders' perceived self-efficacy in their delivery job. According to the suggestion from Bandura (2006), we use a 0-100 slide bar to let riders decide how true it is for 6 items. Although factor loadings for two items and overall AVE do not reach above 0.5, given the good constructs reliability (Cronbach's alpha = 0.808 and Composite Reliability = 0.790), we still decide to keep all the items in the OSE for our analysis (Van der Hauwaert *et al.*, 2022).

SCO captures the social comparison orientation of each rider. We use 11-item social comparison orientation constructs developed by Gibbons and Buunk (1999) to measure these personality characteristics. Riders were asked to what extent they agree with these 11 arguments using a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). To increase the validity of this construct and the survey, two items in this construct are reversely framed. We reversely code the two items before any analysis. All the items' factor loadings and AVE are larger than 0.5, which satisfies the convergent validity. Both Cronbach's alpha (0.914) and composite reliability (0.917) show strong reliability of the construct.

Finally, for our dependent variables of performance measures, we have four measures to measure their performance[19] from different perspectives: first, *MON_ORDER* represents how many orders they delivered in a specific month. We can find that on average each rider delivers 1050 orders in a specific month. Secondly, *BAD_REVIEW* measures how many bad reviews they received this month[20]. On average, riders get 7 bad reviews in our sample. These

---

[19] We notice that these performance measures can mainly reflect big cities with developed on-demand delivery services. However, the results of this study can be generalised to other cities or even other gig economy settings. As for how to define a "good" or "bad" performance, it is mainly from a platform perspective that better performance means deliver more orders and get fewer bad reviews. We realised that different riders or even a same rider in different days might have different definitions for good or bad performance. For the survey purpose, we can only take this platform perspective as a single criterion to measure riders' performance.

[20] Based on the riders' real-time performance of each order, customers can give riders bad reviews for different reasons such as rude services, order delay, the spill-out of the food or drink, or even customers abreaction of its on bad emotion. The number of bad reviews will influence the riders' order assignment by the platform.

two measures exist in riders' applications so that they can easily check for the results[21]. We also let riders estimate the average number of orders they can deliver each month, which is measured as *AVE_ORDER*. The average number of orders is around 1,035. Based on the first three measures, we calculate another two measures. The final performance measure is *QUALITY*, which is calculated using *BAD_REVIEW* divided by *MON_ORDER*. The average bad review ratio is 0.7% in our sample. All these average performance statistics matched with riders' situation in the field study (Chapter 2).

The model includes several control variables. We control the tenure of the rider *(TENURE)* as this delivery job does have its learning curve. As the number of orders delivered increases, drivers become more familiar with different routes and communities in the cities, thus improving the speed of their tasks. We also include controls for riders' seniority and familiarity, using as a proxy the gamified rider level *(LEVEL)* since this indicates accurately the recent seniority of riders as the rider level is updated monthly. Third, we include two city-specific controls for two related to the city: a city-fixed effect *(CITY)[22]* and *(HOMETOWN)*, a dummy variable capturing whether riders are locals or migrants (Local = 1, Migrant = 0). Finally, we also control for two demographic effects: one is a gender *dummy (GENDER)* (Male = 1, Female = 0) and the rider's age *(AGE_CAT)*. To protect respondents' privacy, we only ask riders to choose from five age categories in the questionnaire instead of filling in the actual age. Descriptive statistics of the variables of interest in our study are presented in Table 5 and the correlation matrix of our variable of interest is in Table 6.

To check the discriminant validity, we present the correlation matrix in Table 6. According to Fornell and Larcker (1981), as the square roots of AVE of our constructs are higher than bivariate correlations, the discriminant validity of our constructs is satisfied. Meanwhile, through the correlation table, we find that there are no potential multicollinearity issues within our variables given the low bivariate correlations. We find that *RRPI_USEFUL* is not significantly correlated with OSE, but *RRPI_FREQ* is correlated with OSE. Moreover, *RRPI_USEFUL* is significantly correlated to every performance measure in our model.

---

[21] We ask riders to fill in the previous month performance measures which are shown on their mobile applications. As the current month performance is changing in real-time, we cannot get the accurate current month performance data.

22 We realise that there are different areas in the same city where the number of orders varies, which could potentially influence the performance of riders. However, we didn't collect this information in our questionnaire initially because when we started to collect our data our sample includes both FT and PT kinds of riders. However, for PT riders, they delivered the orders across the city, which is hard to control the specific areas where they are serving. Therefore, we didn't control the area factors for the whole sample.

However, we don't find a correlation between *RRPI_FREQ* and MON_ORDER.

[Insert Table 5 and Table 6 about here]

3.4 Results

3.4.1 Main Results

We use structured equation modelling software AMOS 28 to perform our multivariate analyses. As Full-time riders and Part-time riders use different gamified rider level systems, our main base model results, focus on 753 full-time rider observations to control effectively for the rider type effect. We first use the simple path model, which treats every latent variable as an observed variable. We then use a mixed structural model for the robustness test, which includes each indicator for latent variables in our analysis. For the mixed structural model, all model fit statistics such as $\chi 2$ /df, comparative fit index (CFI) and the root mean square error approximation (RMSEA), show a good model fit. According to Collier (2020), we use in both models composite variables for our moderators of SCO formed by the average of the moderators' indicators and the mean-centred interaction term between SCO and *RRPI_USEFUL* or *RRPI_FREQ*. The mixed structural model confirms the findings from the path models.

[Insert Table 7 and Table 8 about here]

We report in Table 7 and Table 8 the standardised path coefficients, which are based on bootstrapping using 5,000 bootstrap samples. From both *RRPI_USEFUL* and *RRPI_FREQ*, our models largely support H1, which is consistent with some previous RPI literature. *RRPI_USEFUL* is positively related to *MON_ORDER* and it is also positively related to *AVE_ORDER* but not statistically significant. There is a negative relationship between *RRPI_USEFUL* and *BAD_REVIEW* and *QUALITY*, which is the same as what we expect. The RRPI tends to help riders have fewer bad reviews. As for *RRPI_FREQ* results, reported in Table 8, most of the direct effects are similar to *RRPI_USEFUL*, which has a significant positive relationship with *MON_ORODER* and a negative relationship with *BAD_REVIEW* and *QUALITY*. Except for a weak negative relationship between *AVE_ORDER* and *RRPI_FREQ*, other performance measures support our expectations.

For H2 we test the moderation effect of *SCO*. For *RRPI_USEFUL*, Table 7 shows that there is no significant moderation effect of SCO on different performance measures in our FT rider sample, though the signs of coefficients are the same as what we expected. Table 8 shows that the effect of *RRPI_FREQ* on *MON_ORDER* is positively moderated by *SCO*. We interpret this result as evidence that the positive effect of *RRPI_FREQ* on *MON_ORDER* is larger for

riders with higher *SCO*. However, for other performance measures, we don't find statistically significant results for the moderation effect of *SCO*.

For H3, we estimate the indirect effect between RRPI and our performance measures through *OSE*. Table 7 shows that *OSE* has a partial positive mediation effect between *RRPI_USEFUL* and *MON_ORDER* with marginal weak significance. Through checking the public RRPI on the leaderboards, riders felt more capable of delivering more orders and thus finished more orders. Meanwhile, the indirect path from *RRPI_USEFUL* to *QUALITY* via *OSE* is negative and significant, which also supports our H3a. Table 7 also shows that the total effect of *RRPI_USEFUL* on different performance measures is statistically significant except for *AVE_ORDER*. In Table 8, we don't find any significant mediation effect of OSE between *RRPI_FREQ* and performance measures. It reflects that checking the leaderboards more frequently does not help to have better performance through increasing riders' self-efficacy. The total effect of *RRPI_FREQ* on performance is also statically significant for different performance measures. As an example, Table 9 and Table 10 show the detailed results of the models with all control paths and model fit statistics for the *MON_ORDER* performance measure.

[Insert Table 9 and Table 10 about here]

3.4.2 Additional Analysis

When we do a robustness check using the PT rider sample, in Table 11, although the signs of the results meet our expectations and are similar to the results in the main tests using the FT sample, for H1a and H3a, the direct associations or indirect effects through *OSE* between *RRPI_USEFUL* and different performance measures are not statistically significant. However, the interaction effects of H2a are statistically significant for all performance measures. This reflects that in this PT rider sample the positive (negative) effect of *RRPI_USEFUL* on *MON_ORDER* and *AVE_ORDER* (*BAD_REVIEW* and *QUALITY*) is higher (lower) for riders with High *SCO*.

In Table 12, for *RRPI_FREQ*, when we test our model using the PT rider sample, for H1b, the signs of the coefficients are consistent with our expectations. However, they are not statistically significant. For H2b, we find a negative moderation effect of *SCO* on *AVE_ORDER* with a weak significance, which is opposite to what we expected. However, we document a negative moderation effect of SCO on *BAD_REVIEW* and *QUALITY* for PT with high *SCO*, which means they will have fewer bad reviews during the delivery. For H3b, the same as we

find in the main FT sample, we don't find any significant mediation effect of *OSE* between the use frequency of RRPI and different performance measures in the PT sample.

[Insert Table 11 and Table 12 about here]

Among the control variables, we find riders' gamified levels have a significant control effect on the riders' performance. Higher-level riders deliver more orders in a specific month, and they also have a higher estimation of the average number of orders they can deliver. In terms of bad reviews, higher-level riders have relatively more bad reviews. Since they deliver more orders, the probability of getting bad reviews could also increase because of the subjectivity of giving a bad review by customers[23]. Based on our interviews with experienced riders, most of them explained that this situation doesn't indicate that high-level riders will make more mistakes during the delivery. It probably suggests that high-level riders will meet more complicated situations or have a higher probability of meeting immoral customers, given they are the priority of the algorithm to assign more orders.

Given the importance of gamified levels and the mixed results from the main analysis, we provided more detailed analyses on how different levels of riders have different behaviours and attitudes to RRPI and finally influence their performance. According to the findings from the interviews in Chapter 2 of this thesis, high-level riders relatively do not draw much attention to RRPI compared with the freshers. We conduct a regression analysis to test whether this is the case. We use the two measures of RRPI as our dependent variables and *LEVEL* as our independent variables. After controlling different demographic variables, Table 13 and Table 14 confirm the findings from the field paper. We find that *LEVEL* has a significantly negative association with the *RRPI_USEFUL*, which indicates that higher-level riders thought RRPI is less useful than the lower-level riders thought in our sample. We also add the quadratic terms of rider levels to see whether based on the change of the riders' gamified levels riders' attitudes to RRPI also have an inverted-U change. From the result, the quadratic terms of rider level are not statistically significant. With regard to *RRPI_FREQ*, there is a negative association between *LEVEL* and *RRPI_FREQ*. Meanwhile, we also find that the quadratic term of rider level is significant and negative, which means the use frequency of RRPI increases when riders upgrade their levels at the beginning of the gamified level system. And then to some point, riders reduce

---

[23] Although there is a potential counterargument is that riders who want to deliver as many orders as possible would not really care about customers feedback, given the delivery riders need to guarantee their services quality to get the wages from each order in our setting, it is hard for riders to fake the orders themselves. Meanwhile, if their misconducts are found by the system, their rider account will be blocked and cannot work for certain amount of time depending on the situations.

their check frequency when riders reach a certain level. This exactly fits the field findings that at the beginning, some freshers with a low rider level don't check the leaderboards quite often because it is hard for them to get to the Top 50 leaderboards. With their experience and rider level increasing, they started to check more often to see whether they could achieve certain performances. Finally, it is sensible for riders to reduce their check frequency of RRPI when they know their work limit and feel that the RRPI cannot motivate them anymore, as we find that the *RRPI_USEFUL* is negatively associated with riders' level.

[Insert Table 13 and Table 14 about here]

We further break down the whole FT groups into different gamified levels to explore the influence of RRPI on riders' performance. In Table 13, we find that when riders' level becomes higher from Golden to Elite, the influence of *RRPI_USEFUL* on *MON_ORDER* becomes larger first. When riders reach to Elite level, the influence of RRPI starts to reduce. The influence on *MON_ORDER* is weaker for Glory level riders, though the coefficient is not statistically significant. We can understand the direct effect of *RRPI_USEFUL* on MON_ORDER in the main tests is largely driven by the Pioneer and Elite riders. In Table 14, we find a similar pattern for *RRPI_FREQ*. When *LEVEL* goes up, although there is a slight fluctuation of the influence on *MON_ORDER*, the overall trend is also increasing from Golden to Elite. However, when riders become Glory riders, the influence of *RRPI_FREQ* on *MON_ORDER* is significantly negative. Still, the main direct effect of *RRPI_FREQ* on *MON_ORDER* is driven by the Golden, Pioneer, and Elite riders in our sample. From both RRPI measures, the influence of RRPI on performance does change with the change in riders' levels. To some extent, this effect also coincided with riders' attitudes and the use behaviour of RRPI. When they feel that RRPI become less useful or when they check less frequently, the influence of RRPI on their performance will also become weaker.

[Insert Table 15 and Table 16 about here]

3.5 Conclusion

In this study, we draw on the social comparison theory and self-efficacy theory to test the effect of RPI on riders' performance. With the development of the original social comparison theory (Festinger, 1954), social psychologists found that although everyone does social comparison every day, there are differences in the orientation to compare with other people. As management accounting researchers, it is intriguing to explore whether this orientation of social comparison moderates the effect of RPI on individuals' working performance. Meanwhile, according to Bandura (1978), one of the main sources of increasing individuals' self-efficacy is vicarious learning. RPI provides important feedback for individuals

to know others' performance and helps to boost individuals' self-efficacy through vicarious learning channels. Therefore, we predicted the indirect effect of RPI on performance via OSE, which is a specialised construct for self-efficacy in the work setting.

We first confirm the positive effect of RPI on performance. Using the two measures of RPI (the usefulness of RPI and the use frequency of RPI), we find that there is a direct significant effect of RPI on riders' monthly order delivered, bad review numbers and quality measures in our main sample. Secondly, considering different riders' SCO, we find that from the use frequency perspective, riders with high SCO have a stronger influence on monthly order performance compared to riders with low SCO. We do not find any significant moderation effect of SCO on performance for the perceived usefulness of RPI. Finally, for the indirect effect via OSE, we find that there is a positive mediation effect between the monthly order delivered and the perceived usefulness of RPI. Moreover, the mediation effect is negative for the quality measure via OSE, which shows riders' improvement of its delivery quality through the path of OSE increase. However, we do not find any mediation effect of OSE on riders' performance from the use frequency of RRPI.

To further understand the results in our main tests, we did an additional analysis on how the gamified level system linked to the riders' thoughts and user behaviours of RRPI and influenced their performance. We find that with the upgrade of the gamified level, riders felt that the perceived usefulness of RRPI reduces. Meanwhile, the use frequency is changed in a reversed-U shape that increases initially, and riders lose interest in checking when they become more sophisticated. We also find that the influence of RRPI on the riders' monthly number of orders increases until riders reach the Elite level. Then the influence of RRPI on monthly order performance starts to reduce. This reflects that the influence of RRPI will change based on the increase in riders' experience and familiarity with their work. Therefore, not only do we find some evidence of the personality effect of SCO and self-efficacy mediation effect, but also we document that the experience or familiarity of this certain job significantly influences riders' attitudes to RRPI and the use behaviour of RRPI, and thus plays a crucial role in influencing the riders' performance.

As a survey study, it does have some limitations: first of all, the data we collect is from single-source participants, which results in the consideration of common measures bias. Although we cannot get rid of all the biases, we do have control to limit this bias to a minimum from the original design of the questionnaire to the pilot test with delivery riders. We also did Harman's single-factor test to make sure that there are no obvious common bias issues in our model. All the constructs have high reliability and acceptable validity. Secondly, as we only

collect riders' level cross-sectional data in one period to test our model, we can only claim the association between variables instead of the casual effect. In addition, some existing control variables such as City fixed effect might have the limitation to control the influence on the performance of riders based on the different areas where they are located in cities. Although we cannot test the exact causal effect, our findings are supported by psychological theories and our field interview findings, which enhanced the internal validity of our study (Ittner, 2014). Finally, given our data and results come from an on-demand food delivery riders' setting, the generalisability might be a concern for our research. However, as the platform organisations largely influenced people's way of work and life, it can be generalised to other gig economy workers' settings. Moreover, the psychological theories tested in this study are also generalisable to future settings and studies.

Although our study has inherent limitations as a survey study, it does have academic contributions and practical implications. Firstly, we contribute to the stream of RPI literature (Hannan, Krishnan and Newman, 2008; Tafkov, 2013; Newman and Tafkov, 2014; Eyring, Ferguson and Koppers, 2021) to consider the most obvious but overlooked individuals' personality differences in social comparison. Future experiment studies and empirical studies can further test the moderation effect of SCO on performance in different settings. Moreover, as an important incentive, how to provide RPI to different employees needs to be considered in detail. Organisations need to learn about their employees' SCO and customise the RPI to some extent to generate a better incentive effect.

Secondly, contributing to previous incentive literature (Hall, 2008; Burney and Widener, 2013), we investigate the importance of occupational self-efficacy to employees. Through the vicarious learning channel, RPI helps workers to increase their self-efficacy in their work. With the increase in employees' self-efficacy of their work from their peers' performance, employees will feel more capable of dealing with the difficulties in their work and achieve better performance. Future research can further explore how different incentives can increase employees' or workers' occupational self-efficacy and thus improve their performance.

Moreover, in additional analysis, we find that the experience and familiarity of certain work influenced the attitude and use behaviour of RPI. As previous research (Kohler *et al.*, 2023; Eyring and Narayanan, 2018) shows that only about 26-30% of people access RPI that is provided to them, our study gives one explanation for this situation when people are familiar with their current work and focus on their own performance they won't pay much attention to RPI. Future research can continue to explore what factors make people want to access their RPI. Finally, as a potential minor contribution, we use a novel way to collect the survey data in

management accounting literature. By recruiting a professional survey company, our data collection process becomes more efficient. As it is already a data collection approach used widely in management and psychology research, scholars in management accounting can also cooperate with these survey companies to collect survey data for our research in the future.

3.6 References

Ammons, R. B. (1956) 'Effects of knowledge of performance: A survey and tentative theoretical formulation', The Journal of General Psychology, 54(2), pp. 279-299.

Atkinson, A. A., Waterhouse, J. H. and Wells, R. B. (1997) 'A stakeholder approach to strategic performance measurement', MIT Sloan Management Review, 38(3), pp. 25.

Bandura, A. (1977) 'Self-efficacy: Toward a unifying theory of behavioral change', Psychological Review, 84(2), pp. 191-215.

Bandura, A. (1978) 'Self-efficacy: Toward a unifying theory of behavioral change', Advances in Behaviour Research and Therapy, 1(4), pp. 139-161.

Bandura, A. (1997) Self-Efficacy: The Exercise of Control. Worth Publishers.

Bandura, A. (1999) 'Self-Efficacy: The Exercise of Control', Journal of Cognitive Psychotherapy, 13, pp. 158 - 166.

Bandura, A. (2006) 'Guide for constructing self-efficacy scales', Self-efficacy beliefs of adolescents, 5(1), pp. 307-337.

Bandura, A. and Barab, P. G. (1973) 'Processes governing disinhibitory effects through symbolic modeling', Journal of abnormal psychology, 82(1), pp. 1.

Bandura, A. and Jourden, F. J. (1991) 'Self-regulatory mechanisms governing the impact of social comparison on complex decision making', Journal of personality and Social Psychology, 60(6), pp. 941.

Bedford, D. S. and Speklé, R. F. (2018) 'Construct validity in survey-based management accounting and control research', Journal of Management Accounting Research, 30(2), pp. 23-58.

Bonner, S. E. and Sprinkle, G. B. (2002) 'The effects of monetary incentives on effort and task performance: theories, evidence, and a framework for research', Accounting Organizations and Society, 27(4-5), pp. 303-345.

Brislin, R. W. (1980) 'Translation and content analysis of oral and written materials', Methodology, pp. 389-444.

Burney, L. L. and Widener, S. K. (2013) 'Behavioral work outcomes of a strategic performance measurement system-based incentive plan', Behavioral Research in Accounting, 25(2), pp. 115-143.

Buser, T., Gerhards, L. and Van Der Weele, J. (2018) 'Responsiveness to feedback as a personal trait', Journal of Risk and Uncertainty, 56(2), pp. 165-192.

Buunk, A. P. and Gibbons, F. X. (2007) 'Social comparison: The end of a theory and the emergence of a field', Organizational Behavior and Human Decision Processes, 102(1), pp. 3-21.

Buunk, A. P., Gibbons, F. X., Dijkstra, P. and Krizan, Z. (2019) 'Individual differences in social comparison', Suls., J., Collins, RL, Wheeler, L.(Eds.), Social Comparison, Judgment, and Behavior, pp. 77-104.

Buunk, B. P., Zurriaga, R., Gonzalez-Roma, V. and Subirats, M. (2003) 'Engaging in upward and downward comparisons as a determinant of relative deprivation at work: A longitudinal study', Journal of Vocational Behavior, 62(2), pp. 370-388.

Cappelli, P. and Tavis, A. (2016) 'The performance management revolution', Harvard Business Review, 94(10), pp. 58-67.

Carroll, E. and Marginson, D. (2021) 'Relative performance information and social comparisons: Exploring managers' cognitive, emotional and dysfunctional behavioral processes', Management Accounting Research, 53.

Casas-Arce, P., Lourenço, S. M. and Martínez-Jerez, F. A. (2017) 'The performance effect of feedback frequency and detail: Evidence from a field experiment in customer satisfaction', Journal of Accounting Research, 55(5), pp. 1051-1088.

Collier, J. E. (2020) Applied structural equation modeling using AMOS: Basic to advanced techniques. Routledge.

Culpepper, P. D. and Thelen, K. (2019) 'Are We All Amazon Primed? Consumers and the Politics of Platform Power', Comparative Political Studies, 53, pp. 288 - 318.

Eyring, H., Ferguson, P. J. and Koppers, S. (2021) 'Less Information, More Comparison, and Better Performance: Evidence from a Field Experiment', Journal of Accounting Research, 59(2), pp. 657-711.

Eyring, H. and Narayanan, V. G. (2018) 'Performance Effects of Setting a High Reference Point for Peer-Performance Comparison', Journal of Accounting Research, 56(2), pp. 581-615.

Festinger, L. (1954) 'A Theory of Social Comparison Processes', Human Relations, 7(2), pp. 117-140.

Fornell, C. and Larcker, D. F. 1981. Structural equation models with unobservable variables and measurement error: Algebra and statistics. Sage Publications Sage CA: Los

Angeles, CA.

Garcia, S. M., Reese, Z. A. and Tor, A. (2020) 'Social comparison before, during, and after the competition'.

Gibbons, F. X. and Buunk, B. P. (1999) 'Individual differences in social comparison: development of a scale of social comparison orientation', Journal of Personality and social psychology, 76(1), pp. 129.

Gist, M. E. and Mitchell, T. R. (1992) 'Self-efficacy: A theoretical analysis of its determinants and malleability', Academy of Management review, 17(2), pp. 183-211.

Hall, M. (2008) 'The effect of comprehensive performance measurement systems on role clarity, psychological empowerment and managerial performance', Accounting, organizations and society, 33(2-3), pp. 141-163.

Hannan, R. L., Krishnan, R. and Newman, A. H. (2008) 'The effects of disseminating relative performance feedback in tournament and individual performance compensation plans', Accounting Review, 83(4), pp. 893-913.

Hartmann, F. and Schreck, P. (2016) 'Rankings, Performance, and Sabotage: The Moderating Effects of Target Setting', European Accounting Review, 27(2), pp. 363-382.

Hemphill, K. J. and Lehman, D. R. (1991) 'Social comparisons and their affective consequences: The importance of comparison dimension and individual difference variables', Journal of Social and Clinical Psychology, 10(4), pp. 372.

Holderness, D. K., Olsen, K. J. and Thornock, T. A. (2016) 'Who Are You to Tell Me That?! The Moderating Effect of Performance Feedback Source and Psychological Entitlement on Individual Performance', Journal of Management Accounting Research, 29(2), pp. 33-46.

Ittner, C. D. (2014) 'Strengthening causal inferences in positivist field studies', Accounting, Organizations and Society, 39(7), pp. 545-549.

Kohler, M., Mahlendorf, M. D., Seiter, M. and Vogelsang, T. 'Social Comparison on Multiple Tasks–Sacrificing Overall Performance for Local Excellence'. 2022.

Locke, E. and Latham, G. (1991) 'A Theory of Goal Setting & Task Performance', The Academy of Management Review, 16.

Mahlendorf, M. D., Kleinschmit, F. and Perego, P. (2014) 'Relational effects of relative performance information: The role of professional identity', Accounting Organizations and Society, 39(5), pp. 331-347.

Newman, A. H. and Tafkov, I. D. (2014) 'Relative performance information in tournaments with different prize structures', Accounting, Organizations and Society, 39(5), pp. 348-361.

Peng, F. and Zhang, L.-W. (2021) 'The Relationship of Competitive Cognitive Anxiety and Motor Performance: Testing the Moderating Effects of Goal Orientations and Self-Efficacy Among Chinese Collegiate Basketball Players', Frontiers in Psychology, pp. 1965.

Peterson, T. O. and Arnn, R. B. (2005) 'Self‐efficacy: The foundation of human performance', Performance Improvement Quarterly, 18(2), pp. 5-18.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. and Podsakoff, N. P. (2003) 'Common method biases in behavioral research: a critical review of the literature and recommended remedies', Journal of applied psychology, 88(5), pp. 879.

Podsakoff, P. M. and Organ, D. W. (1986) 'Self-reports in organizational research: Problems and prospects', Journal of Management, 12(4), pp. 531-544.

Rigotti, T., Schyns, B. and Mohr, G. (2008) 'A short version of the occupational self-efficacy scale: Structural and construct validity across five countries', Journal of Career Assessment, 16(2), pp. 238-255.

Schwellnus, C., Geva, A., Pak, M. and Veiel, R. (2019) 'Gig economy platforms: Boon or Bane?'.

Speklé, R. F. and Widener, S. K. (2018) 'Challenging issues in survey research: Discussion and suggestions', Journal of Management Accounting Research, 30(2), pp. 3-21.

Sprinkle, G. B. (2000) 'The effect of incentive contracts on learning and performance', Accounting Review, 75(3), pp. 299-326.

Srnicek, N. (2017) Platform capitalism. John Wiley & Sons.

Stajkovic, A. D. and Luthans, F. (1998) 'Self-efficacy and work-related performance: A meta-analysis.', Psychological Bulletin. Sep98, 124(2), pp. 240.

Tafkov, I. D. (2013) 'Private and Public Relative Performance Information under Different Compensation Contracts', Accounting Review, 88(1), pp. 327-350.

Thomas, T. F. and Thornock, T. A. (2017) 'Me vs. We: The Effect of Incomplete Team Member Feedback on Cooperation of Self-Regarding Individuals Tyler', SSRN Electron. J.

Thornock, T. A. (2016) 'How the timing of performance feedback impacts individual performance', Accounting, Organizations and Society, 55, pp. 1-11.

Van der Hauwaert, E., Hoozée, S., Maussen, S. and Bruggeman, W. (2022) 'The impact of enabling performance measurement on managers' autonomous work motivation and performance', Management Accounting Research, 55, pp. 100780.

Wang, L. W. (2017) 'Recognizing the Best: The Productive and Counterproductive Effects of Relative Performance Recognition', Contemporary Accounting Research, 34(2), pp.

966-990.

Appendix: Variable Description

Variable	Definition
RRPI_USEFUL	The perceived usefulness of real-time relative performance information 4-item construct on a 5-point Likert scale. Adapted from Mahlendorf, Kleinschmit and Perego (2014).
RRPI_FREQ	The use frequency of real-time relative performance information measures how often riders check their leaderboard ranking information The frequency scales are: 1 = Every rest after peak-time period; 2= Daily; 3 = Weekly; 4=Monthly; 5 = Less than once every month. We reversely coded the result before any analysis.
OSE	Occupational Self-efficacy 6-items construct on a 1-100 slide bar. Adapted from Rigotti, Schyns and Mohr (2008) and Bandura (2006).
SCO	Social Comparison Orientation 11-item construct on a 5-point Likert scale. Adapted from Gibbons and Buunk (1999).
MON_ORDER	The number of delivery orders riders finished in May 2022.
BAD_REVIEW	The number of bad review riders received in May 2022.
AVE_ORDER	The average number of delivery orders riders estimated per month.
QUALITY	The number of bad reviews divided by the number of delivery orders in May 2022.
GENDER	The dummy variable is one for male riders and zero for female riders.
AGE_CAT	Age categories of riders: 1 = Below 21, 2 = 21 ~ 30, 3 = 31~ 40, 4 = 41 ~ 50, 5 = Above 50.
HOMETOWN	The dummy variable is one for local riders and zero for migrant riders.
TENURE	How many months they have been a delivery rider.
CITY	The dummy variable is one for riders from GZ and zero for riders from SZ.
LEVEL	The gamified rider levels were provided by the platform in May.

Tables and Figures

Figure 1 Theoretical Model
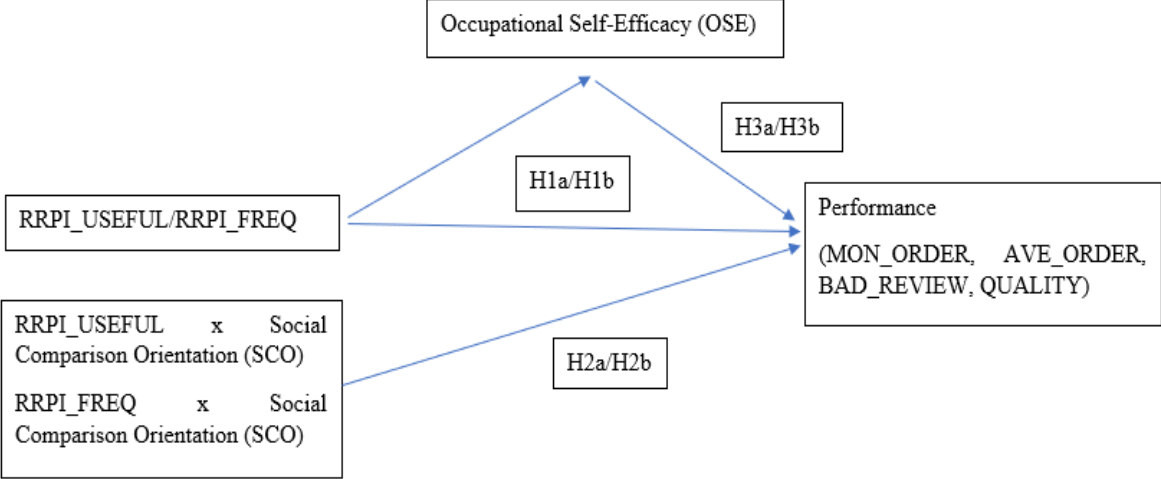
Table 1 Sampling Situation and Response Rate

	GZ		SZ		All	
	N	%	N	%	N	%
Participation Link Sent	3677		4205		7882	
Valid Link Click	1706		1977		3683	
Not Qualified Participants	973		1144		2117	
Unfinished Questionnaires	92		241		333	
Gross Sample	641	17.4%	592	14.1%	1233	15.6%
Invalid Answers and Fast Answers	141		92		233	
Final Valid Sample	500	13.6%	500	11.9%	1000	12.7%

Notes: Table 1 reports the sampling situation. The non-qualified participants are the participants who failed to answer three qualifying questions before the real questionnaire. The invalid answers and fast answers are the participants who fill the questionnaire too fast or who fill the invalid answers such as the same Likert choice for all items across the different constructs.

Table 2 Respondents Demographic Statistics (N = 1,000)

	GZ		SZ		All	
	N	%	N	%	N	%
Panel A: Rider Categories						
FT	400	80.0%	353	70.6%	753	75.3%
PT	100	20.0%	147	29.4%	247	24.7%
Panel B: GENDER						
Male	467	93.4%	459	91.8%	926	92.6%
Female	33	6.6%	41	8.2%	74	7.4%
Panel C: TENURE (Month Range)						
0-10	131	26.2%	75	15.0%	206	20.6%
11-20	223	44.6%	353	70.6%	576	57.6%
21-30	93	18.6%	52	10.4%	145	14.5%
31-40	44	8.8%	19	3.8%	63	6.3%
41-50	9	1.8%	1	0.2%	10	1.0%
Panel D: AGE_CAT						
Below 21	21	4.2%	44	8.8%	65	6.5%
21 ~ 30	158	31.6%	194	38.8%	352	35.2%
31 ~ 40	260	52.0%	202	40.4%	462	46.2%
41 ~ 50	59	11.8%	54	10.8%	113	11.3%
Above 50	2	0.4%	6	1.2%	8	.8%
Panel E: HOMETOWN						
Local	213	42.6%	199	39.8%	412	41.2%
Migrant	287	57.4%	301	60.2%	588	58.8%

Notes: Table 2 reports the demographic characteristics of riders. All variables are explained in the Appendix. We grouped the tenure of riders in this table, but in the analysis, we directly used the exact number of months as a control variable.

Table 3 Rider Gamified Level Statistics (N = 1,000)

| | PT Rider Level | | | | | | FT Rider Level | | | | | |
| | GZ | | SZ | | ALL | | | GZ | | SZ | | ALL | |
Level	N	%	N	%	N	%	Level	N	%	N	%	N	%
Novice Rider	2	2.0			2	0.8	Bronze Knight	8	2.0	1	0.3	9	1.2
Shining Star 1	3	3.0			3	1.2	Silver Knight	40	10.0	13	3.7	53	7.0
Shining Star 2	3	3.0	2	1.4	5	2.0	Golden Knight	105	26.3	73	20.7	178	23.6
Shining Star 3	4	4.0	4	2.7	8	3.2	Pioneer Knight	84	21.0	116	32.9	200	26.6
Glory Elite 1	5	5.0	24	16.3	29	11.7	Elite Knight	51	12.8	72	20.4	123	16.3
Glory Elite 2	2	2.0	21	14.3	23	9.3	Glory Knight	112	28.0	78	22.1	190	25.2
Glory Elite 3	8	8.0	29	19.7	37	15.0							
Strongest Hero 1	28	28.0	43	29.3	71	28.7							
Strongest Hero 2	10	10.0	16	10.9	26	10.5							
Strongest Hero 3	7	7.0	5	3.4	12	4.9							
Supreme God of War	28	28.0	3	2.0	31	12.6							
Total	100	100%	147	100%	247	100%	Total	400	100%	353	100%	753	100%

Notes: Table 3 reports the gamified level for riders provided by riders' mobile application, which represents the current seniority of riders in this month and the familiarity of this job.

Table 4 Loadings of construct variables (N = 1,000)

Items	Loadings	Cronbach's alpha	Composite Reliability	Average Variance Extracted
RRPI_USEFUL		0.874	0.885	0.659
Item 1	0.834			
Item 2	0.710			
Item 3	0.776			
Item 4	0.914			
OSE		0.808	0.790	0.397
Item 1	0.465			
Item 2	0.752			
Item 3	0.461			
Item 4	0.754			
Item 5	0.509			
Item 6	0.748			
SCO		0.914	0.917	0.503
Item 1	0.762			
Item 2	0.690			
Item 3	0.686			
Item 4	0.700			
Item 5 (R)	0.843			
Item 6	0.622			
Item 7	0.743			
Item 8	0.683			
Item 9	0.679			
Item 10	0.673			
Item 11 (R)	0.698			

Notes: Table 4 reports the measurement model for latent variables used in structured equation modelling. The measurement model is well-fitting with $\chi2$ /df = 5.819, CFI = 0.916, NFI = 0.901, TLI = 0.904, RMSEA = 0.069, and SRMR = 0.0517. After Conducting Harman's one-factor test for common method issues, there is little evidence that our constructs have common method issues (Podsakoff and Organ, 1986).

Table 5 Descriptive Statistics for Variables (N = 1,000)

Variable	Theoretical Range	GZ			SZ			ALL		
		Mean	SD	Actual Range	Mean	SD	Actual Range	Mean	SD	Actual Range
RRPI_USEFUL	1.00-5.00	3.44	1.00	1.25-5.00	3.73	0.90	1.25-4.75	3.58	0.96	1.25-5.00
RRPI_FREQ	1.00-5.00	3.84	0.61	2.00-5.00	4.00	0.51	2.00-5.00	3.92	0.57	2.00-5.00
OSE_L	1.00-5.00	4.02	0.46	1.67-5.00	4.06	0.46	1.83-5.00	4.04	0.46	1.67-5.00
OSE	1.00-100.00	71.23	8.85	24.67-86.50	71.68	9.08	26.50-88.33	71.46	8.97	24.67-88.33
SCO	1.00-5.00	3.95	0.72	1.45-4.73	3.92	0.78	1.45-4.64	3.93	0.75	1.45-4.73
MON_ORDER		1046.19	247.28	410-2530	1053.90	138.45	450-1802	1050.04	200.33	410-2530
BAD_REVIEW		7.49	6.06	0-25	7.68	4.09	0-17	7.58	5.17	0-25
AVE_ORDER		1031.69	224.12	400-2400	1040.25	163.56	458-2856	1035.97	196.14	400-2856
QUALITY		0.007	0.005	0-0.027	0.007	0.004	0-0.023	0.007	0.005	0-0.027

Notes: Table 5 reports the descriptive statistics for the main variables in the model. All variables are explained in the Appendix.

Table 6 Correlations Table

	1	2	3	4	5	6	7	8
1. RRPI_USE	0.812							
2. RRPI_FREQ	0.228^{***}	--						
3. OSE	0.036	0.088^{**}	0.630					
4. SCO	0.382^{***}	-0.015	0.201^{***}	0.709				
5. MON_ORDER	-0.068^{**}	-0.035	0.160^{***}	-0.061	--			
6. BAD_REVIEW	-0.200^{***}	-0.109^{***}	0.136^{***}	-0.067^{**}	0.259^{***}	--		
7. AVE_ORDER	-0.081^{**}	-0.147^{***}	0.114^{***}	0.002	0.817^{***}	0.229^{***}	--	
8. QUALITY	-0.191^{***}	-0.126^{***}	0.085^{***}	-0.053^{*}	0.061^{*}	0.960^{***}	0.067^{**}	--

Notes: Table 6 shows the correlation between the main variables of our analyses. All variables are explained in the Appendix. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed). Off-diagonal underlined numbers are square roots of AVE. Square roots of AVE are omitted for single-item and observed variables.

Table 7: The effect of RRPI_USEFUL on different performance measures (N = 753 FT)

Path from:	Path to:	Hypothesis Sign	Path Model	Mixed Model
H1a				
RRPI_USEFUL	MON_ORDER	+	0.104*** (0.004)	0.092** (0.027)
RRPI_USEFUL	AVE_ORDER	+	0.027(0.462)	0.010 (0.779)
RRPI_USEFUL	BAD_REVIEW	-	-0.168*** (0.000)	-0.187*** (0.000)
RRPI_USEFUL	QUALITY	-	-0.200*** (0.000)	-0.217*** (0.000)
H2a				
RRPI_USEFUL x SCO	MON_ORDER	+	0.067 (0.409)	0.063 (0.431)
RRPI_USEFUL x SCO	AVE_ORDER	+	0.058 (0.204)	0.054 (0.243)
RRPI_USEFUL x SCO	BAD_REVIEW	-	-0.039 (0.367)	-0.050 (0.221)
RRPI_USEFUL x SCO	QUALITY	-	-0.047 (0.248)	-0.058† (0.160)
H3a: Hypothesised Indirect Paths				
RRPI_USEFUL > OSE > MON_ORDER		+	0.012* (0.097)	0.010† (0.184)
RRPI_USEFUL > OSE > AVE_ORDER		+	-0.002 (0.767)	-0.004 (0.544)
RRPI_USEFUL > OSE > BAD_REVIEW		-	0.003 (0.573)	-0.008† (0.200)
RRPI_USEFUL > OSE > QUALITY		-	-0.003 (0.547)	-0.015** (0.022)
Total effect RRPI_USEFUL on Performance				
MON_ORDER			0.116*** (0.001)	0.102** (0.015)
AVE_ORDER			0.024 (0.509)	0.006 (0.875)
BAD_REVIEW			-0.165*** (0.001)	-0.195*** (0.000)
QUALITY			-0.203*** (0.000)	-0.232*** (0.000)

Notes: Table 7 summarises the results of our main analysis of the perceived usefulness of RRPI using 753 FT riders' main sample. All variables are explained in the Appendix. We show the results of both the simple path model, which treats every latent variable as an observed variable, and the mixed structural model, which includes each item of RRPI_USEFUL and OSE construct in the model. All demographic characteristics are controlled in these models. The table reports standardized coefficients and P-values in parentheses. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed), † p < 0.10 one-tailed.

Table 8 The effect of RRPI_FREQ on different performance measures (N = 753 FT)

Path from:	Path to:	Hypothesis Sign	Path Model	Mixed Model
H1b				
RRPI_ FREQ	MON_ORDER	+	0.082** (0.018)	0.085** (0.015)
RRPI_ FREQ	AVE_ORDER	+	-0.066* (0.076)	-0.064* (0.093)
RRPI_ FREQ	BAD_REVIEW	-	-0.174*** (0.000)	-0.164*** (0.001)
RRPI_ FREQ	QUALITY	-	-0.215*** (0.001)	-0.206*** (0.001)
H2b				
RRPI_ FREQ x SCO	MON_ORDER	+	0.063* (0.060)	0.062* (0.063)
RRPI_ FREQ x SCO	AVE_ORDER	+	0.034 (0.458)	0.034 (0.467)
RRPI_ FREQ x SCO	BAD_REVIEW	-	0.035 (0.500)	0.034 (0.481)
RRPI_ FREQ x SCO	QUALITY	-	0.042 (0.412)	0.042 (0.400)
H3b: Hypothesised Indirect Paths				
RRPI_FREQ > OSE > MON_ORDER		+	0.009 (0.259)	0.006 (0.405)
RRPI_FREQ > OSE > AVE_ORDER		+	0.000 (0.990)	-0.002 (0.797)
RRPI_FREQ > OSE > BAD_REVIEW		-	0.008 (0.245)	-0.001 (0.784)
RRPI_FREQ > OSE > QUALITY		-	0.003 (0.776)	-0.007 (0.284)
Total effect RRPI_ FREQ on Performance				
MON_ORDER			0.091*** (0.005)	0.091*** (0.006)
AVE_ORDER			-0.066* (0.065)	-0.066* (0.065)
BAD_REVIEW			-0.166*** (0.001)	-0.166*** (0.001)
QUALITY			-0.213*** (0.000)	-0.212*** (0.001)

Notes: Table 8 summarises the results of our main analysis of the use frequency of RRPI using 753 FT riders' main sample. All variables are explained in the Appendix. We show the results of both the simple path model, which treats every latent variable as an observed variable, and the mixed structural model, which includes each item of the OSE construct in the model. All demographic characteristics are controlled in these models. The table reports standardized coefficients and P-values in parentheses. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed), † p < 0.10 one-tailed.

Table 9 The effect of RRPI_USEFUL on MON_ORDER (N = 753)

Path from:	Path to:	Hypothesis Sign	Path Model	Mixed Model
H1a & H2a				
RRPI_USEFUL	MON_ORDER	+	0.104*** (0.004)	0.092** (0.027)
RRPI_USEFUL x SCO	MON_ORDER	+	0.067 (0.409)	0.063 (0.431)
Control Paths				
GENDER	OSE		0.083*(0.052)	0.087* (0.052)
	MON_ORDER		-0.024(0.471)	-0.023(0.486)
AGE_CAT	OSE		0.146*** (0.000)	0.132*** (0.001)
	MON_ORDER		-0.082**(0.027)	-0.081**(0.032)
HOMETOWN	OSE		-0.047† (0.194)	-0.021(0.619)
	MON_ORDER		0.022(0.539)	0.020(0.591)
TENURE	OSE		-0.244*** (0.000)	-0.248*** (0.000)
	MON_ORDER		0.045(0.481)	0.043(0.519)
CITY	OSE		-0.013(0.704)	-0.078*** (0.066)
	MON_ORDER		0.069*(0.052)	0.074**(0.037)
LEVEL	OSE		0.472*** (0.000)	0.547*** (0.000)
	MON_ORDER		0.538*** (0.001)	0.537*** (0.001)
H3a: Hypothesised Indirect Paths				
RRPI_USEFUL > OSE > MON_ORDER		+	0.012* (0.097)	0.010† (0.184)
Total effect RRPI_USEFUL on Performance				
MON_ORDER			0.116*** (0.001)	0.102** (0.015)
Model Fit statistics:				
R-squared OSE			0.183	0.218
R-squared MON_ORDER			0.307	0.303
χ2 (p)			57.129(0.000)	402.572 (0.000)
χ2 /df			28.565	3.908
CFI			0.967	0.943
RMSEA			0.191	0.062
SRMR			0.0326	0.0558

Notes: Table 9 shows one model as an example of our main analysis of the perceived usefulness of RRPI on the number of delivery orders using 753 FT riders' main sample. All variables are explained in the Appendix. The table reports standardized coefficients and P-values in parentheses. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed), † p < 0.10 one-tailed.

Table 10 The effect of RRPI_FREQ on MON_ORDER (N = 753)

Path from:	Path to:	Hypothesis Sign	Path Model	Mixed Model
H1a & H2a				
RRPI_FREQ	MON_ORDER	+	0.082** (0.018)	0.085** (0.015)
RRPI_FREQ x SCO	MON_ORDER	+	0.063* (0.060)	0.062* (0.063)
Control Paths				
GENDER	OSE		0.091** (0.040)	0.092* (0.052)
	MON_ORDER		-0.023 (0.487)	-0.023 (0.498)
AGE_CAT	OSE		0.153*** (0.000)	0.147*** (0.000)
	MON_ORDER		-0.083** (0.020)	-0.081** (0.023)
HOMETOWN	OSE		0.009 (0.788)	0.021 (0.595)
	MON_ORDER		0.048 (0.201)	0.047 (0.214)
TENURE	OSE		-0.277*** (0.000)	-0.292*** (0.000)
	MON_ORDER		0.026 (0.672)	0.025 (0.697)
CITY	OSE		-0.006 (0.834)	-0.060† (0.116)
	MON_ORDER		0.069** (0.033)	0.071** (0.030)
LEVEL	OSE		0.488*** (0.000)	0.562*** (0.000)
	MON_ORDER		0.559*** (0.001)	0.562*** (0.000)
H3a: Hypothesised Indirect Paths				
RRPI_FREQ > OSE > MON_ORDER		+	0.009 (0.259)	0.006 (0.405)
Total effect RRPI_FREQ on Performance				
MON_ORDER			0.091*** (0.005)	0.091*** (0.006)
Model Fit statistics:				
R-squared OSE			0.184	0.222
R-squared MON_ORDER			0.308	0.307
$\chi 2$ (p)			77.110(0.000)	282.887(0.000)
$\chi 2$ /df			38.555	5.052
CFI			0.945	0.926
RMSEA			0.223	0.073
SRMR			0.0328	0.0559

Notes: Table 10 shows one model as an example of our main analysis of the use frequency of RRPI on the number of delivery orders using 753 FT rider's main sample. All variables are explained in the Appendix. The table reports standardized coefficients and P-values in parentheses. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed), † p < 0.10 one-tailed.

Table 11 The effect of RRPI_USEFUL on different performance measures (N = 247 PT)

Path from:	Path to:	Hypothesis Sign	Path Model	Mixed Model
H1a				
RRPI_USEFUL	MON_ORDER	+	0.049 (0.361)	0.037 (0.508)
RRPI_USEFUL	AVE_ORDER	+	0.060 (0.284)	0.041 (0.478)
RRPI_USEFUL	BAD_REVIEW	-	-0.051 (0.490)	-0.209 (0.533)
RRPI_USEFUL	QUALITY	-	-0.058 (0.402)	-0.061 (0.429)
H2a				
RRPI_USEFUL x SCO	MON_ORDER	+	0.146** (0.016)	0.143** (0.018)
RRPI_USEFUL x SCO	AVE_ORDER	+	0.165*** (0.004)	0.157*** (0.005)
RRPI_USEFUL x SCO	BAD_REVIEW	-	-0.141** (0.042)	-0.558* (0.076)
RRPI_USEFUL x SCO	QUALITY	-	-0.146** (0.021)	-0.138** (0.029)
H3a: Hypothesised Indirect Paths				
RRPI_USEFUL > OSE > MON_ORDER		+	0.002 (0.430)	0.005 (0.368)
RRPI_USEFUL > OSE > AVE_ORDER		+	0.005 (0.211)	0.016 (0.200)
RRPI_USEFUL > OSE > BAD_REVIEW		-	-0.010 (0.462)	-0.017 (0.215)
RRPI_USEFUL > OSE > QUALITY		-	-0.008(0.413)	-0.012 (0.213)
Total effect RRPI_USEFUL on Performance				
MON_ORDER			0.051 (0.353)	0.042 (0.470)
AVE_ORDER			0.064 (0.258)	0.057 (0.352)
BAD_REVIEW			-0.061 (0.412)	-0.069 (0.389)
QUALITY			-0.066 (0.345)	-0.073 (0.326)

Notes: Table 11 summarises the results of our additional analysis of the perceived usefulness of RRPI using 247 PT riders' samples. All variables are explained in the Appendix. We show the results of both the simple path model, which treats every latent variable as an observed variable, and the mixed structural model, which includes each item of RRPI_USEFUL and OSE constructs in the model. All demographic characteristics are controlled in these models. The table reports standardized coefficients and P-values in parentheses. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed), † p < 0.10 one-tailed.

Table 12 The effect of RRPI_FREQ on different performance measures (N = 247 PT)

Path from:	Path to:	Hypothesis Sign	Path Model	Mixed Model
H1b				
RRPI_FREQ	MON_ORDER	+	0.026 (0.731)	0.032 (0.694)
RRPI_ FREQ	AVE_ORDER	+	0.058 (0.435)	0.061 (0.431)
RRPI_ FREQ	BAD_REVIEW	-	-0.010 (0.868)	-0.027 (0.715)
RRPI_ FREQ	QUALITY	-	-0.021 (0.728)	-0.035 (0.600)
H2b				
RRPI_ FREQ x SCO	MON_ORDER	+	-0.076 (0.284)	-0.079 (0.266)
RRPI_ FREQ x SCO	AVE_ORDER	+	-0.135* (0.095)	-0.138* (0.088)
RRPI_ FREQ x SCO	BAD_REVIEW	-	-0.246*** (0.004)	-0.234*** (0.003)
RRPI_ FREQ x SCO	QUALITY	-	-0.254*** (0.004)	-0.245*** (0.003)
H3b: Hypothesised Indirect Paths				
RRPI_ FREQ > OSE > MON_ORDER		+	-0.004 (0.391)	-0.011 (0.355)
RRPI_ FREQ > OSE > AVE_ORDER		+	-0.009 (0.240)	-0.002 (0.308)
RRPI_ FREQ > OSE > BAD_REVIEW		-	0.019 (0.278)	0.032 (0.333)
RRPI_ FREQ > OSE > QUALITY		-	0.016 (0.249)	0.027 (0.319)
Total effect RRPI_ FREQ on Performance				
MON_ORDER			0.022 (0.764)	0.023 (0.752)
AVE_ORDER			0.049 (0.501)	0.050 (0.491)
BAD_REVIEW			0.009 (0.926)	0.005 (0.953)
QUALITY			-0.005 (0.910)	-0.008 (0.893)

Notes: Table 12 summarises the results of our additional analysis of the use frequency of RRPI using 247 PT riders' sample. All variables are explained in the Appendix. We show the results of both the simple path model, which treats every latent variable as an observed variable, and the mixed structural model, which includes each item of the OSE construct in the model. All demographic characteristics are controlled in these models. The table reports standardized coefficients and P-values in parentheses. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed), † p < 0.10 one-tailed.

Table 13 The effect of the gamified rider level on RRPI_USEFUL

	RRPI_USEFUL	
RIDER_LEVEL	-0.326***	-0.340***
	(-7.971)	(-8.032)
RIDER_LEVEL^2		-0.044
		(-1.279)
SCO	0.335***	0.329***
	(10.123)	(9.815)
GENDER	0.005	0.008
	(0.168)	(0.262)
AGE_CAT	-0.008	-0.011
	(-0.230)	(-0.344)
TENURE	0.062	0.078*
	(1.483)	(1.784)
HOMETOWN	0.026	0.029
	(0.764)	(0.864)
CITY	0.197***	0.190***
	(6.196)	(5.902)
N	753	753
R-square	0.271	0.273

Notes: Table 13 shows the results of the additional analysis on the relationship between the perceived usefulness of RRPI and the gamified rider level. All variables are explained in the Appendix. The table reports standardized coefficients, and t-value in parentheses. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed), † p < 0.10 one-tailed.

Table 14 The effect of the gamified rider level on RRPI_FREQ

	RRPI_FREQ	
RIDER_LEVEL	-0.430***	-0.458***
	(-9.799)	(-10.102)
RIDER_LEVEL^2		-0.086**
		(-2.353)
SCO	-0.001	-0.014
	(-0.036)	(-0.394)
GENDER	-0.007	-0.001
	(-0.215)	(-0.041)
AGE_CAT	-0.037	-0.045
	(-1.051)	(-1.261)
TENURE	0.163***	0.194***
	(3.622)	(4.148)
HOMETOWN	-0.217***	-0.210***
	(-5.947)	(-5.760)
CITY	0.161***	0.148***
	(4.711)	(4.272)
N	753	753
R-square	0.160	0.166

Notes: Table 14 shows the results of the additional analysis on the relationship between the use frequency of RRPI and the gamified rider level. All variables are explained in the Appendix. The table reports standardized coefficients, and t-value in parentheses. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed), † p < 0.10 one-tailed.

Table 15 The effect of RRPI_USEFUL in different rider-level groups (MON_ORDER)

	Golden	Pioneer	Elite	Glory
RRPI_USEFUL	-0.120†	0.141*	0.290***	0.086
	(-1.496)	(1.741)	(3.715)	(1.189)
GENDER	0.256***	-0.072	-0.082	-0.101†
	(3.733)	(-0.984)	(-1.144)	(-1.512)
AGE_CAT	0.077	-0.115†	-0.045	-0.220***
	(1.043)	(-1.590)	(-0.637)	(-3.226)
TENURE	-0.118	-0.149†	0.581***	0.357***
	(-1.638)	(-1.623)	(7.619)	(5.199)
HOMETOWN	-0.017	0.114†	0.093	0.163**
	(-0.259)	(1.412)	(1.203)	(2.382)
SCO	0.110	-0.086	-0.346***	0.074
	(1.275)	(-1.039)	(-4.104)	(0.872)
SCO x RRPI_USEFUL	0.159*	-0.102†	-0.139†	0.135†
	(1.938)	(-1.352)	(-1.561)	(1.504)
CITY	0.483***	0.072	-0.108†	-0.204***
	(6.932)	(0.930)	(-1.439)	(-3.049)
N	178	200	123	190
R-square	0.313	0.101	0.446	0.237

Notes: Table 15 shows the results of the effect of the perceived usefulness of RRPI on the number of delivery orders within each FT rider gamified level group. Bronze and Silver rider groups are not included because of the small number of samples. All variables are explained in the Appendix. The table reports standardized coefficients, and t-value in parentheses. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed), † p < 0.10 one-tailed.

Table 16 The effect of RRPI_FREQ in different rider-level groups (MON_ORDER)

	Golden	Pioneer	Elite	Glory
RRPI_FREQ	0.152**	0.138*	0.216**	-0.544***
	(2.122)	(1.687)	(2.412)	(-3.767)
GENDER	0.186***	-0.071	-0.049	-0.123*
	(2.699)	(-0.972)	(-0.639)	(-1.906)
AGE_CAT	0.046	-0.124*	-0.031	-0.218***
	(0.657)	(-1.726)	(-0.421)	(-3.365)
TENURE	-0.150**	-0.164*	0.504***	0.361***
	(-2.075)	(-1.847)	(6.517)	(5.116)
HOMETOWN	0.021	0.141*	0.149*	0.110
	(0.293)	(1.690)	(1.879)	(1.616)
SCO	-0.058	-0.031	-0.175**	-0.079
	(-0.869)	(-0.376)	(-2.220)	(-1.140)
SCO x RRPI_FREQ	0.035	0.043	0.063	0.525***
	(0.508)	(0.497)	(0.805)	(3.856)
CITY	0.442***	0.067	-0.164**	-0.188***
	(6.705)	(0.899)	(-2.010)	(-2.862)
N	178	200	123	190
R-square	0.311	0.095	0.402	0.287

Notes: Table 16 shows the results of the effect of the use frequency of RRPI on the number of delivery orders within each FT rider gamified level group. Bronze and Silver rider groups are not included because of the small number of samples. All variables are explained in the Appendix. The table reports standardized coefficients, and t-value in parentheses. ***, **, *. Indicates the significance of the p-value at <0.01, 0.05, 0.10 (2-tailed), † p < 0.10 one-tailed.

Paper 3

Motivations from Gamification: Internalisation of external incentives in an online learning platform

Motivations from Gamification: Internalisation of external incentives in an online learning platform

Abstract

A good gamification design needs to integrate both intrinsic and extrinsic motivations of people to influence their behaviour. I conducted interactive netnography research on an online learning platform by immersing myself in this setting as a superuser and interviewing other users. Mobilising self-determination theory, I analyse how this platform used gamification to change the users' behaviour through the internalisation of external gamified elements. Specifically, the platform lets users get used to some gamified elements such as the check-in function, gamified badges, and learning credit with learning performance information, and form their user behaviours as a part of their own value and life. Therefore, these internalisations further complement the intrinsic motivation to satisfy the user's three basic psychological needs autonomy, competence, and relatedness. The findings shed light on how to design a gamified incentive system in a workplace to motivate workers not only in an organisation preferable way but also in an integrating way of themselves.

4.1 Introduction

According to the incentive literature of management accounting, instead of using traditional monetary incentives to motivate employees' behaviours (Bonner and Sprinkle, 2002; Van der Stede, Wu and Wu, 2020), many organisations start to use novel ways to incentivise their employees such as giving leisure time (Vogelsang, 2022), tangible rewards (Kelly, Presslee and Webb, 2017; Presslee, Vance and Webb, 2013; Choi and Presslee, 2023), prosocial rewards (Berger, Guo and Presslee, 2023), or gamification (Kelly, Valtchanov and Webb, 2021). Moreover, some platforms such as TripAdvisor (Jeacle and Carter, 2011), Foursquare (Chapman, Chua and Fiedler, 2021), and Airbnb (Leoni and Parker, 2019; Van den Bussche and Dambrin, 2020) also used some gamification elements to control and seduce their users' behaviour, making their "labour" a core part of the business model and commercial success of these platforms even as they are not employees. This paper further draws attention to gamification to elaborate on how it can motivate people's behaviours through satisfying basic psychological needs and internalising the external controls as parts of oneself.

Gamification, which is defined as the use of game elements in a non-game

context (Deterding *et al.*, 2011), has been popularly used in organisations to transform their business. According to Callan, Bauer and Landers (2015), there will be 40% of Global 1000 organisations to use gamification by 2015. Through different gamified mechanics such as badges, points and leaderboards, the organisations changed the way to motivate and control people's behaviour, which is the key issue for accounting researchers. Using proprietary data, Kelly, Valtchanov and Webb (2021) find that after using an online slot machine, there is a positive association between the extent to which employees play the online slot machine and the increase of training modules finished by the employees and the effort that employees put into better performance. However, they find that the employees, who are inclined to gamified designs, gradually lose their interest in playing the online slot machine and training activities over time. Moreover, the gamification even reduces the training activities significantly finished by employees who chose not to play. The effect of gamification is short-term and only influences the sub-set of employees. Regarding the platform companies, Chapman, Chua and Fiedler (2021) find that Foursquare uses gamified elements such as virtual rewards and badges to seduce users to keep checking in different places to exchange feelings of pleasure and a sense of community. Jeacle and Carter (2011) also explain how these symbolic designs such as ratings and leaderboards build trust from travellers around the world on TripAdvisor.

To further zoom in on how gamification works to motivate behaviour, I use the self-determination theory (SDT) to explain the logic behind it. SDT is mainly related to people's basic psychological needs and different kinds of motivation (Deci and Ryan, 2000). There are three basic psychological needs for people to satisfy: competence (people feel they are capable of achieving performance), autonomy (people feel their own violation and freedom to decide what they want to do), and relatedness (people feel that they are connected to and accepted by others) (Deci and Ryan, 2000). Normally, intrinsically motivated behaviours, which are defined as the activity people purely enjoy for the sake of this activity itself (Deci, 1975), can satisfy these needs. On the contrary, extrinsically motivated behaviours triggered by external rewards, regulations or controls might not fully satisfy people's needs, depending on the internalisation process of external factors (Ryan, Connell and Deci, 1985). According to different extents of internalisation, extrinsic motivation is categorised into four kinds: external regulated, introjected, identified, and integrated motivations from fully external forced

to fully internalised to the part of people's own value. Since gamification tries to provide people with both intrinsic motivation and extrinsic motivation (Zichermann and Cunningham, 2011), it is a good opportunity to find out how the combination of different motivations influences people's behaviour and performance (Van der Hauwaert *et al.*, 2022; Bouten and Hoozée, 2022; Pfister and Lukka, 2019).

Inspired by Jeacle and Carter (2011, p. 309), "positioning oneself in front of the social and organisational context", I used the interactive netnography method (Kozinets, 2019) to undertake a field study in a Chinese online learning platform. I immersed myself as a superuser of this platform for more than five years before jumping in as a researcher. When planning to conduct research in this setting, I recorded my feelings, user behaviours, and the change in the interface of the platform application. I conducted 39 interviews with the users of the platform to understand how they use this application. Meanwhile, I collected the learning data of most of these participants to further confirm my findings.

Through this field study, I find that different from a single incentive mechanism, in a gamified incentive system, gamified elements such as a daily check-in button, gamified challenges and badges, and learning credits integrate together to influence the users' behaviour significantly. Some of these gamified elements fully internalise to the part of users' value and life, which further enhanced the original intrinsic motivation of learning at the beginning and increased users' stickiness. Moreover, these gamified elements also serve as symbolic social currency to build a strong connection and trust among users on this platform, which also internalises external motivations and significantly satisfies the need for relatedness. Finally, a vague learning credit system without clear rules and goals also motivates users to try to boost their learning credits by improving five different performance measures together.

The study has certain implications from both academic and practical practices. Firstly, I respond to Kelly, Valtchanov and Webb's (2021) call for using interviews to develop a deeper understanding of gamification issues. I conduct 39 interviews to discover users' attitudes and habits toward those gamified elements in a gamified incentive system, which have influenced their behaviours and motivations. Meanwhile, only a few papers explore the effect of gamification (Chapman, Chua and Fiedler, 2021; Kelly, Valtchanov and Webb, 2021) in accounting research. I continue to discover the

motivational effect from gamification in a learning platform setting. Gamification as one of the novel incentives (Berger, Guo and Presslee, 2023; Vogelsang, 2022; Kelly, Presslee and Webb, 2017; Presslee, Vance and Webb, 2013) has many potential and possibilities for use by organisations.

Secondly, for the designer of gamified systems, my study suggested that vague gamified performance rules can prevent certain gaming behaviours of users by chasing a specific performance measure in the system. Users in this learning platform need to improve every performance measure to increase their learning credits instead of only improving one performance measure. Furthermore, designers of gamified systems need to consider the existence of different types of "players" in a system and try to use different kinds of game mechanics to incentivise their behaviours, which also extends the understanding of the individual differences of gamification (Cardador, Northcraft and Whicker, 2017; Kelly, Valtchanov and Webb, 2021).

Thirdly, mobilising SDT in accounting research (Van der Hauwaert *et al.*, 2022; Bouten and Hoozée, 2022; Chen, Lill and Vance, 2020; Pfister and Lukka, 2019), my study further explains the internalisation of extrinsic motivations in an original intrinsically motivated setting. Different from other settings, my study shows the internalisation of external gamified elements based on the foundation of intrinsic motivation and the crowding-in effect of the combination of two kinds of motivations. Moreover, I discuss the different kinds of internalisation processes and further show the cognitive differences of users in my setting.

I will explain the SDT in detail and introduce my research question in the next section. The interaction netnography method will be explained in section 3. Then I will show my findings on this online learning platform in section 4. Sections 5 and 6 will be discussion and conclusion.

4.2 Theory and Research Questions
4.2.1 Self-determination theory:
The self-determination theory (SDT) is one of the most influential psychology theories to conceptualise human beings' basic psychological needs and motivation. According to SDT, alongside basic peoples' physical needs such as nutrients for body health and safety, three basic psychological needs are important for people's psychological interests and development: autonomy, competence, and relatedness (Deci

and Ryan, 2000). Firstly, autonomy is defined as the need to self-regulate one's actions and experiences. In other words, people experience behaviours of their own volition and full self-endorsement rather than being controlled, coerced, or seduced by some external forces of the self. It is important to satisfy the autonomy needs to let people choose what they are willing to do or not to do. These actions fully driven by the sense of volition and endorsement do have larger access to people's physical, cognitive, and emotional capabilities than the behaviours mainly motivated by external regulations or controls. More importantly, with the full engagement of people's abilities and energies, autonomous actions generally generate better outputs and performance (Ryan and Deci, 2017).

Secondly, competence refers to the need that people have feelings of capability and mastery of what they are doing (Deci, 1975). This is an inherent need for people to feel that they can achieve performance on their certain actions. When people achieve a good performance or receive positive feedback in their jobs or in any games they play, they feel more competent and confident to continue on this task and hope to perform even better later. The competence of certain actions motivates people to have better performance. However, the need for competence can be also easily unsatisfied when people face tasks with an unachievable and difficult target, when people receive negative feedback from their supervisors, or when they underperform their peers after social comparison (Festinger, 1954). These situations will probably demotivate individuals to perform later. Moreover, Ryan and Moller (2017) stressed that to get more positive effects from competency activities, the activities should be autonomous and self-regulated instead of controlled by external forces.

Thirdly, relatedness is defined as the need that people feel connected to others. Beyond the tangible helps with each other, people want to have feelings of belongings to certain groups and communities and to feel a sense of mattering to others. It is a basic need for people to get respect and responses from others and avoid rejection and disconnectedness from others (Ryan and Deci, 2017). As people want to be better connected with others, they tend to have an interest in what others think and expect from them. Based on these interests, people will internalise those values or behaviours to become part of their own values to get acceptance from others. However, whether these outside beliefs and behaviours are fully integrated as a part of the self or are still

externally controlled forces is significant to justify whether people satisfy the need for relatedness. To fully satisfy the basic need of relatedness, people need to feel the enhancement of their self-esteem and feel that their related actions are accepted and affirmed by other members.

The true satisfaction of these three basic psychological needs has one common prerequisite is the internal locus of causality, which means that people attribute certain behaviours to their own endorsement and volition (DeCharms, 1968). Only if people have the autonomy to decide what they want to do, achieve some performance in those autonomous activities, and internalise fully other people's interests as part of themselves, they can satisfy the needs of autonomy, competence and relatedness. This locus of causality also differentiates different kinds of motivation on the spectrum from intrinsic motivation to extrinsic motivation (Weibel, Wiemann and Osterloh, 2014), which brings out the main motivation part of SDT.

Intrinsic motivation is defined as the motivation purely from the feelings of efficacy and enjoyment brought by the activity itself instead of other interests or outside rewards (Deci, 1975). On the contrary, extrinsic motivation is defined as the motivation from separate consequences or interests such as external rewards, social recognition, punishment avoidance or the success of valued outcomes (Deci and Ryan, 2000). More specially, there are different kinds of external motivations based on the extent to congruent with one's self, which means vary based on the different levels of external locus of causality or one's autonomy (Ryan and Connell, 1989; Landers *et al.*, 2015).

External regulation is the strongest controlled external motivation where people are purely motivated by the rules from authority, fear of punishment or any kind of external contingencies (Ryan and Deci, 2017). For example, employees need to finish certain tasks they don't like or want to do purely because of the salary or the codes of conduct of the company. Introjection is an intermediate kind of external motivation which also connects people's shame, self-esteem, pride, and self-satisfaction (Ryan and Deci, 2017). People will feel guilty and anxious if they don't perform well under this kind of motivation. Identified motivation, as the third type of extrinsic motivation, means that people can truly feel the personal relevance and importance when they finish the task, and they can recognise the value of doing this task. At this level, people can make more decisions on what they choose to do. Finally, integrated motivation, which

is the most similar to intrinsic motivation, further internalised the identified motivations so that people can almost feel the full sense of autonomy and interest in finishing the tasks themselves.

Based on the extent of the locus of causality, these different kinds of motivation are also divided into autonomous motivation (Intrinsic motivation, Identified motivation and Integrated motivation) and controlled motivation (Externally regulated motivation and Introjected motivation). According to Ryan and Deci (2017), these concepts of autonomous or controlled motivation are quite theoretical and hard to directly reflect the psychological process of different people. However, based on certain manipulation strategies of people's autonomy, psychology studies do find different impacts on people's behaviours and performance under different kinds of motivations. To conclude, as the main parts of SDT, not only do people's motivations have different magnitudes, but also they are influenced by different kinds of external motivation and internal motivation.

4.2.2 Gamification and SDT

Gamification is defined as the use of game elements in non-game contexts (Deterding *et al.*, 2011). One significant part of this definition to highlight is game design elements, which means rather than developing a whole serious game gamification uses the different mechanics of games in different contexts (Sailer *et al.*, 2017). By implementing different mechanics such as points, levels, badges and challenges, gamification uses game-thinking to engage users and solve their actual problems in real life (Zichermann and Cunningham, 2011). These mechanics are implemented by some platform companies such as Foursquare (Chapman, Chua and Fiedler, 2021), Airbnb (Leoni and Parker, 2019; Van den Bussche and Dambrin, 2020) and eBay (Kornberger, Pflueger and Mouritsen, 2017) to seduce and control their users' behaviour. Not only does gamification trigger the intrinsic motivation for users to enjoy the gamified designs themselves, but also provides extrinsic motivation for users to change their behaviours (Landers *et al.*, 2015).

On the one hand, gamification serves three basic psychological needs to intrinsically motivate users in certain contexts. Normally, in a gamified setting, users can autonomously choose what elements they want to play and what elements they choose to ignore (Deterding *et al.*, 2011). These elements give different users

opportunities to explore how they want to engage in this setting. This serves the autonomy need that users have a quite strong perceived locus of causality of certain behaviours. Bartle (1996) suggested a famous model for different kinds of gamers to show the autonomy of different users.

There are four kinds of players: Achievers, Explorers, Socialisers, and Killers. Achievers want to achieve certain goals in the game such as collecting more points and levering up. Explores are mainly to explore the setting to find out more interesting and new things. The fun of the game is from the exploration by themselves. Another part of players are socialisers whose main purpose of playing the game is to interact and socialise with other users. They care about this platform so they can build up more inter-player relationships. Finally, killers just want to beat other players in the game. The more distress they cause to others, the happier they are in the game. However, Zichermann and Cunningham (2011) argue that users might not only be one kind of player exclusively. Therefore, the platforms using gamified designs need to consider the different types of players or users to attract as many as users they can (Bartle, 1996).

With the different purposes of different users, users can satisfy their competence needs by finishing certain challenges and achieving some performance on the tasks (Ryan, Rigby and Przybylski, 2006). From what users autonomously choose to do on the platform, they get feelings of self-efficacy and a sense of achievement. Moreover, no matter whether the user is a socialiser or not, the users on the same platform will more or less serve their relatedness need to feel accepted by other users and generate a sense of community among users (McMillan and Chavis, 1986; Chapman, Chua and Fiedler, 2021). All three basic needs were satisfied for users to have intrinsic motivation.

On the other hand, gamification uses different kinds of external rewards or rules to provide users with different kinds of extrinsic motivations. Chapman, Chua and Fiedler (2021) show how the gamified external rewards such as badges, virtual stickers and virtual titles "mayorship" keep seducing users to use Foursquare to check in the places where they went. Van den Bussche and Dambrin (2020) describe how the external evaluation rules from Airbnb change the behaviours of the hosts and guests on the platform.

Given that gamification designs provide both intrinsic and extrinsic motivations, there will be potential problems with the crowding-out effect (Frey, 1997). Deci and

Ryan (1985) find that pay for performance negatively influences performance. Frey (1997) defined the crowding-out effect that intrinsic motivation can be reduced by external rewards, punishments, or other forms of control. According to Weibel, Wiemann and Osterloh (2014), there are some different explanations for this effect. The first explanation is about the change of locus of causality (DeCharms, 1968). Some kinds of external rewards change the original internal locus of causality for people to do certain tasks to the external locus of causality to conduct the tasks because of certain external controls. Another explanation is the signalling effect that external controls signal agents that their principals have no trust in them or signal agents that the tasks are not enjoyable or difficult which needs to give some external rewards to find people to finish. No matter which explanation, Frey and Jegen (2001) argue that performance-related compensation creates the crowding-out effect to reduce the effort at work when an original intrinsically motivated activity is motivated by the external reward interpreted as a control tool, which cannot counterbalance the loss of intrinsic motivation.

Therefore, to reduce the conflict and crowd-out effect, Zichermann and Cunningham (2011) argue that at the beginning of designing the gamified system, the designer needs to let users be who they are and experience their motivational status. Guiding them to where they want to go, the designer also needs to help them reach their greater potential and explore different things they didn't know before. This coincides with the crowding-in effect (Frey and Osterloh, 2001), which is defined that some specific external institutional factors such as participation, procedural fairness and clear normative signals might boost the intrinsic motivation for certain tasks. How to align intrinsic and extrinsic motivation to get the crowding-in effect from gamification is related to internalisation.

Ryan, Connell and Deci (1985) defined internalisation as the process of transforming external values, beliefs or regulations into people's own. Through internalisation, extrinsically motivated behaviours can become part of people's internal motivation, which helps them satisfy three basic psychological needs. The different kinds of external motivations do have different levels of internalisation from externally regulated motivation to integrated motivation. When the external factors are fully internalised into the integrated motivation, these external controls are congruent with

people's own identity, and one part of self (Ryan and Deci, 2017). Zichermann and Cunningham (2011, p. 29) mentioned, "A good extrinsic motivation is a good map to intrinsic motivation". Given the online learning platform uses gamification to boost users' intrinsic motivation to learn new knowledge, it is a good setting for me to look into the details of this dynamic of internalisation between extrinsic gamified elements and intrinsic motivation of learning and reading. Therefore, my research question is: ***Among different kinds of users, how does gamification internalise the external gamified elements and rewards to boost the intrinsic motivation of users and influence their behaviours?***

4.3 Research Method

4.3.1 Interacting and Immersing in the Setting

According to Kozinets (2019), netnography is a fruitful approach to conducting qualitative research using social media as the basis of data sources. It is an innovative approach for accounting researchers to collect data (Jeacle and Carter, 2011; Bialecki, O'Leary and Smith, 2017; Van den Bussche and Dambrin, 2020; Chapman, Chua and Fiedler, 2021). Although the setting of the study is an online learning platform, it does have social media features such as an online community (Knowledge City State, described later in detail) to post users' comments or evaluations of what they learnt and even their daily life such as Instagram. Every user can choose to verify their account to display their identities in real life and follow other users' accounts to see their posts and what they are learning on this platform. Kozinets (2019) mentioned that the biggest common issue for non-participation in netnography or ethnography is the risk of only describing the setting superficially instead of capturing the cultural dynamics. My natural engagement previously helps to solve the potential superficiality issues of the depth of netnographic research.

Before starting this project, I have been using this application and starting on this platform for more than five years. As the starting user of this application, I witnessed every step of change in the design of this application and the business development of this platform. When the first time platform introduced the gamified elements three years ago, I was quite intrigued by those designs, and my user behaviours also started to change from that moment. After immersing in this platform for many years, as one of the research serendipities (Chapman *et al.*, 2023; Bebbington *et al.*, 2020), connecting to my research interests in gamification and incentive systems, I felt

that it is a good opportunity to dig into this setting from a researcher perspective to conduct a netnography field study.

Within the five engagement strategies introduced by Kozients' book, I automatically fit into three strategies. From the historical engagement perspective, I started to engage in this platform when it was founded and understood quite well the history of the platform, the superusers' behaviour and the future potential development of the platform. From an emotional perspective, I witnessed many significant moments or updates of this platform with many other superusers. Same as many users, I was moved by some movements of this platform. All these moments serve as evidence to show my emotional engagement on this platform. From a social engagement perspective, during my six-year journey, I post more than 5,000 comments or reviews for courses or books on this platform and get many likes and comments from users on this platform. After starting with this research project, I have deeper social engagements with users on this platform through online interviews[24].

Extending my role from a pure superuser to a researcher, one important role in the netnography is research disclosure (Kozinets, 2019). It is important to disclose your researcher's identity and profile before starting the research. Given my account is verified by the platform, every user on the platform can see my identity as a management accounting PhD student and my profile[25]. Furthermore, when getting access to each user for an interview, I sent them the consent form to introduce this project and the purpose of the interview to make sure they were satisfied to participate happily and voluntarily. The participants to get access are also verified by their study data on the platform. I will check whether they have some minutes' of studying on the day I am getting access. More than half of the participants are active users recently and some of them are even superusers, who are from different provinces around China.

For the choice of the participants, the first batch of the participants is from commenting on their posts. Given there is no private message function within the platform, I first commented on some randomly selected users on recent posts in the community function and at the same time introduced my project to let them add my

---

[24] The research project and interview were approved by the ethical committee of the business school.
[25] The platform took a serious process to verify an account. The users have to get over 30 followers, follow over 30 users, have 5 posts in recent months, and send them proof of work or proof of study to verify.

WeChat to get further information. The second batch is from an organiser list of the New Year's Eve speech offline event hosted by the platform. In this list, the platform displays the WeChat accounts of these users in different cities. I directly added them to the WeChat with an invitation note about my research project. Most of them agreed to my invitation and wanted to further discuss the project. The final batch of the participants is from the snowballing strategy from the first two batches of participants. There are three face-to-face interviews and 36 online interviews through WeChat, which is the most popular and convenient teleconferencing application for participants to use in China. In total, I conducted 39 interviews with different users from 18 provinces in China and across different industries and age ranges, and the gender is balanced coincidently with 19 female and 20 male participants. Their total learning time on this platform ranges from 38 to 7,661 hours since some are loyal users since the platform was established and some are new users just joining the platform[26]. With the different strategies to get access, I eliminated large parts of selection bias on the participants. All the demographic and interview information are shown in Table 1.

[Table 1 about here]

The interviews were conducted in a semi-structured way. As an insider of this setting, this superuser identity made me discuss with these participants using our familiar languages or terminologies in this platform, which helps me understand and decipher their behaviour better (Kozinets, 2019). This also helps me quickly build a close rapport at the start of the interviews (Baxter and Chua, 1998). To let users openly share their thoughts, we started with some open questions such as "Could you describe your daily use of this application?" or "Which courses do you like most and Why?". Following the guidance from Patton (2002) and Marginson (2004), we added some prompt questions to help participants continue their sharing. To not create some demanding effect in interviews given my specific role as a superuser, most often when the participants started to ask my opinions or preferences, I explained to them and postponed their questions or my answers to the end of the interviews, unless the conversation was flowing naturally, or the structured questions have been answering by participants thoroughly. As suggested by Kozinets (2019), I stopped recording until the window closed on my screen for online interviews. In some interviews, this behaviour

---

[26] 3 participants didn't share their learning data publicly.

does generate some interesting points and discussions which are relevant to this project during the final chat.

Alongside the interview data, through every user's account profile, with the permission of the participants, I also collected the study data from most of the participants, which is a good supplement to the interview data to verify and triangulate what they said in the interviews. With the good performance data in my user account profile, the users from this platform are quite willing to help me with my research project. Most of them mentioned that it is their honour to participate in this research project to help their cohort (me) study on the same platform and sent the invitation to meet in person if possible.

4.3.2 Data analysis

According to Kozinets (2019), as a netnographer, when we are immersed in the setting and record reflectively what we are experiencing, we will get both closeness and distance from the people and the platform. We need to let our real experiences and emotions flow in our immersion notes as our internal human reality is the key element of netnography. Since I started this project, I have taken some immersion notes of my daily use of the application and recorded some key changes and findings of this platform. At the same time, I made some immersion reflective notes after a few interviews to record what I felt in those chats. As mentioned by Kozinets (2019, p. 301), "Ultimately, those reflective notes are what will allow the story of your vulnerable humanity to suffuse your work". These immersion notes with some theoretical lenses also fit my abductive approach to data analysis.

According to Ahrens and Chapman (2006), I used the abductive approach to go back and forth between data and theories. With my personal user experience, I have some empirical and theoretical expectations for creating interview protocol and conducting the interviews. Based on the interview data, I find some fits with previous expectations but meanwhile, we also find the anomalies in the data to give me the chance to go back to the theories to find the appropriate explanation and further develop the understanding of theories. I followed the three steps to analyse my data: collating, coding and combining (Kozinets, 2019). For collating, it is mainly the preparation for coding that I transcribed all the interviews manually and combined them with the study data I collected from the platform and some screenshots of the interface of the

application from my personal profile. During the transcribing process, I highlighted some key sentences in the interviews and wrote some reflective notes as well to further record my feelings when I relistened to the interviews.

After collating, all the data were inserted into Nvivo software to start the coding process. Kozinets (2019) explained that coding is the process of breaking down the data into pieces and labelling each piece with a specific theme. For the coding process, I had some initial codes based on my interview protocol and expectations. At the same time, when I was coding transcripts, I added some new codes inductively based on my understanding. Every transcript was coded sentence by sentence in Chinese to keep consistency. There are 20 main codes such as Absolute Study Performance Information, Gamification Incentive Design and Motivation. In total, 46 codes in total with some subcodes are included. Through the coding process, it is clear that the gamified designs influence the internal and external motivations of users, and thus the change in the users' behaviours.

The final step is combining: I combined the relevant codes and theorised the pattern of my data with previous literature. At the same time, I triangulate those patterns with the study data of users to further confirm my understanding. From these three steps, it is helpful for me to shed light on the gamified incentive system driven by intrinsic and extrinsic motivation and generate the discussion with previous literature at the end of this paper.

4.4 Empirical Findings

4.4.1 The Online Learning Platform

The empirical setting is a gamified learning mobile application setting in China. This medium-sized platform company has around 500 employees and provides online and in-person education services for millions of Chinese users. As I have been a 5-year loyal user of the mobile application from its start, I witnessed every step of the development of this platform from a customer perspective. It was started by selling recorded audio courses on different subjects such as economics, psychology, finance etc. to its customers. Every course has a daily 10-minute audio session which is written and recorded by some professors in economics, psychology, or management from high-reputational Chinese universities with product editors in this company. Later, they also cooperated with Tsinghua University and Harvard Business School to create some short

video courses for Chinese audiences. As the company has its product quality control SOP, the quality of its audio courses is relatively higher compared with its competitors in the market. The platform also has been editing and updating its product quality control SOP into a book and sharing it with the public to explicitly let its competitors and users learn from it every year.

Another main part of the online service in this application is digital library subscription services. The platform cooperated with many book and magazine publishers to build up a digital library in its mobile application. Users can buy a monthly or annual membership to read any books in its library. As more and more publications cooperate with the platform, this digital library becomes one of the biggest digital libraries in China. Meanwhile, the platform has a team that invite some relevant scholars or "professional paraphraser" trained to read and paraphrase the main ideas of some famous or popular books through around 30-minute recorded audio.

With the development of this platform, in recent years, the platform has started to emphasise offline activities. It tries to attract its existing online users and new customers to come to its mini-MBA program in different cities around China. One of the unique selling points of its program is the high quality of cohorts through its strict selection process. For the online application process, either the candidates need to achieve certain learning credits in the mobile application which will be introduced later, or they need to provide one reference from their current jobs to recommend them to this program. After candidates pass the online application and video interviews on this platform, they can get the offer to attend this mini-MBA program.

This platform aims to foster lifelong learning values or habits in its users. It hopes to broaden users' horizons by studying different audio or video courses outside their current scopes and reading books using their mobile application at any time at any place. To form the users' habits of study, the mobile application also has a check-in button for users to click every day. Based on the number of days that users have already checked in, the platform will reward some vouchers or tokens for users to buy other courses or make subscriptions later on the platform. The application also provides the "Study plan" function for users to add all the courses they plan to learn recently in a list and just click the "play all" button to listen to the list of audio sessions from different courses, which makes users more convenient to study.

To let its users keep studying using its mobile application, during these two years, the platform started to use gamified mechanics in the application such as badges, virtual graduation certificates, points, levels, and weekly and monthly learning reports. Users will get special badges by finishing specific tasks in the application such as learning a session from a course every day for 10 days or reading any books for 15 minutes every day for the whole month (Figure 1, Figure 2). Users will also get a special digital graduation certificate with the professor's digital signature (Figure 3). Moreover, based on every behaviour that users conduct in this application, they will get different level-up badges (Figure 4). For example, based on the number of books finished reading or based on the number of times that users use the internal search engine to search for information, users' "reader" and "searcher" badges will keep upgrading from level 1 to level 10. Some bonuses such as vouchers to buy other courses or renew membership will be given to users with the new level of the badges they reach.

[Figure 1-4 about here]

Each user also has an overall study level from embarking level to top-notch level through accumulating learning credits (Figure 5). Every week, the platform will give you learning credits based on the overall performance from 5 performance perspectives (Notes-taking, Curiosity, Knowledge Sharing, Amount of Study and Study Continuity). From Fresher to Top Leaners, it is getting harder and harder to acquire a learning credit when users' level becomes higher. At this moment, as a top learner, I can only get one point every two weeks. This learning credit and overall user level don't just stop as a symbolic incentive. As this platform cooperates with more and more companies and universities to create new courses, users' learning credits and levels play some roles in real life. Users can show their learning credits and certain courses' graduation certificates to directly get job interview opportunities in some cooperated companies or get some tuition fee and unit credit exemption in some universities when they are doing their undergraduate or postgraduate degrees.

[Figure 5 about here]

The platform also provides social community functions called "Knowledge City-State". Every user can follow other users and comment on others' posts about certain courses or books. Users can also discuss any topic about the courses or post their daily life posts in the City-State. The identity of users can be verified in this community

by this platform if users provide certain proof of job or certificate. Depending on the users' privacy settings, users can get access to certain learning data of other users. For example, if other users' setting is public, I can check how many badges and certificates they get, how many hours they spent today in this application, how many hours they spent since they downloaded this application and which four courses or books they are recently learning (Figure 6 and 7). Users discuss some interesting topics together in this virtual community and share their lives through posts, which is similar to the function of Instagram. However, the platform doesn't include any private chat functions in its application because the founder of this platform claimed that they don't want this study mobile application to become the next social media application.

[Figures 6 and 7 about here]

4.4.2 Joining and Learning on this platform

Before joining this platform, most users knew the founder of this platform from his own shows on a streaming platform. The founder made a weekly online show to discuss a famous book and its relevance to everyone's life, which attracted many audiences to follow the show. When he created this learning platform, previous audiences became the first wave of users. Given that the founder's program is quite a knowledge-intense rather than an entertaining program, these users quite enjoyed learning new things. Therefore, when the platform was established, users followed their interests to choose courses in different subjects such as history, economics, and psychology.

> *I am very interested in psychology… I used various parts of services, whether it's daily audiobook paraphrasing or courses, all related to psychology. Whenever new psychology-related courses are released, I purchase and listen to them. (User 1)*

Based on this intrinsic motivation, many users can feel personal growth and change from learning on this platform. Firstly, they cultivate their habit of learning after they join the platform. Listening to some new courses or reading books on the platform has become a part of their life. Secondly, most participants mentioned the improvement of their cognitive ability, meaning that learning different things is significant to help them think differently, broadly, and proactively. Moreover, another perspective mentioned by many participants is that what they learned from the platform builds up a toolbox to

129

face their real-life challenges, and even keeps company with them in their dark moments.

> *It has turned learning into a habit, the concept of lifelong learning, right? It's not like you stop learning once you graduate from school, but maybe in the past, the idea of continuous learning was indifferent to me. (User 15)*

> *Personally, I think its inspiration for me is not in any single course. I felt it elevates my perspectives when approaching problems (User 9)*

After the initial two years of the establishment of the platform, the platform started to add some gamified elements to its application. Users with intrinsic motivation for learning faced some new external attractions.

4.4.3 Check-in function: domesticating users' behaviour

The check-in function in the application is one of the most often functions discussed by participants. When they open the application, there is a check-in button in the top right-hand side of pages with a red point as a reminder to let users check-in. As one part of their life for some users, they check in every day before they study on the platform.

> *I have a habit now. Every morning, after waking up at 6:00 AM, the first thing I do is open the application and check-in. So far, I have been continuously checking in for over 870 days. (User 31)*

> *After waking up in the morning every day, I open the application to check in and then start to listen to some courses I like (User 29)*

To attract users to check in every day before studying, the platform gives some small bonuses such as coupons or tokens at some key check-in days such as 7 days and 30 days. With the number of check-in days increasing, the bonuses will increase as well. More importantly, if one user reaches 365 days, the platform will send a physical badge with a unique series number to this user's home address. The check-in, as a small task every day, lets users form this habit before learning in the application:

> *I believe the initial reason for using the check-in function was because of its design. After a certain number of days, there would be a small reward, right? … I remembered that initially, my first goal was to get a physical badge*

after one year of check-ins. That made me feel very good at that time, which gave me a sense of purpose. However, after obtaining this badge, if you consistently check in for a year, the act of checking in becomes a habit. Nowadays, you will probably check in first before listening to any courses. (User 3)

Users were seduced by the small rewards at the beginning of check-in. The physical badges after one-year check-in keep motivating users to do so, which also satisfies users' competency needs from this small task. However, more importantly, through this small task, the platform domesticated their users' behaviour to cultivate users' habit of opening the application. When check-in becomes a habit, users won't stop checking in every day:

This hasn't brought any psychological burden to my life. It's like a subconscious action for me now, which becomes as natural as I grabbed my phone when I wake up in the morning. (User 5)

A few days ago, when I reached 900-day check-in, I posted it on my social media. The application leads you to do this action. I think there will be a greater element of loss aversion. I don't want to break the streak after accumulating for such a long time, even though it is just a number. I don't want it to be disrupted. (User 24)

From the quotations above, not only does the check-in function become a habit of the users, but it also becomes a symbolic motivation to let users keep doing this action to keep the number. The number of check-in days already becomes an important intangible asset for users to reserve. If the number disappeared, there would be a great emotional loss for their effort over these years. To avoid those psychological pains, it is better to keep this habit as the first thing to do before starting learning.

Some users do not like this check-in function for some reason. Firstly, although they tried to check in every day, they were annoyed about this design after they failed to keep this action a few times, especially after 100 or 200 days. The strong loss aversion effect demotivated them to check-in. Therefore, they complained about the function during the interviews:

I felt this check-in function in the application is not good at all. It is quite

frustrating when I logged in and listen to the courses every day but sometimes, I just forget to check in. (User 30)

When I used this function, I still remember the first break of my 300 days of continuous check-in. I was almost there to get a physical badge, but because of the international flight and time differences, I didn't click the button in time and then everything started from zero. It frustrated me a lot for the following weeks that I stopped checking in for a few weeks though I was using this platform to learn during that period. Secondly, some users are not interested in those small bonuses and the physical badges to keep doing this action. They just focused on the courses they want to learn:

I don't think I need to check in every day. The application interface has many attention-grabbing elements. It's like walking into a shopping mall where I know what I want to buy. I will only look at those specific things. If I start to browse around, that will distract my attention. (User 21)

Though there are some users not motivated by this design, the existence of this function does domesticate some users' behaviour to open this application at least to check in as one part of their life no matter what kinds of extrinsic motivation incentivise them. However, letting users open the application every day is only the first step. How to make them learn actively, gamified challenges and badges stepped in.

4.4.4 Gamified Challenges and Badges: Boosting Intrinsic Users' Motivation temporarily

With other entertaining social media such as TikTok becoming more popular, it is hard for an online learning application to get some users' spare time. Therefore, most users who want to spend their spare time studying some new courses within or outside their career do keep their intrinsic motivation to broaden their horizons:

This application actually broadens my knowledge in various areas. We often talk about the importance of being a T-shaped person, expanding our horizontal bar… For example, before Professor Liu's course, I never thought I would listen to any courses related to Chinese traditional culture before. I find it fascinating and interesting now. (User 39).

During the interviews, all of the participants expressed their interest in learning certain courses or reading books in this application. It is also one of the most popular factors

why choosing and buying certain courses on this platform. To further boost users' intrinsic motivation to finish what they have already bought and keep reading books on their platform, the platform introduced gamified challenges and badges. For instance, if users finish certain challenges such as monthly reading challenges by reading any book for 15 min every day, they will get a specific monthly reading badge. When we discussed this new gamified element with participants, I find that these extrinsic motivations from gamification do generate motivational effects at the beginning of the introduction. However, this motivational effect does not last in the long term:

> *I think its study of gamification design, including the sense of ritual for users, is quite well done, similar to the incentive we receive when playing games. I believe this design is beneficial. There was a time when I was quite intrigued by this, but I don't care anymore. It's because I eventually realised that it doesn't hold much meaning or significance (User 27).*

> *In the beginning, my friends who used this application would share and find these badges to be motivated. I also found it quite novel, and there was a sense of achievement after certain challenges. However, as time went on, I feel that the motivational effect of these badges became less noticeable for me. I tend to be more inclined towards self-motivation to learn courses that I am genuinely interested in, rather than relying solely on badges or external rewards. (User 23)*

> *You don't feel the novelty or the sense of interest anymore. It's like the relationship between dopamine and endorphins. Happiness is fleeting and provides a temporary sense of achievement … I think those gamified features might have some impact on new users, but personally, I normally ignore them automatically. (User 12)*

From these quotes, the gamified badges did generate a temporary motivational effect on the users' study behaviour. Complementing users' intrinsic motivation, it added some fun and novelty at the beginning. However, when users realise that these symbolic badges have less real meaning and get bored with this design, they won't pay much attention to the challenges and badges but are more inclined to their intrinsic motivation to learn new things from this motivation.

Except for one participant who is enthusiastic about collecting certain badges to buy and learn different courses, most of the users just only tried to get the badges related to the courses they have already bought, and services they have already subscribed to. In other words, getting gamified challenges and badges only play a subordinate role in motivating users to learn certain courses, which won't influence too much on the users' purchase decisions. I still have been trying to finish as many challenges as I can to get different special badges within my purchase. For me, these challenges help me organise my study behaviours on the platform: first I listen to some courses I bought. Then I listened to the daily paraphrase of a book, and I finish my learning by reading books I am interested in for 15 minutes every day. At the end of the month, I would get certain badges if I keep this behaviour. Therefore, like the check-in function, these challenges with badges can also to some extent domesticate users' behaviour.

Interestingly, when participants are asked about their feelings about the physical badges they received from check-in and the virtual badges they got from finishing challenges in the application, most of the participants who get both kinds of gamified badges tend to prefer the motivation and happiness brought by the physical badge:

> ***It's different when you have something tangible that you can hold in your hands… If it were a virtual badge, you would have to open it in the application. Initially, you might open and take a look when you first receive it, but later on, you might never open it again, right?... If you have been spending time learning on this platform, and then it also gives you a physical badge, it's like that time is transferred or embodied in that item. When you see it, you can feel a connection to what you've learned before, and it brings this kind of feeling. (User 22)***

When we explored further where this motivation for gamified badges comes from, most participants mentioned a few reasons also shown in the quotes above. First of all, the gamified badges and the virtual graduate certificates bring a sense of ritual to some users who are more inclined to have this feeling in their daily lives. They felt that these designs help them witness and record their important moments of using this application as a milestone, given there is a specific date and time provided when they get these badges or certificates. Secondly, these badges and certificates, as a kind of symbolic social currency, provide users with a good resource for them to post on their social

media such as WeChat. It creates a good impression to friends and families that they like learning new things and reading books, which serves some users' vanity:

> *Sometimes I will push myself to complete the challenges. I think it is still a form of vanity. Like the May or June Reading Badges, it's like a personal label. I am willing to get the badges. In another sense, I want others to see I have many badges and recognise my achievement, right? (User 10)*

> *It is important to acknowledge that the love for learning itself can sometimes stem from a sense of superiority. Where does this sense of superiority come from? It comes from the fact that this platform provides you with high-quality materials. These materials such as gamified badges or virtual graduation certificates give you materials to share on your social media. (User 8)*

From these quotes, these badges bring recognition from others around users. The external gamified elements internalise to users to create part of their value. Not only do users get motivation and recognition from badges and certificates they got in the application, based on their data of every use behaviour, but they also get a learning credit to evaluate them.

4.4.5 Learning Credit: Building up the Symbolic Recognition System among Users

When the platform introduced the learning credit system, it becomes a popular and representative data among users of this application. Based on users' five performance measures (Notes-taking, Curiosity, Knowledge Sharing, Amount of Study and Study Continuity) on the platform, it splits users into 5 categories (Embarking, Advancing, Excellent, Outstanding, Top-notch)[27]. Most users are not familiar with the title of each category, but they are sensitive to the number of the learning credit itself. When they heard some users reach above 800 credits, they give great respect to these users:

> *When you said your credit is over 890, I was like 'Wow!' It immediately raised my admiration for you. Although I hold a certain level of respect for individuals with a doctoral degree as I met some doctors around me, the admiration level is not that high. (User 39)*

---

[27] Learning credit: Embarking: 300-350, Advancing: 350-500, Excellent: 500-700, Outstanding: 700-800, Top-notch: 800-950

The learning credit motivated users through different channels. It guides users to conduct certain behaviours in the application. Based on the five performance measures, some users tried to improve their credit. For example, some users will share more courses and books on their social media to improve the Knowledge sharing measure. Some users will try to take more notes and comment on others' notes to improve their notes-taking perspective. However, the learning credit system only gives an overall performance level of each measure[28] that users reached without giving users clear rules such as a specific number of times they share courses or a number of notes they take to reach the next level. Moreover, for the overall number of learning credits, the platform will decide whether to give users 1 point or more points at the end of the week based on every behaviour the user did on this platform during the work, which is all calculated by the algorithm according to users' study data. Therefore, some users complained that the platform only gives five performance measures without giving specific rules to improve each measure. From the platform perspective, it avoids some gaming behaviours that want to boost their learning credit in the short term. Some users do give them a certain learning credit target to achieve:

> *Although my credit is not high at this moment, I will set a goal for myself. For example, this year my goal is to surpass 700 credits. It gives me some motivation to continue learning and maintain a sense of purpose, which is quite important. (User 29)*

Given the difficulty of increasing the credit, the second channel of motivation is the recognition of other users and the symbolic meaning behind it, which coincides with the gamified badges to serve individuals' sense of achievement and vanity:

> *There aren't many people who like learning. So, someone like me who has around 700 credits will be always labelled as someone who likes learning. Actually, I felt it was pretty good. (User 9)*

> *If you have high credit, sometimes there may be a halo effect, establishing a hierarchical system. (User 19)*

More importantly, the learning credit, as the most representative performance measure, plays a crucial role in generating social comparison among users on this platform, which

---

[28] The levels of each performance measure are B+, A-, A, A+, S.

further motivates users to use this application. When users meet in some in-person activities, comparing the learning credit often happens to show everyone' study situation on this platform:

> *Because the learning credit is a visualised measure, people might compare them. For example, we organised a New Year's Eve speech event where there is an award based on the learning credit. We compared our credits and gave the award to the top performer. (User 7)*

> *We encouraged everyone to share their learning credit and create a ranking list of who is the first and who is the second…Whoever has the highest scores is the top performer and gets our likes and recognition. This is a kind of social currency, right? This creates a positive experience for us. (User 12)*

The learning credit system creates a symbolic recognition system to motivate users to keep using this application through three channels: providing vague five performance measures to influence users' performance, creating symbolic meaning for high performers and generating social comparisons among users to further enhance the recognition of top performers.

4.4.6 Learning Performance Information: Providing Feedback to Users

As a crucial part of gamification, learning performance information is shown in the application. Every user can check different data of almost every activity they did from how many hours they studied since they joined the platform which exacts to one decimal place to how many minutes they used today, from how many notes they took to how many times the search engine they used. Some of them are real-time data, the others are accumulated data. Based on this data, the platform provides weekly, monthly, and annual learning reports to each user to summarise their learning situations. Most users reflected on the annual report they received at the end year.

> *It has some annual summary, which is a good thing. It summarises your entire year of learning including study time and contents. It gives you feedback, which allows you to have a general understanding of how you invested your time. (User 32)*

> *It's like giving you a chance to be seen. You might not have realised that you are a hardworking person, but the platform uses data to show that your past*

effort was visible and memorable to others. Being acknowledged satisfies psychological needs. (User 34)

I understand it as a positive reinforcement. Only positive feedback can motivate people to continue doing things (User 35)

These quotes show that the annual report gives users positive feedback to help them remember what they did on this platform this entire year. This firstly serves their competency needs as they spend many hours finishing different courses and reading books. From those learning, they felt personal growth as shown before. More importantly, this growth is recorded by different figures in the report, which makes it observable and quantifiable. This annual report serves the relatedness need meanwhile because through this design users' growth can be seen by the platform and by people close to them. Like badges and certificates, they shared their report on their social media to show their learning performance.

Last year, I made a special post on my social media for the report. I thought it was a great record. The report said the course I learned most is Dr. Wan's course. (User 14)

By the end of the year, it is a milestone for everyone. Therefore, the annual report from the platform has a sense of ceremony every year. I also downloaded it and shared it in the reading WeChat group among users. (User 12)

Reflecting on this annual report, I felt that this is a quite touching element to influence users emotionally. It seems that the platform or the algorithm and data know us better than ourselves. In the report, when it told me that on which day I was hardworking to study certain courses until quite late, or when it told me how many word counts I wrote the comments or took notes while learning during this year, it was quite a powerful and touching moment that in front of the figures I realised that I spent lots of effort in my spare time to learn new things and broaden my horizons. Since this report measures our study behaviour, it is much more meaningful for users than some other annual reports from other entertaining applications.

To conclude, by using different gamified elements and providing study performance information, this platform further motivates intrinsically motivated users

to keep learning different courses and using different services. Next section, I will discuss these findings with SDT and previous literature.

4.5 Discussion

Previous literature argued the importance of SDT, as one of the main mechanisms for designing gamified elements and explaining the motivational effect of gamification (Landers *et al.*, 2015; Cardador, Northcraft and Whicker, 2017; Sailer *et al.*, 2017), provides a complete picture of how intrinsic motivations and extrinsic motivations incentivise people's behaviour by satisfying three basic psychological needs (Deci and Ryan, 2000). From the last part, users started from their intrinsic motivation of learning new things and recognition of the founder to join the platform, and then the different extrinsic gamified elements jumped into influencing users' behaviour. This particular field setting gives me the chance to discuss how both kinds of motivation function together to change users' behaviour and shed light on how organisations design their gamified incentive systems. In this part, I will first discuss how different users treated the gamified designs differently, then how external motivation internalises to an internal part of users' self, and finally how this setting implicates the motivation and incentive systems in the workplace.

4.5.1 Different kinds of users on the platform

The autonomy need is defined as people having chances to decide what they are willing to do, and their behaviour depends on their own self-regulation and volition (Ryan and Deci, 2017). To satisfy this psychological need, gamification normally includes different game elements and gives the users a large autonomy to choose whether and how they want to engage in the systems. Based on the different characteristics, Bartle (1996) categorised these players into four different roles: achiever, explorer, socialiser, and killer. In our setting, users have different particular focuses and habits to engage in different designs on this platform, which is hard to categorise into four parts. Meanwhile, given this is not a tournament game and there is no user learning leaderboard provided at this moment, it is also hard to have a killer of this particular kind of user who has the main purpose of beating other users. More importantly, sometimes based on different needs they want to satisfy, the role of users is also changing in different circumstances instead of holding one role on this platform.

Most users join the platform to study as an explorer they start to discover which

courses or books they want to learn or read. After exploring and deciding on the content to learn on this platform, they explored the relevant gamified elements on this platform. Users can decide which element of gamification they want to engage in. From the study data collected from participants, 8 users engaged in most of the gamified elements. They checked in every day, finished different kinds of challenges to get many badges, finished different courses to get virtual certificates, and thus had a high learning credit. It is important to notice here that at this moment, these users are not only an explorer but also an achiever. They tried to get more badges and certificates during their learning and achieved a high learning credit and a high number of continuous check-in days. The gamified elements add more fun to the learning and guide users to certain achievements, which further helps to satisfy the competency needs of users. They feel that not only is their cognitive ability enhanced but also their improvement of ability and their effort are seen and recognised by the platform through these gamified elements.

The same is true for other explorers on the platform. Holding their autonomy in their hand, they choose the elements which they feel meaningful to play. For example, some users stick with checking in every day but are not interested in badges, but others like to collect the badges during their learning without bothering by the check-in button. No matter which elements users choose to engage shows the satisfaction of users' autonomy needs in gamification. Even if some participants care nothing about the gamified elements but the learning contents only, being a pure explorer and noticing they have some fun designs to engage if they want is also evidence of autonomy and need satisfaction. The autonomy need doesn't mean that people have to choose from different options. Instead, noticing that they have the right to choose nothing from options also shows their autonomy. As Kelly, Valtchanov and Webb (2021) suggested, the platform gives users choices of different gamified elements to motivate themselves above the original intrinsic motivation.

However, the same temporary effect also picked up in this setting (Kelly, Valtchanov and Webb, 2021). The role of achiever might last for a short period. As shown in the findings, when users still feel novelty and fun from the gamified elements, they will try to get more badges and certificates. With time going by, most of them get back to their existing habits to learn on the platform at their own pace, especially for gamified badges. There are some mechanisms behind it (Zichermann and Cunningham,

2011). To let gamified badges work well, firstly, they need to have scarcity. When users can easily get a badge for any activity they are doing on the platform, some participants mentioned that the value of badges is deflated because of the lack of scarcity. More importantly, Zichermann and Cunningham (2011) also argued for the sociality of gamified badges. Chapman, Chua and Fiedler (2021) show how important the social symbolic value of badges is in seducing users to keep collecting the badges. On the contrary, if the badges can not be a social currency among users to generate discussion, users lose interest in collecting them. This is an exact problem for this platform. Compared with the learning credit and the annual report, gamified badges are hidden in the application. At the same time, given that users focus on the different parts of learning on the platform motivated by too many different badges, it is hard to compare the number of badges or kinds of badges with other users. Therefore, the gamified badge system without scarcity and sociality can only generate a short-term effect on motivating behaviours.

Some users on this platform are more inclined to the socialiser role and are more willing to develop their network through using this platform. There are 10 participants in our sample following more than 500 users on the platform. As the learning characteristic of this platform, it attracts many middle-class or above users who are the talents of their industries. Therefore, it is a good platform to link to some experts based on their verification of the profiles. During their learning, socialisers like to engage in the City State Community function to participate in some discussions about certain courses or books with other participants. The gamified designs also satisfy the relatedness need of users (Ryan and Deci, 2017), especially for socialisers. They shared their learning credit or annual reports on their social media to feel vanity and recognition from others.

This connection also extended to in-person activities. Some users meet together at a certain point to share their learning experiences, deal with certain challenges in their lives, or have certain social events together, which gives a sense of community among them (McMillan and Chavis, 1986; Chapman, Chua and Fiedler, 2021). As mentioned in the findings part, during these in-person events, users compared their learning credit and study data with others, and top performers get huge recognition and respect from others. Other users treated them as a model to learn from, which is an

upward assimilative social comparison (Smith, 2000; Carroll and Marginson, 2021). These online and in-person events further satisfy the relatedness need and power of the existing intrinsic motivation of learning.

To conclude, to satisfy basic psychological needs, users play different player roles on the platform. Initially, most of them are explorers to discover what they want to engage in. With the gamified elements added in, some users are more inclined to become an achiever to perform different tasks to get badges and other rewards or improve their learning credit. The other users are more inclined to link to other users as a socialiser during their learning and enjoy the feeling of learning with many other talented users online or in person. From this setting, I argued that the autonomy from the gamification created this dynamic to let users change in the different player roles or even play multiple roles at the same time. Which roles they want to play mainly depends on psychological needs satisfaction. As described by Zichermann and Cunningham (2011), people might actually have different proportions of certain player roles in the same gamified system. Future research can further quantify and analyse the different player roles in other gamified settings by using some simulation experiments and further discover the effect of needs satisfaction and motivation (Sailer *et al.*, 2017).

4.5.2 The internalisation of the extrinsic gamified motivations

Internalisation is the process in which people transform extrinsic values, beliefs, and controls into part of themselves (Ryan, Connell and Deci, 1985). Ryan and Deci (2017) emphasise the importance of internalisation to help people satisfy their basic psychological needs for relatedness, competence, and autonomy. How to internalise the extrinsic motivations from gamification is important for its design (Zichermann and Cunningham, 2011). Pfister and Lukka (2019) explained how employees internalised their stretched goals through the interrelation of the goals with personal and cultural controls of the company. Different from their setting, users joined the platform because of their intrinsic motivation for learning. When extrinsic gamified elements are added to this platform, there will be the potential risks of crowding out effect, which means the extrinsic motivation will reduce users' effort and thus performance (Frey, 1997). To keep learning on the platform, users internalised those gamified elements in different types: introjection, identification and integration.

As a function with small bonuses and physical badge rewards, the check-in

function is internalised by some users through an introjection process initially. Whether users click the check-in button or not complies with some internal demands of users such as the influence on self-esteem and feelings of pride themselves (Ryan and Deci, 2017). Users felt quite upset and doubted themselves when there was a break. The strong loss aversion psychological effect motivates users to continue checking in. When users achieved 365 days of check-in, they felt proud of themselves, which is not only about the physical badges they got but also the activity itself they achieved. After they achieved this target, the internalisation process of some users changed to an identification process or even an integration process, which are two internalisation processes with more autonomy so that people can see the importance and value of doing certain things which are partially or fully congruent with people's own value and beliefs (Ryan and Deci, 2017). When users treated the check-in function as the part of their life that they naturally finish this activity every morning before learning, this external gamified design fully integrated into users' minds to become a "highly stable and mature form of self-regulation" (Ryan and Deci, 2017, p. 189). However, not every user successfully internalises or chooses to internalise this check-in function because of the negative emotion from the previous break or simply ignorance of the rewards from check-in in the first place.

The internalisation process of gamified badges and virtual certificates is relatively hard to achieve for most users, especially in the long term. These kinds of external rewards from learning on the platform also internalise in an introjected way at the beginning that users can influence self-esteem and self-satisfaction when they get a new badge. However, when badges are not scarce and have a low social value, the internalisation actually fails, and these badges and certificates are only external rewards to motivate users without achieving a good integration with users' intrinsic motivation. Future research can dig into the scarcity effects of gamified badges to explore how to reward the badges to achieve a deeper internalisation process.

Compared with gamified badges, the learning credit and the annual report elements in the application are better internalised into users' intrinsic motivation. These elements, as symbolic social currency in the platform, tried to satisfy users' relatedness needs. Through social comparison with other users (Festinger, 1954), they can get recognition from this learning community and meanwhile, learn from users who have

a higher learning credit. This can make some users feel the real value and importance of learning on the platform, which achieves an identification process. Since the rules of the learning credit system are vague, this further enhances the value and respect to users with high credit. To reach a high credit, they have to improve most of the 5 performance measures during the learning process instead of gaming the learning credit score system. Similar to the learning credit, the annual report makes users feel that their hard-working is seen and recorded by the platform. At the end of the year, the platform offers annual reports to users to explicitly point out the remarkable data of learning during this year, which motivates users to keep using it by satisfying their competence needs. Therefore, some users identified the value and importance of this annual report as positive feedback.

To improve the users' experience and retain old users, the platform tries to implement different gamified elements and internalise them as part of internal motivation to further complement the existing intrinsic motivation of learning. Although some users quite focus on the learning content itself and are immune to these gamified elements, different kinds of internalisation processes of these external elements do happen naturally to most users during their use to transform external motivations into part of internal motivations. This study extends Pfister and Lukka (2019) to explain how different kinds of internalisation processes happen in a gamified setting with both intrinsic and extrinsic motivation. Moreover, this study also shows how psychological need-supportive gamified designs are important to change people's behaviour (Bouten and Hoozée, 2022). Given the limitation of field research, future research can further examine the magnitude of the effect of different kinds of internalisation on the change of users' or workers' effort and performance.

4.5.3 Inspiration of gamified incentive system in work settings

Although this is a setting about how the platform attracts and retains its users to keep learning through gamification, it does shed light on the control and incentive systems of employees in organisations. Cardador, Northcraft and Whicker (2017) argue that work gamification also tries to motivate employees more intrinsically by making them feel rewarded for their work itself. They also suggested that work gamification is not transforming work into a game but implementing the game elements to incentivise expected work behaviour. Kelly, Valtchanov and Webb (2021) find that there is a positive effect of using online slot machines on the number of training modules

employees finish and the effort they make in the related quizzes. However, they also find the temporary effect of gamification that the motivational effect only lasts in the short term, and the influence only applies to a subset of employees.

Consistent with what they found, I realised that some gamified elements such as badges and certificates can only motivate users at the beginning of initiatives. The reason is mainly the lack of novelty, scarcity, and social currency value of the gamified badges. On the contrary, Chapman, Chua and Fiedler (2021) show the constant attractiveness and fun of the gamified badges designed by Foursquare because of their strong social value among users. To use gamification effectively in a work setting, we have to be aware of the engagement of the gamification. With time going by, engagement will start to diminish because the workers get bored of the gamified system (Hamari and Koivisto, 2015). Therefore, as proposed by Cardador, Northcraft and Whicker (2017), organisations can apply gamification to some contexts where the tasks have a clear time horizon, where the sub-goal achievement is closely linked to the final task outcome, and where there is a change group of workers. Future research can further examine and find the optimal period length to introduce the gamified elements through lab or field experiment designs.

As for only engagement for part of employees, this study used Bartle (1996)'s model of four different player roles to explain why different users have different preferences and attitudes to different gamified elements. If there is only one single game design in the system (Kelly, Valtchanov and Webb, 2021), it is hard to attract every worker's attention. Zichermann and Cunningham (2011) argued that including different gamified mechanics in the system, which achievers, explorers, socialisers, and killers can benefit from, is the optimal way to implement gamification. We have to realise the different preferences and characteristics of workers. This is also related to the satisfaction of their autonomy needs. When organisations implement different gamified elements in an incentive system, workers can choose to focus on the elements they are interested in autonomously and gamify their work in their own way (Cardador, Northcraft and Whicker, 2017). The workers will have more control over what they decide to do, which can achieve a better internalisation process to integrate these external elements into their intrinsic motivation for the work itself.

Related to previous accounting literature (Bouten and Hoozée, 2022; Pfister and

Lukka, 2019), this study continued to mobilise SDT to further explain the different internalisation processes of different external gamified elements. Instead of integrating the stretched targets with a traditional control system, using gamification to internalise the external rewards and controls into people's part of self is more natural and enjoyable without generating a backlash from users or workers. Based on the extent to satisfy three psychological needs, different kinds of internalisation processes exist for different gamified elements (Ryan and Deci, 2017). Organisations should try their best to reach more autonomous internalisation such as identification and integration. Only if organisations internalised their gamified incentive systems or other controls to a deeper level, workers would enjoy their tasks more and realise the value and importance of their tasks for themselves. This study also confirms the motivational effect of peer recognition systems (Black, 2023). Through the learning credit and annual reports, users got recognition and respect from other users, which satisfied their needs for competency and relatedness. The process of getting access to my participants in this study is also evidence of generating a help effect among peers (Black, 2023). As these participants recognise my superuser identity through my learning credit and data, most of them are quite willing to help with my study. Therefore, gamification can also help an organisation to create a peer recognition system to further motivate workers' performance and helping behaviours within the organisation.

4.6 Conclusion

In this study, I conduct a field study in an online learning platform using gamification to motivate their users' behaviours. Through the interactive approach of netnography (Kozinets, 2019), as a superuser of this platform, I immersed myself in this setting to feel the change in my behaviours and conducted 39 semi-structured with users on this platform to collect the field data. Mobilising SDT (Deci and Ryan, 2000; Pfister and Lukka, 2019; Bouten and Hoozée, 2022), I find that the whole gamified incentive system with many gamified elements such as daily check-in button, gamified challenges and badges, and learning credits influence the users' behaviour significantly. Based on different preferences, users play different player roles to engage in the different gamified elements. Some users fully internalise some gamified elements such as the check-in button as part of their own daily lives. These gamified elements to some extent further enhance users' intrinsic motivation to learn and retain the users on this platform. Furthermore, some gamified elements such as the learning credit system, as

a symbolic social currency, build up a peer recognition system to create a strong connection and trust among users.

This study contributes to literature on the gamification literature (Kelly, Valtchanov and Webb, 2021; Chapman, Chua and Fiedler, 2021; Cardador, Northcraft and Whicker, 2017; Sailer *et al.*, 2017) to explain how gamification integrates both intrinsic and extrinsic motivation of users or workers to satisfy their basic psychological needs. Meanwhile, this study adds more understanding of the novel incentive and control system in management accounting literature (Choi and Presslee, 2023; Black, 2023; Berger, Guo and Presslee, 2023; Vogelsang, 2022; Kelly, Presslee and Webb, 2017; Presslee, Vance and Webb, 2013) to generalise the potential of gamified incentives in the working setting. As the limitation of the field study, it cannot examine the magnitude of the effect of gamification on the users' behaviours and performance. Future studies can further test the effect by using experiments or survey methods.

In terms of practical implications, organisations need to consider the different preferences of workers to include vivid gamified mechanics to let workers choose their interested mechanics autonomously. Moreover, organisations should consider the temporary effects of gamification to use it in some suitable contexts and observe the engagement situation consistently to decide whether to stop or update the gamified elements to regain the novelty. Finally, the design of gamification should be natural and clearly engage in the intrinsic motivation of workers. As Zichermann and Cunningham (2011, p. 29) mentioned "The better a designer knows his players, resulting in a better game design, the less it will feel to the player like being on a wheel, and the more it will feel like it was her idea to begin with. That's the holy grail of gamification: a game so well designed that the player's actions just feel normal. We believe it can be done in almost any experience."

4.7 Reference

Ahrens, T. and Chapman, C. S. (2006) 'Doing qualitative field research in management accounting: Positioning data to contribute to theory', Accounting, Organizations and Society, 31(8), pp. 819-841.

Bartle, R. (1996) 'Hearts, clubs, diamonds, spades: Players who suit MUDs', Journal of MUD research, 1(1), pp. 19.

Baxter, J. A. and Chua, W. F. (1998) 'Doing field research: practice and meta-theory in counterpoint', Journal of Management Accounting Research, 10, pp. 69-87.

Bebbington, J., Österblom, H., Crona, B., Jouffray, J.-B., Larrinaga, C., Russell, S. and Scholtens, B. (2020) 'Accounting and accountability in the Anthropocene', Accounting, Auditing & Accountability Journal, 33(1), pp. 152-177.

Berger, L., Guo, L. and Presslee, A. (2023) 'Motivating Employees with Goal-Based Prosocial Rewards', Contemporary Accounting Research, 40(1), pp. 231-256.

Bialecki, M., O'Leary, S. and Smith, D. (2017) 'Judgement devices and the evaluation of singularities: The use of performance ratings and narrative information to guide film viewer choice', Management Accounting Research, 35, pp. 56-65.

Black, P. W. (2023) 'The effect of peer-to-peer recognition systems on helping behavior: The influence of rewards and group affiliation', Accounting, Organizations and Society, pp. 101454.

Bonner, S. E. and Sprinkle, G. B. (2002) 'The effects of monetary incentives on effort and task performance: theories, evidence, and a framework for research', Accounting Organizations and Society, 27(4-5), pp. 303-345.

Bouten, L. and Hoozée, S. (2022) 'Hybridity in a hotel chain: designing a package of controls to sustain a hybrid mission', Accounting, Auditing & Accountability Journal, 35(8), pp. 1862-1891.

Callan, R. C., Bauer, K. N. and Landers, R. N. (2015) 'How to avoid the dark side of gamification: Ten business scenarios and their unintended consequences', Gamification in education and business, pp. 553-568.

Cardador, M. T., Northcraft, G. B. and Whicker, J. (2017) 'A theory of work

gamification: Something old, something new, something borrowed, something cool?', Human resource management review, 27(2), pp. 353-365.

Carroll, E. and Marginson, D. (2021) 'Relative performance information and social comparisons: Exploring managers' cognitive, emotional and dysfunctional behavioral processes', Management Accounting Research, 53.

Chapman, C., Chua, W. F. and Fiedler, T. (2021) 'Seduction as control: Gamification at Foursquare', Management Accounting Research, 53, pp. 100765.

Chapman, C., Sawabe, N., Kimura, A. and Selmes-Suzuki, H. (2023) 'Pursuing serendipity in the governance of academic performance and participation in the field'.

Chen, C. X., Lill, J. B. and Vance, T. W. (2020) 'Management control system design and employees' autonomous motivation', Journal of Management Accounting Research, 32(3), pp. 71-91.

Choi, J. W. and Presslee, A. (2023) 'When and why tangible rewards can motivate greater effort than cash rewards: An analysis of four attribute differences', Accounting, Organizations and Society, 104, pp. 101389.

DeCharms, R. (1968) Personal Causation: The Internal Affective Determinants of Behavior. Academic Press.

Deci, E. L. (1975) 'Conceptualizations of Intrinsic Motivation', in Deci, E.L. (ed.) Intrinsic Motivation. Boston, MA: Springer US, pp. 23-63.

Deci, E. L. and Ryan, R. M. (1985) 'Conceptualizations of intrinsic motivation and self-determination', Intrinsic motivation and self-determination in human behavior, pp. 11-40.

Deci, E. L. and Ryan, R. M. (2000) 'The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior', Psychological inquiry, 11(4), pp. 227-268.

Deterding, S., Dixon, D., Khaled, R. and Nacke, L. 'From game design elements to gamefulness: defining" gamification"'. 2011, 9-15.

Festinger, L. (1954) 'A Theory of Social Comparison Processes', Human Relations, 7(2), pp. 117-140.

Frey, B. S. (1997) Not just for the money. Citeseer.

Frey, B. S. and Jegen, R. (2001) 'Motivation crowding theory', Journal of economic surveys, 15(5), pp. 589-611.

Frey, B. S. and Osterloh, M. (2001) Successful management by motivation: Balancing intrinsic and extrinsic incentives. Springer Science & Business Media.

Hamari, J. and Koivisto, J. (2015) 'Why do people use gamification services?', International Journal of Information Management, 35(4), pp. 419-431.

Jeacle, I. and Carter, C. (2011) 'In TripAdvisor we trust: Rankings, calculative regimes and abstract systems', Accounting, Organizations and Society, 36(4-5), pp. 293-309.

Kelly, K., Presslee, A. and Webb, R. A. (2017) 'The Effects of Tangible Rewards versus Cash Rewards in Consecutive Sales Tournaments: A Field Experiment', The Accounting Review, 92(6), pp. 165-185.

Kelly, K., Valtchanov, D. and Webb, A. (2021) 'Behavioral implications of using an online slot machine game to motivate employees: A cautionary tale', Accounting, Organizations and Society, 89.

Kornberger, M., Pflueger, D. and Mouritsen, J. (2017) 'Evaluative infrastructures: Accounting for platform organization.', Accounting, 60, pp. 79.

Kozinets, R. V. (2019) Netnography: The essential guide to qualitative social media research. Sage.

Landers, R. N., Bauer, K. N., Callan, R. C. and Armstrong, M. B. (2015) 'Psychological theory and the gamification of learning', Gamification in education and business, pp. 165-186.

Leoni, G. and Parker, L. D. (2019) 'Governance and control of sharing economy platforms: Hosting on Airbnb', The British Accounting Review, 51(6), pp. 100814.

Marginson, D. E. W. (2004) 'The case study, the interview and the issues: a personal reflection', The real life guide to accounting research: Elsevier, pp. 325-337.

McMillan, D. W. and Chavis, D. M. (1986) 'Sense of community: A definition and theory', Journal of Community Psychology, 14(1), pp. 6-23.

Patton, M. Q. (2002) 'Two decades of developments in qualitative inquiry: A personal, experiential perspective', Qualitative social work, 1(3), pp. 261-283.

Pfister, J. A. and Lukka, K. (2019) 'Interrelation of controls for autonomous motivation: A field study of productivity gains through pressure-induced process innovation', The Accounting Review, 94(3), pp. 345-371.

Presslee, A., Vance, T. W. and Webb, R. A. (2013) 'The effects of reward type on employee goal setting, goal commitment, and performance', The Accounting Review, 88(5), pp. 1805-1831.

Ryan, R. M. and Connell, J. P. (1989) 'Perceived locus of causality and internalization: examining reasons for acting in two domains', Journal of personality and social psychology, 57(5), pp. 749.

Ryan, R. M., Connell, J. P. and Deci, E. L. (1985) 'A motivational analysis of self-determination and self-regulation in education', Research on motivation in education: The classroom milieu, 2, pp. 13-51.

Ryan, R. M. and Deci, E. L. (2017) Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Publications.

Ryan, R. M. and Moller, A. C. (2017) 'Competence as central, but not sufficient, for high-quality motivation', Handbook of competence and motivation: Theory and application, pp. 216-238.

Ryan, R. M., Rigby, C. S. and Przybylski, A. (2006) 'The motivational pull of video games: A self-determination theory approach', Motivation and emotion, 30, pp. 344-360.

Sailer, M., Hense, J. U., Mayr, S. K. and Mandl, H. (2017) 'How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction', Computers in human behavior, 69, pp. 371-380.

Smith, R. H. (2000) 'Assimilative and contrastive emotional reactions to upward and downward social comparisons', Handbook of social comparison: Springer, pp. 173-200.

Van den Bussche, P. and Dambrin, C. (2020) 'Peer-to-peer evaluations as

narcissistic devices: fabricating an entrepreneurial community', Accounting, Auditing & Accountability Journal.

Van der Hauwaert, E., Hoozée, S., Maussen, S. and Bruggeman, W. (2022) 'The impact of enabling performance measurement on managers' autonomous work motivation and performance', Management Accounting Research, 55, pp. 100780.

Van der Stede, W. A. A., Wu, A. A. and Wu, S. Y.-C. A. (2020) 'An Empirical Analysis of Employee Responses to Bonuses and Penalties.', Accounting Review. Nov2020, 95(6), pp. 395.

Vogelsang, T. (2022) 'Time is (Not) Money: Incentive Effects of Granting Leisure Time', European Accounting Review, pp. 1-25.

Weibel, A., Wiemann, M. and Osterloh, M. (2014) 'A behavioral economics perspective on the overjustification effect: Crowding-in and crowding-out of intrinsic motivation', The oxford handbook of work engagement, motivation, and self-determination theory, pp. 72-84.

Zichermann, G. and Cunningham, C. (2011) Gamification by design: Implementing game mechanics in web and mobile apps. " O'Reilly Media, Inc.".

Tables and Figures

Table 1 User Interview Details

Participant	Length	Gender	AGE	Industry	Study Time (hours)
1	0:59:38	M	20-30	Psychology	1481.7
2	0:24:32	F	20-30	Law	100.8
3	0:43:36	F	40-50	Psychology	2747.5
4	00:34:36	M	20-30	Marketing	NA
5	01:11:15	M	30-40	Sales and IT	2430.8
6	00:34:31	F	30-40	Design	77.8
7	00:38:42	F	30-40	Public Service	2488.6
8	00:49:47	M	30-40	Consulting	3035.9
9	00:33:36	F	40-50	Sales	2295.8
10	00:47:00	M	40-50	Food Supply Chain	4089.7
11	00:43:17	M	50-60	Insurance	104.1
12	00:50:17	F	50-60	HR	3528.7
13	00:30:42	F	30-40	Special Education	NA
14	00:40:19	F	50-60	Tea Entrepreneur	1247.9
15	00:43:14	F	30-40	Banking	3883.8
16	01:00:54	M	30-40	SOE Director	1687
17	00:22:14	M	50-60	Lawyer	4912.5
18	00:31:33	M	20-30	Student	NA
19	00:59:17	F	30-40	Branding	3993.6
20	00:49:06	M	30-40	Hotel Maintenance	3507.6
21	00:30:02	F	40-50	Insurance	2569.6
22	00:39:12	M	30-40	Restaurant	7518.8
23	00:54:00	F	30-40	Training	669
24	00:47:59	M	40-50	Supply Chain	7661
25	00:39:29	F	30-40	Wine Entrepreneur	163.2
26	00:45:50	M	30-40	Gym Entrepreneur	1530.6
27	00:48:39	F	40-50	Business Real-estate	6704.1
28	00:37:16	M	30-40	Hotel Management	5752.9
29	00:42:42	M	30-40	Children Education	797.2
30	00:27:38	M	30-40	Insurance	1764.4
31	00:43:57	M	50-60	After-sale Service	5859.7
32	00:33:06	M	20-30	Clothes Sale	38.3
33	00:47:53	F	20-30	New Media Writing	3245.6
34	00:31:02	M	30-40	Entrepreneur	3336.1
35	00:24:02	M	50-60	Higher Education	988
36	00:48:19	F	40-50	Project Management	4426.6
37	00:27:30	F	30-40	Hotel	187.5
38	00:35:23	F	20-30	Student	102.5
39	00:43:23	F	40-50	Insurance	1351.1

Figure 1 Reading Challenge and Special Virtual Badge

Figure 2 Special Challenge Badges

特别成就勋章

7日职场冲刺	7日「每天读…	知识股民
每天读点书–1…	年中听书王	每天听点书
每天多读点书	了不起的我	商界高手
7天每天读书…	连续21天读点书	听书·金融大亨
日更专栏连学…	更懂医学更健康	更懂心理更懂你

Figure 3 Virtual Certificate

Figure 4 Level-up badges

升级勋章

课代表Lv.1

终身学习Lv.7

鉴书官Lv.2

天天想上Lv.8

以得服人Lv.1

得到大神Lv.7

刻意练习Lv.6

得到赞助人Lv.1

心流Lv.7

笔记侠Lv.8

搜商Lv.5

人气王Lv.7

Figure 5 Learning Credit

Figure 6 Personal Profile

Figure 7 Personal Learning Data

5. Conclusion for the whole thesis

In this thesis, I conducted three studies in two platform settings implementing gamified incentive systems. As a novel way of incentive system design, gamification has been implemented in different kinds of organisations. Based on big data and algorithms, platform organisations are especially developing different gamified elements to motivate and change people's behaviour and thus influence organisations' performance. Whilst there is a growing literature that explores the potential for power and exploitation in platform and gamified settings, there is relatively little work that addresses the extant incentive literature seeking to understand its motivational effects. To fill this research gap, mobilising classic psychological theories, my thesis shed light on the vivid possibility and potential of gamification as incentive and control systems. On the one hand, this thesis raises the attention of management accounting scholars to further explore the effectiveness and efficiency of gamification in different settings. On the other hand, it also triggers considerations of gamification designers in organisations to rethink their current use of gamification and make sensible adjustments.

In the first chapter, I chose an on-demand delivery rider platform to investigate their gamified incentive system. When I did the first-round interviews with riders, I found out those gamified elements such as gamified delivery challenges with bonuses, rider leaderboards, and gamified rider levels do have a motivational effect on riders' performance. However, more importantly, when I talked with riders and analysed the first batch of transcripts, I felt that strong group or community feelings play an important role in retaining riders to keep working as a rider. Therefore, I turned my attention to the relations between riders and the riders' community to revise my interview protocols in my third-round interviews to discuss more riders' community and relationship topics with participants.

I mobilise the sense of community theory by McMillan and Chavis (1986) to explain how a sense of community complements the individual gamified incentive systems to further motivate and retain riders on this platform. The study introduced the sense of community as an incentive factor to management accounting literature to inspire future survey or experimental research to examine the effect of sense of community on workers' performance. Moreover, the study contributes to the emotion literature in accounting to explain how the gamified incentive system influences the riders' emotions on a digital platform organisation. This study also has practical

implications for platform organisations to rethink the use of gamification. Instead of isolating workers to chase their own goals using gamification designs (Anderson, 2021), the platforms could get motivational benefits from letting the workers generate a sense of community from the gamified incentive system.

During the interviews for the first chapter, I found that the riders' attitude to their leaderboards is quite interesting and different. This triggers my interest in why the same gamified leaderboards generate such different emotions and motivations for different riders. Digging into the development of social comparison theory (Festinger, 1954), I found that the individual difference in social comparison is one of the significant branches of this classic social psychology theory. I realised that social comparison orientation is an important personality characteristic which is overlooked by previous RPI literature in management accounting. As this personality influences people's behaviour to compare with others and process RPI, I expected it to have a certain moderation effect on riders' performance. Meanwhile, from field interviews, I found riders' RPI from leaderboards serves as a source for riders to conduct some vicarious learning to improve their capability of working. This inspires me to link this empirical evidence to the self-efficacy theory from Bandura (1997) to hypothesise the mediation effect of occupational self-efficacy between relative performance information measures to riders' performance.

Through a survey study, I find evidence of the moderation effect of social comparison orientation and the mediation effect of occupational self-efficacy. This study contributes to previous RPI literature to introduce the moderation effect of SCO. Future RPI research can further test its effect in other settings and consider this characteristic when designing surveys and experiments related to social comparison. In practice, when organisations use RPI to motivate their employees, they need to be aware of the characteristics of employees. This study also contributes to incentive literature to explain the mediation effect of occupational self-efficacy from RPI. By providing RPI to employees, organisations could increase their employees' self-efficacy in their jobs. Moreover, in the additional analysis, the study also brings an explanation of why workers don't access the RPI which is provided to them. Future research can continue to explore the reasons why employees don't check their RPI (Kohler et al., 2023; Eyring and Narayanan, 2018).

Finally, I chose an online learning platform setting for my third chapter. It is a coincidence that during the second year of my PhD, the platform introduced gamified elements in its mobile application. Although my participation in using this platform is mainly driven by the intrinsic motivation of learning new things and broadening my horizons in different subjects, the introduction of gamification did change my own user behaviour to chase for some gamified badges and higher learning credits. My user behaviour change triggered my research motivation to conduct a field study. I mobilised self-determination theory to look into both intrinsic and extrinsic motivations behind and internalisation of extrinsic motivations. I try to find out how these gamified designs satisfy people's three basic psychological needs (autonomy, competency, and relatedness).

From this study, I find that different users are interested in different gamified elements, which are internalised into their own value or part of life to a different extent. The designer of gamification should use different gamified elements to let people choose which they want to engage with. I also find the temporal effect of gamification (Kelly, Valtchanov and Webb, 2021). When people get bored with the gamified designs, their motivational effects only last for a short term. Therefore, the designers need to observe the change of interests of users to adjust the existing gamified designs to keep the motivational effect. Moreover, the study also finds that using a vague performance measurement system helps to prevent users from gaming the system to improve their learning credits.

As for the boundary conditions of findings in this thesis, given there are two platform settings for these three papers, these findings can shed light on the use of gamification for platforms powered by advanced technologies such as big data and artificial intelligence. More specially, the first two papers provide new insights into how the gamified mobile application influences the behaviours of the gig economy workers, who are under-researched by previous management accounting research. The third paper continues exploring gamification from the mobile application users' perspective (Chapman et al., 2021) to understand different kinds of motivation and satisfaction of psychological needs.

From a methodology perspective, although I interviewed only 27 riders, our results are triangulated by revising the short videos related to riders' daily life and work

filmed by thousands of riders around China. Therefore, it has a comprehensive external validity to reflect what is happening to riders every day. Moreover, based on my observation in the UK, the delivery riders also gather together during their spare time to chat with each other at different corners of the city. Therefore, our field research findings could be also applicable to Western countries' gig economy workers' settings. For the third paper, with the immersion approach of netnography, the findings reflect the holistic picture of users' behaviour on this learning platform based on my observation on this platform. The design of those gamified elements in the mobile application is also worthy of consideration by any platform companies around the world.

Moreover, from theoretical perspectives, these findings from three papers further extend the understanding of the important role of psychological theories for incentive and control literature, which is not only limited to platform settings but also applies to any organisation setting.

From this thesis, some potential research directions are emerging. Firstly, the online learning platform in the third chapter just announced a significant update on its gamified elements such as the check-in button, learning credit system and badges system in October. Future studies could take further investigation on this update of this gamified system.

Secondly, during the interviews for the third chapter, some participants warmly invited me to their in-person learning activities and reading groups. It would be interesting in the future to participate in some of these events to explore this strong connection between users on this platform. Finally, with a good understanding of users, future studies can take a platform perspective to discover how these gamified incentive designs help it retain users and improve its operating performance.

All these potential projects are closely related to what I did in the last four years to discover gamified incentives and control systems. With more and more organisations using gamification, this will be a fascinating research avenue which has much potential for management accounting researchers to explore.

Reference

Anderson, G. (2021) 'Amazon Gamifies Warehouse Worker Tasks, Raising Experts' Concerns'. Available at: https://www.forbes.com/sites/retailwire/2021/03/23/amazon-gamifies-warehouse-worker-tasks-raising-experts-concerns/?sh=54f909103810 (Accessed 08/09/2023).

Bandura, A. (1997) Self-Efficacy: The Exercise of Control. Worth Publishers.

Baxter, J., Carlsson-Wall, M., Chua, W. F. and Kraus, K. (2019) 'Accounting and passionate interests: the case of a Swedish football club', Accounting, Organizations and Society, 74, pp. 21-40.

Berger, L., Fiolleau, K. and MacTavish, C. (2019) 'I Know Something You Don't Know: The Effect of Relative Performance Information and Individual Performance Incentives on Knowledge Sharing', Journal of Management Accounting Research, 31(2), pp. 19-35.

Berger, L., Guo, L. and Presslee, A. (2023) 'Motivating Employees with Goal-Based Prosocial Rewards', Contemporary Accounting Research, 40(1), pp. 231-256.

Bialecki, M., O'Leary, S. and Smith, D. (2017) 'Judgement devices and the evaluation of singularities: The use of performance ratings and narrative information to guide film viewer choice', Management Accounting Research, 35, pp. 56-65.

Birnberg, J. G., Luft, J. and Shields, M. D. (2006) 'Psychology Theory in Management Accounting Research', in Chapman, C.S., Hopwood, A.G. and Shields, M.D. (eds.) Handbooks of Management Accounting Research: Elsevier, pp. 113-135.

Bonner, S. E., Hastie, R., Sprinkle, G. B. and Young, S. M. (2000) 'A Review of the Effects of Financial Incentives on Performance in Laboratory Tasks: Implications for Management Accounting.', Journal of Management Accounting Research. 2000, 12, pp. 19.

Bonner, S. E. and Sprinkle, G. B. (2002) 'The effects of monetary incentives on effort and task performance: theories, evidence, and a framework for research', Accounting Organizations and Society, 27(4-5), pp. 303-345.

Bouten, L. and Hoozée, S. (2022) 'Hybridity in a hotel chain: designing a package of controls to sustain a hybrid mission', Accounting, Auditing & Accountability

Journal, 35(8), pp. 1862-1891.

Burney, L. L. and Widener, S. K. (2013) 'Behavioral work outcomes of a strategic performance measurement system-based incentive plan', Behavioral Research in Accounting, 25(2), pp. 115-143.

Buunk, A. P. and Gibbons, F. X. (2007) 'Social comparison: The end of a theory and the emergence of a field', Organizational Behavior and Human Decision Processes, 102(1), pp. 3-21.

Cardador, M. T., Northcraft, G. B. and Whicker, J. (2017) 'A theory of work gamification: Something old, something new, something borrowed, something cool?', Human resource management review, 27(2), pp. 353-365.

Carroll, E. and Marginson, D. (2021) 'Relative performance information and social comparisons: Exploring managers' cognitive, emotional and dysfunctional behavioral processes', Management Accounting Research, 53.

Chapman, C., Chua, W. F. and Fiedler, T. (2021) 'Seduction as control: Gamification at Foursquare', Management Accounting Research, 53, pp. 100765.

Dale, S. (2014) 'Gamification: Making work fun, or making fun of work?', Business Information Review, 31(2), pp. 82-90.

Deci, E. L. and Ryan, R. M. (2000) 'The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior', Psychological Inquiry, 11(4), pp. 227-268.

Deterding, S., Dixon, D., Khaled, R. and Nacke, L. 'From game design elements to gamefulness: defining" gamification"'. 2011, 9-15.

Eyring, H., Ferguson, P. J. and Koppers, S. (2021) 'Less Information, More Comparison, and Better Performance: Evidence from a Field Experiment', Journal of Accounting Research, 59(2), pp. 657-711.

Festinger, L. (1954) 'A Theory of Social Comparison Processes', Human Relations, 7(2), pp. 117-140.

Frederickson, J. R. (1992) 'Relative performance information: The effects of common uncertainty and contract type on agent effort', Accounting Review, pp. 647-

669.

Gibbons, F. X. and Buunk, B. P. (1999) 'Individual differences in social comparison: development of a scale of social comparison orientation', Journal of personality and social psychology, 76(1), pp. 129.

Gist, M. E. and Mitchell, T. R. (1992) 'Self-efficacy: A theoretical analysis of its determinants and malleability', Academy of Management Review, 17(2), pp. 183-211.

Hall, M. (2008) 'The effect of comprehensive performance measurement systems on role clarity, psychological empowerment and managerial performance', Accounting, organizations and society, 33(2-3), pp. 141-163.

Hannan, R. L., Krishnan, R. and Newman, A. H. (2008) 'The effects of disseminating relative performance feedback in tournament and individual performance compensation plans', Accounting Review, 83(4), pp. 893-913.

Hannan, R. L., Towry, K. L. and Zhang, Y. (2013) 'Turning Up the Volume: An Experimental Investigation of the Role of Mutual Monitoring in Tournaments', Contemporary Accounting Research, 30(4), pp. 1401-1426.

Hartmann, F. and Schreck, P. (2016) 'Rankings, Performance, and Sabotage: The Moderating Effects of Target Setting', European Accounting Review, 27(2), pp. 363-382.

Jeacle, I. and Carter, C. (2011) 'In TripAdvisor we trust: Rankings, calculative regimes and abstract systems', Accounting, Organizations and Society, 36(4-5), pp. 293-309.

Kelly, K., Presslee, A. and Webb, R. A. (2017) 'The Effects of Tangible Rewards versus Cash Rewards in Consecutive Sales Tournaments: A Field Experiment', The Accounting Review, 92(6), pp. 165-185.

Kelly, K., Valtchanov, D. and Webb, A. (2021) 'Behavioral implications of using an online slot machine game to motivate employees: A cautionary tale', Accounting, Organizations and Society, 89.

Kohler, M., Mahlendorf, M. D., Seiter, M. and Vogelsang, T. 'Social Comparison on Multiple Tasks–Sacrificing Overall Performance for Local Excellence'. 2022.

Kornberger, M., Pflueger, D. and Mouritsen, J. (2017) 'Evaluative infrastructures: Accounting for platform organization.', Accounting, 60, pp. 79.

Kozinets, R. V. (2019) Netnography: The essential guide to qualitative social media research. Sage.

Krath, J., Schürmann, L. and Von Korflesch, H. F. O. (2021) 'Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning', Computers in Human Behavior, 125, pp. 106963.

Latham, G. P. and Locke, E. A. (1991) 'Self-regulation through goal setting', Organizational behavior and human decision processes, 50(2), pp. 212-247.

Libby, R., Bloomfield, R. and Nelson, M. W. (2002) 'Experimental research in financial accounting', Accounting, organizations and society, 27(8), pp. 775-810.

McDaid, E., Andon, P. and Free, C. (2023) 'Algorithmic management and the politics of demand: Control and resistance at Uber', Accounting, Organizations and Society, pp. 101465.

McDaid, E., Boedker, C. and Free, C. (2019) 'Close encounters and the illusion of accountability in the sharing economy', Accounting, Auditing & Accountability Journal, 32(5), pp. 1437-1466.

McMillan, D. W. and Chavis, D. M. (1986) 'Sense of community: A definition and theory', Journal of Community Psychology, 14(1), pp. 6-23.

Merchant, K. and Van der Stede, W. (2017) Management Control Systems 4th Edition : Performance Measurement, Evaluation and Incentives. Harlow, United Kingdom, UNITED KINGDOM: Pearson Education, Limited.

Otley, D. (1999) 'Performance management: a framework for management control systems research', Management accounting research, 10(4), pp. 363-382.

Pfister, J. A. and Lukka, K. (2019) 'Interrelation of controls for autonomous motivation: A field study of productivity gains through pressure-induced process innovation', The Accounting Review, 94(3), pp. 345-371.

Presslee, A., Vance, T. W. and Webb, R. A. (2013) 'The effects of reward type

on employee goal setting, goal commitment, and performance', The Accounting Review, 88(5), pp. 1805-1831.

Repenning, N., Löhlein, L. and Schäffer, U. (2021) 'Emotions in Accounting: A Review to Bridge the Paradigmatic Divide', European Accounting Review, 31(1), pp. 241-267.

Ryan, R. M., Connell, J. P. and Deci, E. L. (1985) 'A motivational analysis of self-determination and self-regulation in education', Research on motivation in education: The classroom milieu, 2, pp. 13-51.

Schwellnus, C., Geva, A., Pak, M. and Veiel, R. (2019) 'Gig economy platforms: Boon or Bane?'.

Smith, R. H. (2000) 'Assimilative and contrastive emotional reactions to upward and downward social comparisons', Handbook of social comparison: Springer, pp. 173-200.

Sun, P. (2019) 'Your order, their labor: An exploration of algorithms and laboring on food delivery platforms in China', Chinese Journal of Communication, 12(3), pp. 308-323.

Sun, P. and Chen, J. Y. (2021) 'Platform Labour and Contingent Agency in China', China Perspectives, 2021(1), pp. 19-27.

Tafkov, I. D. (2013) 'Private and Public Relative Performance Information under Different Compensation Contracts', Accounting Review, 88(1), pp. 327-350.

Thomas, T. F. and Thornock, T. A. (2017) 'Me vs. We: The Effect of Incomplete Team Member Feedback on Cooperation of Self-Regarding Individuals Tyler', SSRN Electron. J.

Van den Bussche, P. and Dambrin, C. (2020) 'Peer-to-peer evaluations as narcissistic devices: fabricating an entrepreneurial community', Accounting, Auditing & Accountability Journal.

Van den Bussche, P. and Morales, J. (2019) 'Performing (in) the community: accounting, biopower and the sharing economy', Handbook of the Sharing Economy: Edward Elgar Publishing.

Van der Hauwaert, E., Hoozée, S., Maussen, S. and Bruggeman, W. (2022) 'The impact of enabling performance measurement on managers' autonomous work motivation and performance', Management Accounting Research, 55, pp. 100780.

Vogelsang, T. (2022) 'Time is (Not) Money: Incentive Effects of Granting Leisure Time', European Accounting Review, pp. 1-25.

Wang, L. W. (2017) 'Recognizing the Best: The Productive and Counterproductive Effects of Relative Performance Recognition', Contemporary Accounting Research, 34(2), pp. 966-990.

Webb, R. A., Williamson, M. G. and Zhang, Y. (2013) 'Productivity-Target Difficulty, Target-Based Pay, and Outside-the-Box Thinking', The Accounting Review, 88(4), pp. 1433-1457.

Yang, Y. (2021) 'How China's delivery apps are putting riders at risk', Financial Times.

Yu, A. (2021) 'Accountability as mourning: Accounting for death in the time of COVID-19', Accounting, Organizations and Society, 90, pp. 101198.Baxter, J., Carlsson-Wall, M., Chua, W. F. and Kraus, K. (2019) 'Accounting and passionate interests: the case of a Swedish football club', Accounting, Organizations and Society, 74, pp. 21-40.

Appendix

A. Interview Protocol for Paper 1

- Engagements: Introduce me and the Project.

- Exploration Interview questions:

  ➢ **Building Rapport with participants:**

  1. Why do you want to be a rider initially?

  2. How is everything going about your delivery job recently?

  ➢ **Delivery job of riders:**

  3. Please describe today's delivery work to me in detail (from your morning preparation to the last shifts you just finished)

     - Do those assigned targets on the app or your self-targets can be easily achieved?

     - How does the leaderboard function influence your working behaviour and emotions?

     - During your rest time, what topics do you chat with nearby riders?

  ➢ **Relationship and Influence Among Riders:**

  4. Share an unforgettable or recent event between you and your nearby rider friends.

  5. How do you usually contact other riders? Where do you usually get other riders' information?

  6. Why do you want to film a short video of your job and life? (Rider Vlogger only)

  ➢ **Future Expectation:**

  7. What is your feeling about being a member of the delivery rider? Why?

  8. What things for this delivery job would you like to see improved in the future?

B. Questionnaire for Paper 2
Informed Consent Question:

1. Are you willing to participate in this research project to fill in this questionnaire?
    - o   Yes
    - o   No

Rider Information:

2. Categories of riders:
    - o   Full-time
    - o   Part-time

3. Your gender:
    - o   Male
    - o   Female

4. Are you currently working in your hometown city or not?
    - o   Yes
    - o   No

5. Your Age:
    - o   Below 21
    - o   21-30
    - o   31-40
    - o   41-50
    - o   Above 50

6. Tenure as a rider (How many months have you worked as a delivery rider? For example, if you have been a delivery rider for 1 year, you can fill in 12)
    _____ Month

7. What is your rider level in May? (For Part-time Rider)
    - o   Novice Rider
    - o   Shining Star 1
    - o   Shining Star 2
    - o   Shining Star 3
    - o   Glory Elite 1
    - o   Glory Elite 2
    - o   Glory Elite 3

- o Strongest Hero 1
- o Strongest Hero 2
- o Strongest Hero 3
- o Supreme God of War

8. What is your rider level in May? (For Full-time Rider)
   - o Bronze Knight
   - o Silver Knight
   - o Golden Knight
   - o Pioneer Knight
   - o Elite Knight
   - o Glory Knight

Functions of your mobile application:

9. Think about the information you are provided (e.g., Daily/Monthly rankings information in the riders' leaderboard). To what extent do you agree or disagree with the following statements (1 = strongly disagree to 5 = strongly agree):
   - o The provision of Leaderboard ranking information increases my productivity.
   - o The provision of Leaderboard information improves the quality of the work I do.
   - o The provision of Leaderboard ranking information gives me greater control over my work.
   - o Overall, I find the provision of this leaderboard ranking information useful in my job.

10. How frequently do you check your Daily Leaderboard Ranking?
    - o Every rest after the peak-time period
    - o Daily
    - o Weekly
    - o Monthly
    - o Less than once every month

About You:

11. To what extent do you think the statements below are true or not, give a score for each item below between 0 to 100 (0 for Never True, 100 for Always True)
    - I can remain calm when facing difficulties in my job because I can rely on my abilities.
    - When I am confronted with a problem in my job, I can usually find several solutions.
    - Whatever comes my way in my job, I can usually handle it.
    - My past experiences in my job have prepared me well for my occupational future.
    - I meet the goals I set for myself in my job.
    - I feel prepared for most of the demands in my job.

12. To what extent do you agree or disagree with the statements (1 = strongly disagree to 5 = strongly agree):
    - I often compare how my loved ones (boy or girlfriend, family members, etc.) are doing with how others are doing.
    - I always pay a lot of attention to how I do things compared with how others do things.
    - If I want to find out how well I have done something, I compare what I have done with how others have done.
    - I often compare how I am doing socially (e.g., social skills, popularity) with other people.
    - I am not the type of person who compares often with others (reversed).
    - I often compare myself with others with respect to what I have accomplished in life.
    - I often like to talk with others about mutual opinions and experiences.
    - I often try to find out what others think who face similar problems as I face.
    - I always like to know what others in a similar situation would do.
    - If I want to learn more about something, I try to find out what others think about it.

o   I never consider my situation in life relative to that of other people (reversed).

About Your Work:

13. How many orders did you deliver on the platform in May? _____

14. What was your cycling distance on the platform in May?   (in kilometres) _____

15. How many bad reviews did you have on the platform in May? _____

16. On average, how many orders do you deliver each month on the platform? _____

C. Interview Protocol for Paper 3

- Engagements: Introduction me and the Project.
- Exploration Interview questions:

  ➢ **Building Rapport with participants:**

1. Talk about yourself briefly. Where are you from in China? What is your current job?
2. How/where do you know about this mobile application initially? When do you start to use this mobile application?

  ➢ **Users' purchase and study behaviour:**

3. Please describe in detail how you use this mobile application to learn every day.
4. Why do you purchase certain online courses in this application? What are the factors to motivate you to purchase certain courses to study in this mobile application?

  ➢ **Gamification and Information Functions:**

5. In recent years, the application has started to create gamified badge systems and introduce many different challenge badges and virtual graduation certificates for you to get. How do you feel about these gamified badges and virtual certificates?
6. Do these gamified badges influence your course purchase decisions? Do you have the experience to buy certain courses or memberships just for getting certain badges you want to get?
7. Have you ever read about your weekly and monthly study reports? How does this information influence your study behaviour in a later week or month study?
8. What do you think about physical awards and the virtual coupons you get based on your study behaviours? Will that motivate you to purchase more courses to study or extend your membership?
9. What do you think about the function of "the town of knowledge"?

  ➢ **Future Expectation:**

10. What are your overall feelings about the update of this mobile app?
11. What things or functions would you like to see be improved in the future?