Rating-based compensation systems as a commitment tool on crowdworking platforms. An empirical analysis of four platforms[1]

Paul Hemsen
University of Paderborn, chair of personnel economics, Paderborn, Germany
paul.hemsen@uni-paderborn.de

Abstract Crowdworking (CW) or paid digital work on intermediary internet platforms is usually associated in public and academic discourses as a highly flexible work organization without long-term working relationships. This view misses the point that although CW platforms are working environments without an employment contract, platforms and their registered crowdworkers establish long-term working relationships often stretching over years. In addition, some platforms that deal with challenging tasks offer long-term compensation based on the individual rating levels of the crowdworkers and, therefore, show some extent of self-commitment. Nevertheless, research on commitment and rating-based compensation systems (RBCS) on CW platforms is rather rare. This paper examines how RBCSs on CW platforms motivate and commit crowdworkers on the platform; whether crowdworkers report higher affective and calculative commitment to a platform and perform better when the platform operates a RBCS; and whether the affective and calculative commitment of crowdworkers and their performance increase with their rating level on a CW platform. It is argued that a RBCS uses elements of internal labor markets and deferred compensation, both concepts developed for regular employment. In addition, goal-setting theory can explain how a platform hierarchy and its associated rewards set desirable goals for crowdworkers, and why these motivate and commit in the long run. It is therefore assumed that a CW platform with a RBCS motivates and commits crowdworkers much like a regular organization its employees. The hypotheses are tested with cross-sectional questionnaire data that includes 378 crowdworkers involved in text creation tasks from four CW platforms, two of which have implemented a RBCS, and the other two non-reputational fixed task prizes. The analyses show significant positive effects of a RBCS on the affective commitment and the weekly working hours of a crowdworker, while no significant effect on the calculative commitment was observed. Furthermore, each higher rating level is consistently associated with an increase in both commitment facets. To a certain extent,

---

[1] This work was funded by the Ministry of Culture and Science of the German State of North Rhine-Westphalia within the research program “Digital Future” [grant number 324-8.03.02.01.03-131116].

a higher rating level is also related to more weekly working hours of a crowdworker, which however stagnates at the second highest rating level and decreases at the highest possible rating level. Building on this, long-term relationships in CW agreements and underlying mechanisms that promote them should be given more attention in research. A deeper understanding would be particularly useful for CW platforms that struggle with commitment and incentive issues. It can also support the desirability of this accessible work environment for different groups of people, as it gives them the opportunity to actively improve their working conditions.

1 Introduction

From the public and academic discourses on paid work on intermediary Internet platforms (Standing & Standing, 2018), namely crowdworking (CW), one may quickly conclude that CW is just an unusually flexible way of working without ongoing employment relationships. Actors who make extensive use of this flexibility by switching platforms frequently or engaging in one-time exchanges is only one possible way. Another way is for crowdworkers to stay on one or a few other platforms for years and feel committed to them (Giard et al., 2019; Schulten & Schaefer, 2015). Nonetheless, those aspects of crowdworkers who are committed to platforms and platforms with commitment tools play a rather minor role in research (Buettner, 2015; Ghezzi, Gabelloni, Martini, & Natalicchio, 2018; Pedersen et al., 2013; Zheng, Li, & Hou, 2011) and therefore leave the one-sided view that CW is merely digital day laboring unchanged.

One important reason for CW platforms to promote the commitment of crowdworkers is that they rely on the skills of crowdworkers (Schulten & Schaefer, 2015). This applies especially for platforms that mediate demanding tasks such as graphic design, programming or sophisticated text creation. Advanced skills are obviously not so easy to replace compared to platforms for very simple and repetitive microtasks such as video and photo tagging, which require only basic technology knowledge. Only if a sufficient number of suitable crowdworkers are registered and willing to participate will a specialized platform be able to provide a consistent quality of service and stay in business (Boons, Stam, & Barkema, 2015). Like conventional employers, platforms must attract and retain skilled crowdworkers (Boons et al., 2015; Schulten & Schaefer, 2015). But unlike conventional employers, platforms are not willing to offer formal employment to their valuable crowdworkers because they then have to accept the corresponding legal obligations (Prassl & Risak, 2015).

In fact, there are several differences between platform providers and regular employers. In general, crowdworkers registered on a CW platform are free to choose the tasks they want to do from a task pool and are paid after they have delivered a satisfactory result. These satisfactory results in turn confirm his or her self-reported skills, as CW platforms are not able to carefully screen and select their experts during the registration process (Gadiraju, Fetahu, Kawase, Siehndel, & Dietze, 2017). But crowdworkers can also leave without notice and often work for a number of CW platforms. Because of this flexibility on both sides, CW platforms struggle with information asymmetry regarding the motives and skills of crowdworkers and lack important resources to promote the ongoing participation of qualified crowdworkers (Boons et al., 2015; Schulten & Schaefer, 2015). Therefore, an appropriate incentive design on CW platforms is important for both the motivation and commitment of suitable crowdworkers.

Rating-based compensation systems (RBCS), as the author calls them (Hemsen, 2021), are a promising way to address these commitment and incentive issues of CW platforms. This is achieved by allocating monetary and non-monetary incentives to crowdworkers based on their rating level on the platform (Hemsen, 2021). The rating level is based on a crowdworker's past performance, typically in terms of quantitative output such as the number of tasks solved and the behavior that is rated by the platform and clients. By achieving a higher rating level, the crowdworker will be given access to more demanding and better paid tasks. In addition to this monetary aspect, platforms with a rating system are likely to address the social need for self-presentation of crowdworkers, self-efficacy, desired social bonds and playfulness (Boons et al., 2015; Feng, Jonathan Ye, Yu, Yang, & Cui, 2018), since ratings and thus the reputation of crowdworkers are visible to others and includes gamified elements (Eickhoff, Harris, Vries, & Srinivasan, 2012; Feng et al., 2018; Kawajiri, Shimosaka, & Kahima, 2014). Nevertheless, research remains largely unclear whether RBCSs are actually effective in increasing crowdworkers' commitment to a CW platform and their performance on the platform. Understanding the logic of RBCSs is important for understanding the ongoing relationships between CW platforms and their registered crowdworkers. Especially since RBCSs are already used by a number of highly specialized CW platforms for demanding task types, such as the text creation platforms Content.de and Textbroker, the graphic design platforms 99Designs, Designenlassen as well as the software and web interface test platforms uTest by Applause and test.io.

This paper addresses three questions: How do RBCSs on CW platforms motivate crowdworkers and commit them to the platform? Do crowdworkers show higher affective and calculative commitment to platforms with a RBCS and do they perform better there? Does the affective and calculative commitment of crowdworkers to a platform and their performance increase with their rating level?

In order to explain how a RBCS motivates and commits crowdworkers, this paper transfers two concepts developed for regular employment to the CW context and also supports it with a theoretical foundation. In particular, it is argued that the hierarchy of RBCS is similar to the concept of internal labor markets and that the gradual improvements in compensation with moving up in the platform hierarchy resemble the concept of deferred compensation. How a platform hierarchy and its associated rewards set desirable goals for crowdworkers that meet their economic and social needs, and why these goals can motivate and commit also in the long run is argued by the goal-setting theory of Locke and Latham (2002).

In order to consider the different needs of crowdworkers and their impact, two facets of organizational commitment and weekly working hours are examined as the main outcome variables. By transferring the definition of organizational commitment from Meyer and Allen (1991) to the CW context, affective commitment is defined as the willingness of a crowdworker to stay with the platform based on emotional reasons. Calculative (originally termed “continuance”) commitment is defined by the degree to which a crowdworker needs to stay because he or she lacks more advantageous alternatives or it is too expensive to leave the platform. Seven hypotheses are derived on how the existence of a RBCS and its rating levels affect the two facets of commitment as well as the weekly working hours.

The hypotheses were tested with a cross-sectional questionnaire data set of 803 crowdworkers from four German-speaking CW platforms. Of these, 378 crowdworkers involved in text creation tasks were analyzed in detail to increase the comparability between these platforms. The use of ordinary least squares (OLS) regression analyses and inverse-probability-weighted regression adjustments (IPWRA) produces two main findings: First, the findings show significant positive effects of a RBCS on the affective commitment and weekly working hours of crowdworkers who perform text creation tasks. In contrast to the affective facet, no significant effect of a RBCS on the calculative commitment of crowdworkers to the platform were observed. A possible reason could be that some crowdworkers may feel committed to a CW platform on a calculative level because they lack other income-generating alternatives and are therefore even willing to accept unbalanced working conditions and compensation systems. Second, a rising rating level, however, on the text creation platform is related to a gradual increase in both commitment facets and, to a certain extent, to the weekly working hours of crowdworkers. Different from the commitment facets, the weekly working hours increases until they stagnate at the second highest rating level and significantly decreases at the highest possible rating level. There are indications that the reason for this progression could be insufficient incentives, especially since the monetary incentives for crowdworkers of the respective rating levels describe a similar progression as their weekly working hours. Overall, the pattern shows strong support for the idea that the RBCS draws on mechanisms which are known from regular employment. Therefore, a RBCS seems to be a promising tool to commit and incentivize crowdworkers on CW platforms.

2 Related work

There is some related work on the commitment of individuals to intermediary platforms and on incentives that govern their performance and behavior. However, previous work on commitment does not deal directly with CW and rarely considers the different facets of organizational commitment to a platform. In addition, the CW literature emphasizes the importance of incentives that direct crowdworkers' performance and behavior on platforms (e.g. Dalle, den Besten, Martínez, & Maraut, 2017; Goh, Pe-Than, & Lee, 2017; Hsieh & Kocielnik, 2016; Liu, Yang, Adamic, & Chen, 2014). The success of platforms depends on the continued participation of crowdworkers (Boons et al., 2015; Schulten & Schaefer, 2015). But the extant literature does not explain how and why incentives systematically create a commitment to a CW platform. Despite these gaps, this paper benefits from the important insights and findings of this related work.

In general, it seems that research on commitment between crowdworkers and CW platforms plays a rather minor role. Only Schulten and Schaefer (2015) address directly the affective or emotional commitment, which is one of the three facets of Meyer and Allen's (1991) concept of organizational commitment, and the loyalty to crowdsourcing platforms. They conclude that satisfaction with the crowdsourcing process and a sense of virtual community are particularly relevant factors that have a positive impact on affective commitment and loyalty of individuals to a platform (Schulten & Schaefer, 2015). However, it should be noted that the authors refer to crowdsourcing and not to CW. The main difference between these working arrangements is that people who participate in online activities on crowdsourcing platforms are not necessarily compensated financially (Schulte, Schlicher, & Maier, 2020). Since CW is the financially compensated variant of crowdsourcing, it remains to be shown to what extent the findings can be plausibly transferred. Besides this single study on affective commitment, there are other related studies that indirectly consider commitment by examining ongoing engagement or identification of crowdworkers with a platform. Corresponding studies show how engagement or identification with a CW platform could be promoted by perceived pride and respect (Boons et al., 2015), fairness (Franke, Keinz, & Klausberger, 2013), and by considering the cost and benefit considerations of crowdworkers (Ye & Kankanhalli, 2017). Although only one study actually deals with affective commitment, while the calculative and normative facets of organizational commitment were not directly considered, these studies at least provide evidence that crowdworkers are willing to work regularly on certain CW platforms. Nevertheless, it remains unclear how incentive systems can promote and perpetuate this tendency to work on one platform regularly and over a longer period of time.

Why the distinction between different facets of commitment is important is explained by a brief look at the literature on organizational commitment. By transferring the concept of organizational commitment from Meyer and Allen (1991) to the CW context, affective commitment could be defined as the willingness of a crowdworker to stay on the platform for emotional reasons; he or she does not have to stay, but wants to stay. Calculative (originally termed "continuance") commitment is defined by the degree to which a crowdworker needs to stay with the current platform because he or she lacks advantageous alternatives or the switching costs are too high for him or her to leave the platform. Lastly, normative commitment is defined as a crowdworker's urge to stay with the platform because of his or her personal values and norms. Following by Meyer and Allen (1991), Meyer and Herscovitch (2001) argued that these three facets of organizational commitment – affective, calculative, and normative commitment – combine to form a commitment profile. Furthermore, they proposed that behavior varies in predictable ways across potential profile groups (Meyer, Stanley, & Parfyonova, 2012). Accordingly, it is important to consider different facets of commitment to understand the underlying needs of crowdworkers.

Once the needs of particular crowdworker groups have been identified, compensation systems are an important tool for CW platforms to meet those needs (Acar, 2018; Leimeister, Huber, Bretschneider, & Krcmar, 2009; Pee, Koh, & Goh, 2018). In the CW context, these may be grouped into economic and social needs (Zhao & Zhu, 2014). According to Meyer and Allen's (1991) definition of organizational commitment, it can be assumed that fulfilled economic needs are more likely to promote the calculative commitment of crowdworkers to a platform and that fulfilled social needs positively influence the affective commitment. To what extent these needs promote the normative commitment of crowdworkers is not considered here. Unlike regular employment, normative pressure from family and culture to engage in work and feelings of guilt, e.g., through advance payments or costly investments in trainings (Wiener, 1982), are highly unusual for CW agreements. Other studies examining economic and social exchanges also excluded normative commitment and emphasized the importance of affective and calculative commitment (e.g. Gong, Law, Chang, & Xin, 2009; Shore, Tetrick, Lynch, & Barksdale, 2006; Whitener & Walz, 1993).

A RBCS in particular is a compensation system that allows a CW platform to cover both types of crowdworkers' needs. There are crowdworkers who participate in CW to earn a living, supplement other types of income (Archak, 2010; DiPalantino & Vojnovic, 2009; Horton & Chilton, 2010; Stewart, Lubensky, & Huerta, 2010), improve job prospects (Brabham, 2008, 2010), and signal their ability to potential clients or employers (Lakhani & Wolf, 2003). The communication of a crowdworker's standing on the platform by means of a rating level should also meet the social needs of crowdworkers. A higher standing can meet the needs for recognition (Brabham, 2008, 2010), glory (Archak, 2010), social identification (Lakhani & Wolf, 2003) or virtual communities (Brabham, 2010; Zhong, Wang, & Qiu, 2011) by belonging to a specific status group on the platform. In addition, crowdworkers can experience enjoyment and fun (Brabham, 2008, 2010; Stewart et al., 2010), individual skill development (Brabham, 2010; Zhong et al., 2011), curiosity and interest (Brabham, 2010), self-affirmation (Zhong et al., 2011) or enjoyable time passing (Paolacci, Chandler, & Ipeirotis, 2010) by striving for and achieving a higher rating level. Since meeting the desired needs of crowdworkers leads to satisfaction, satisfied crowdworkers are likely to become more committed to the respective platform. Hence, by recognizing the past performance and behavior of crowdworkers, RBCSs are also likely to encourage their commitment to the CW platform.

It still remains empirically largely unclear whether such RBCSs can effectively increase crowdworkers' motivation and commitment to a platform. Unfortunately, to the author's knowledge, there is no study yet that combines monetary and non-monetary rating systems to examine the effect on motivation and commitment to a CW platform. There are only a few studies that address the effect of non-monetary ratings on the performance or participation of crowdworkers or online users. Studies on the performance of crowdworkers emphasize the positive impact of ratings as a visualized form of their reputation on the platform (Basili & Rossi, 2020; Goes, Guo, & Lin, 2016; Peer, Vosgerau, & Acquisti, 2014) or as a signal for crowdworkers to adapt their behavior to the requirements of the platform and clients (Riedl & Seidel, 2018). Studies on the participation of crowdworkers show a positive impact through ratings as a virtual reward system with gamified elements that address the intrinsic motivation of crowdworkers (Feng et al., 2018; Goh et al., 2017) or as a form of direct performance evaluation by clients and peers, even during an ongoing work process (Jian, Yang, Ba, Lu, & Jiang, 2019). Schörpf, Flecker, Schönauer, and Hubert (2017) support the idea of RBCSs conceptually, with their more holistic view of the concept of rating systems on platforms. They point out that ratings on a CW platform are a form of a control feature and thus, part of the platform design. Therefore, rating systems have a particularly strong impact on the time use, income and creativity of crowdworkers and thus on their working and living conditions (Schörpf et al., 2017). In summary, these studies provide preliminary evidence of the positive effects of the non-monetary rating component of a RBCS on crowdworkers' performance and participation, but lack a direct theoretical basis and empirical evidence for RBCs as a whole.

3 Theory and hypotheses

To explain how RBCSs motivate and commit crowdworkers, the argumentation of this paper refers to supportive concepts and a theoretical basis developed for regular employment. Among other things, it is basically argued that the hierarchy of RBCSs is similar to internal labor markets. With these internal platform hierarchies and the associated incentives, platform providers show a willingness to commit valuable crowdworkers and make the working relationship more mutually beneficial. Therefore, crowdworkers are given the opportunity to improve their working conditions with each higher rating level. An important part of these improvements is a compensation that could exceed the crowdworker's productivity, similar to a deferred compensation in regular employment. This is a concept for motivating and committing valuable employees to the company by paying them an income that is below their productivity at the beginning of their career, but which exceeds their productivity with their seniority (Lazear, 1990).

Looking at these two concepts on a more abstract level, moving up in the hierarchy and thus improving one's own income and working conditions are basically the rewards for challenging goals. The counterpart to well specified and challenging goals in regular employment are rating levels on a CW platform. In addition to transferring these two established concepts to the CW context, the goal setting theory of Locke and Latham (2002) provides a theoretical mechanism explaining why the combination of a hierarchy, each rating level and the rewards of a RBCS can steadily increase the performance and commitment of crowdworkers over time. This argumentation therefore not only emphasizes similarities with the incentive design for regular employment, but also supports the positive effects of RBCSs on the motivation and commitment of crowdworkers.

As briefly mentioned, some of these positive effects are due to the similar functioning of RBCSs and internal labor markets. According to Doeringer and Piore (1985: 8–9), internal labor markets function as an "[…] administrative unit, […], within the pricing and allocation of labor is governed by a set of administrative rules and procedures". In these internal labor markets, the internal labor force has exclusive rights to jobs filled internally, continuity of employment, and even at the entry port, they are protected from direct competition by workers in the external labor market (Doeringer & Piore, 1985). RBCSs on CW platforms have similar characteristics, with performance and behavior thresholds as a set of administrative rules and rating levels as administrative procedures that determine, among other things, income and availability of tasks (Hemsen, 2021). This also gives crowdworkers similar exclusive rights as an internal labor force. Each higher rating level is also the next higher position in the platform's hierarchy, internally staffed with registered crowdworkers and accompanied by assigned compensation. As long as these crowdworkers do not violate the platforms' guidelines they can rely on the continuity of this relationship. In addition, crowdworkers with a higher rating level are even protected to a certain extent from further competitors from outside the platform, as new competitors have to start with a lower rating level.

These points not only reveal similarities with internal labor markets, but also show that CW platforms with a RBCS are willing to commit themselves to valuable crowdworkers by improving their working conditions for moving up the hierarchy. This also makes it more likely that the working relationships between these parties will become more mutually beneficial over time than on platforms with no room for improvement. A number of authors also support this assumption, arguing that employees perform better, demonstrate more citizenship behavior, and express higher level of commitment to an employer when they worked in relationships that tend to be equally beneficial (Atchison, 1991; Lawler III, 1986; Osterman, 1988). To a certain extent, this can also apply to crowdworkers on platforms with a RBCS.

Despite the similarities, there are also some differences between internal labor markets and RBCSs, but these differences also strengthen the effects which RBCSs exert on crowdworkers' motivation and commitment. For example, a higher rating is associated with more demanding tasks and higher quality requirements, but the basic task structures are quite similar to those for lower ratings (Hemsen, 2021). Therefore, it is unlikely that crowdworkers who move up to the next higher rating level will suddenly find themselves not sufficiently qualified for the tasks, a situation known as the Peter Principle (Lazear, 2004). Even if this were the case, these crowdworkers would be downgraded again, something that is uncommon in internal labor markets (Doeringer & Piore, 1985). Hence, RBCSs lead to groups with more homogenous qualification levels and more adequately compensated crowdworkers according to their qualification level. Moreover, in internal labor markets, compensation is strongly linked to an employee's position and less to his or her actual performance, as performance is difficult to measure for positions with non-quantifiable output (Lazear & Gibbs, 2014). While the performance of crowdworkers can be measured directly per task, compensation on CW platforms with a RBCS is linked to the actual performance level and behavior of the crowdworker (Hemsen, 2021). These prompt and comprehensible influences by one's own performance and behavior seem likely to contribute to the effects on the motivation and commitment of crowdworkers.

Less obvious to crowdworkers, however, is that the gradual increase in compensation for each higher rating level in this platform hierarchy corresponds to a deferred compensation. According to Lazear (1990: 275), deferred compensation describes a situation in which "a worker who remains with the firm for a significant amount of time receives as a "bonus" wage that exceed his productivity and, presumably, his alternative use of time". According to this concept, wages can be increased gradually to attract and retain employees until the value of their marginal product or productivity is reached. Even a wage that is higher than an employee's productivity can be advantageous as long as the savings from the initially low wages that were below the value of an employee's marginal product are not exhausted (Lazear, 1990). The compensation increases of CW platforms with a RBCS is similar. Once crowdworkers have registered on a CW platform with a RBCS, crowdworkers typically start with a low rating level. A crowdworker's initial rating level is not necessarily the lowest, as some CW platforms also conduct qualification tests at the beginning of the membership. Therefore, these platforms have the possibility to assign task prizes below the value of the marginal product to newly registered crowdworkers and to increase the task prizes for crowdworkers with a higher rating level and a longer membership duration. Ideally, this in turn can lead to task prizes or an earned income above the market average for motivated and committed crowdworkers.

Nevertheless, there are differences compared to deferred compensation in regular employment. Although a crowdworker may eventually be paid above his or her productivity, a certain minimum level of performance and behavior is required at each level. If these are not met, the crowdworker will be downgraded regardless of his or her membership duration and paid according to his or her lower rating level. Therefore, performance and behavior thresholds for each rating level can be used for regulatory purposes (Gadiraju et al., 2017). They create incentives through high compensation, but can keep the difference between productivity and task prizes as well as the proportion of above-average paid crowdworkers at an appropriate level. In this way, CW platforms can offer higher compensation to motivated, engaged crowdworkers and a competitive pricing to their clients.

The underlying neoclassical economic view of deferred compensation could also explain why CW platforms that specialize in demanding task types are more likely to offer task prizes that rise with experience. According to this view, organizations would hire new employees as long as the additional revenue from an additional employee is less than or equal to their pay (Lazear, 1990). Due to the strong competition between CW platforms and the need for general human capital, CW platforms for challenging task types have difficulties in replacing or recruiting qualified crowdworkers (Boons et al., 2015; Schulten & Schaefer, 2015). Since they cannot easily increase the number of qualified crowdworkers, they gradually increase task prizes to motivate and commit their already registered crowdworkers. On the other hand, platforms that offer a very easy entry and only require crowdworkers with basic technology skills, such as microtask platforms (Gadiraju et al., 2017), can rely on the available mass of crowdworkers and thus keep the task prizes for their crowdworkers low.

In addition to the earlier comparisons, the more abstract view of the goal setting theory of Locke and Latham (2002) provides a theoretical basis why a platform hierarchy and rewards induced by a RBCS can influence the motivation and commitment of crowdworkers. The use of the theory is based on the premise that rating levels and what they represent, e.g. the position in the platform hierarchy with associated income and working conditions, can serve as challenging and well-specified goals on CW platforms. Locke & Latham 2006 show in their goal-setting theory that specific, challenging goals lead to a higher level of task performance than simple goals or vague, abstract goals. They particularly emphasize the importance of the feeling of success. Applied to the CW context, by achieving a higher rating level, crowdworkers can experience that they are able to progress and overcome challenges that are important or meaningful. This progress is particularly relevant for crowdworkers because it has a major impact on their working conditions on the platform. Building on the performance improvements through goals, the theory also states that satisfaction with one's own performance and appreciation by the organization through rewards can have a correspondingly positive effect on job satisfaction and thus on commitment to the organization. This in turn can encourage crowdworkers to commit themselves to the next higher rating level. Therefore, the goal-setting theory describes a cycle that potentially promotes the performance and commitment of a crowdworker for each higher rating level.

According to the goal-setting theory (Locke & Latham, 2002), not only can performance and commitment gradually increase, but crowdworkers can probably also gradually improve other factors that further influence the relationship between a rating level and their performance. For example, a crowdworker's qualifications or human capital, whether general or specific, can progress through a combination of task-specific feedback and the associated holistic feedback on their performance and behavior based on their rating level (Riedl & Seidel, 2018). This in turn can have a positive effect on their perceived self-efficacy, i.e. the knowledge of the crowdworker's own abilities (Bandura, Freeman, & Lightsey, 1997), and thus on the assessment of task complexity.

The argumentation of this paper is based on the assumption that by systematically designing RBCSs, the platform actually takes into account the needs of its crowdworkers to create desired incentives and goals that motivate and commit crowdworkers. The different needs are particularly relevant as they were also the initial motives for the crowdworkers' decision to engage in CW (Zhao & Zhu, 2014; Zheng, Li, & Hou, 2011). Furthermore, different results can be expected depending on the needs addressed. As mentioned in the previous section, crowdworkers may participate in CW to earn a living or supplement other types of income. Satisfying such economic needs with compensation based on reputation and qualifications is more likely to encourage a calculative commitment to a CW platform because the crowdworker lacks more advantageous alternatives or it is too expensive to leave the platform once a certain rating level is reached. Crowdworkers can also have social needs that have also been mentioned above, such as the desire for recognition, social identification, virtual communities, joy and fun, the development of individual skills or self-affirmation. Individual rating levels and corresponding visualized gamified elements such as star ratings, also can address these needs (Chittilappilly, Chen, & Amer-Yahia, 2016; Goes et al., 2016; Schörpf et al., 2017), thus rather promote the affective commitment of a crowdworker to its platform. In contrast to the three-dimensional organizational commitment from Meyer and Allen (1991), the normative facet is not considered. This is because RBCSs appear less likely to satisfy needs related to crowdworkers' normative values and beliefs, especially when those needs arise from a sense of obligation or guilt based on family or cultural socialization (Wiener, 1982). Furthermore, other studies on economic and social exchange have also particularly emphasized the importance of affective and calculative commitment, thus excluding the normative facet (e.g. Gong et al., 2009; Shore et al., 2006; Whitener & Walz, 1993). Based on the focus of economic and social needs, it can be assumed that meeting them also tends to have a positive impact on crowdworker performance, especially since previous work in the CW context supports this (e.g. Acar, 2018; Frey, Lüthje, & Haag, 2011; Goh et al., 2017; Goncalves, Hosio, Rogstadius, Karapanos, & Kostakos, 2015; Ho, Slivkins, Suri, & Vaughan, 2015). Overall, a RBCS can certainly fulfill different needs of crowdworkers and thus represents a promising compensation system on CW platforms to address incentive and commitment issues.

In order to systematically analyze RBCSs and to be able to make comprehensive statements about their effectiveness for the first time, seven hypotheses were developed from the argumentation of this paper. The first three hypotheses focus in particular on the question whether a RBCS significantly improves the performance and commitment of crowdworkers compared to crowdworkers who are confronted with non-reputational compensation systems with equal task prizes. These three hypotheses are:

Hypothesis 1a: Crowdworkers perceive a higher affective commitment to a platform with a RBCS compared to a platform with a non-reputational fixed task prize system.

Hypothesis 1b: Crowdworkers perceive a higher calculative commitment to a platform with a RBCS compared to a platform with a non-reputational fixed task prize system.

Hypothesis 1c: Crowdworkers perform more hours per week on a CW platform with a RBCS compared to a platform with a non-reputational fixed task prize system.

The following two hypotheses are used to empirically analyze whether there are differences in commitment between differently rated crowdworkers within a CW platform with a RBCS. These two hypotheses are:

Hypothesis 2a: The affective commitment of crowdworkers to a CW platform with a RBCS increases with each higher rating level.

Hypothesis 2b: The calculative commitment of crowdworkers to a CW platform with a RBCS increases with each higher rating level.

In contrast to Hypothesis 2a and Hypothesis 2b, the next and last two hypotheses do not assume a steady increase in performance up to the highest possible rating level, but rather a stagnation or decline after a certain point. Even if the task prize and thus the potential income on the platform or the immaterial incentives of a crowdworker are increased over time, the mere increase does not necessarily lead to an increase in performance according to the neoclassical economic view (Lazear & Gibbs, 2014). An increase in performance can also turn into stagnation or a decrease in performance (Lazear & Gibbs, 2014). For example, if task prizes increase, the crowdworker might substitute their free time for higher income, as it would be too expensive not to work. However, the crowdworker has the possibility to turn to other activities above a certain income level and thus stagnate or reduce his or her performance on the platform. This point can be reached earlier for crowdworkers on the CW platform, as many people work as crowdworkers on the side and accordingly only want to improve other sources of income (e.g. Archak, 2010; DiPalantino & Vojnovic, 2009; Horton & Chilton, 2010). By setting further sufficiently high incentives, this point can possibly be postponed for a while, but is also associated with correspondingly higher costs for the platform. In the specific case of a RBCS, however, there is another natural end to these incentive increases, namely the highest rating level. Thus, until more desirable rating levels are implemented, it seems likely that performance will eventually stagnate or decline, either by maintaining a desired level of income, a desired rating level, or the lack of desirable incentives that compensate the crowdworker for the cost of the additional effort. Based on this reasoning, the last two hypotheses are stated.

Hypothesis 2c: The weekly working hours of crowdworkers on a CW platform with a RBCS increase with each higher rating level, provided that the platform offers at least one next-higher rating level as an incentive for further improvement.

Hypothesis 2d: The weekly working hours of crowdworkers on a CW platform with a RBCS stagnate or decrease at the latest when they are at the highest possible rating level.

4 Methods and data

4.1 Research context

The empirical analyses in this paper are based on 378 crowdworkers engaged in text creation tasks from four German-speaking CW platforms. These four CW platforms can be shortly described as follows: The text creation platform operates an online platform that is specialized in the mediation of text creation tasks of varying complexity on the topics of finance and insurance, medicine, industry, management consultancy and press releases. It was founded in 2010 and has more than 6,500 registered crowdworkers. Among the surveyed platforms, only the text creation platform has implemented a RBCS for each available task. The second platform deals with a broad spectrum of rather simple and repetitive tasks such as answering questionnaires, writing simple texts such as product descriptions or categorizing data, so-called microtasks. The platform was founded in 2005 and reports more than 225,000 registered crowdworkers. Concerning the compensation system, this microtask platforms uses a mixed approach consisting of a RBCS for text creation tasks and non-reputational, predefined task prizes for the remaining task types. The third platform is a mobile microtask platform that trades in locally available microtasks such as test shopping's, checking product placements in stores or writing store and product reviews. This platform reports about 300,000 registered crowdworkers in Europe and was launched in 2010. In contrast to the previous platforms, this platform only pays non-reputational, predefined prizes per task. The last platform is specialized in software and user interface test cases. These tasks include testing websites and software for usability, quality, security or functionality, and writing detailed descriptions of performed test cases. The test platform has about 300,000 crowdworkers worldwide, was founded in 2011 and also pays non-reputational, predefined prizes per task.

Although all four platforms are specialized in different types of tasks, their crowdworkers report that text creation is a recurring task type. This means, crowdworkers on both platforms for simple and repetitive microtasks have to write texts about products or services and describe experiences. Crowdworkers from the platform for software and web interface testing document test cases and write reports. While the crowdworkers on the remaining platform for text creation, obviously write texts of varying complexity on several topics. Another point that promotes the comparability of these four platforms is that their task processing, i.e., from task offer to successful task completion, follows very similar routines. The CW processes starts when clients outsource a task to the respective intermediary CW platform. This platform offers tasks to its registered crowd in a highly standardized way, and ideally the tasks are completed successfully. The solutions are evaluated by clients and in case of disagreement sometimes also by the platform according to predefined criteria. A satisfactory result will be compensated with a predefined task prize, whereby the type of compensation depends on the platform's compensation system. The use of all four platforms is without any contractual obligations and without a direct charge of crowdworkers. However, crowdworkers are indirectly priced, because the platform charges their clients and withdraws a share of the task prize for offering a working environment, mediating between crowdworkers and clients, and operating as a trustee.

Despite the similarities in the task types and working environments, a certain heterogeneity among these crowdworkers remains. Due to a simple registration process on all four CW platforms, crowdworkers with different personal circumstances participate in this type of work. In order to further improve the comparability of these crowdworkers, variables controlling personal circumstances such as employment status and additionally platform-related information such as length of membership were included in the analysis. Further, inverse-probability-weighted regression adjustments were applied. This method accounts for the contrasts in mean treatment-specific predicted outcomes, addressing the common problem, also in OLS regression analysis, that each subject is observed in only one of the possible outcomes (Imbens & Wooldridge, 2009; StataCorp. L.L.C., 2019).

Overall, the data set allows comparisons between crowdworkers of CW platforms with a RBCS and platforms with non-reputation fixed task prize systems as well as comparisons of differently rated crowdworkers within a platform with a RBCS. Thus, it is an interesting opportunity to study the effects of RBCSs on the performance of crowdworkers and their commitment to the CW platform.

4.2 Measures, data and steps in the analysis

The data of these 378 crowdworkers are a subset of a dataset with a total of 803 observations and were collected as part of the interdisciplinary research program "Digital Future" (Giard et al., 2019). This research program was funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia and is a cooperation between the University of Paderborn and the University of Bielefeld. The aim of the research program and the researchers from the fields of business administration, computer science, engineering, psychology, and sociology involved in it was to study the topics of CW as well as data security and data protection in digitalized work processes. For this reason, the interdisciplinary questionnaire with its 71 questions comprehensively addresses working conditions, including data security and privacy, as well as working conditions in CW (Giard et al., 2019). It was conducted anonymously as an online questionnaire, offered as a paid task on four German-speaking platforms, and closed on the platform once 200 crowdworkers had answered the questionnaire. Participating crowdworkers took an average of about 26 minutes (with a standard deviation of 21 minutes) to complete the survey. Depending on the platform, it took between 3.5 hours and 12 days to reach the required number of 200 complete questionnaires. For technical reasons, some platforms exceeded the maximum number of 200 completed questionnaires. Nevertheless, additional questionnaires were paid and included in the dataset. A total of 9 crowdworkers who did not pass the attention test were subsequently excluded from the analyses. Of the 803 surveyed crowdworkers in the dataset, 204 were on the texting platform, 195 on the microtask platform, 198 on the mobile microtask platform, and 206 on the testing platform.

The following three variables are considered as dependent variables: The depending variable *affective commitment* is measured using the German short inventory (G-OCQ) (Maier & Woschée, 2014) of the organizational commitment questionnaire (OCQ) by Mowday, Porter, and Steers (1982). The affective commitment index comprises nine items, each measured on a five-point scale. One item reads: “I am willing to put in a great deal of effort beyond that normally expected in order to help this platform be successful”. The wording has been changed only slightly by replacing the word "organization" with "platform".

The second commitment index *calculative commitment* is measured using the German version (COBB: “Commitment Organisation, Beruf und Beschäftigungsform”; translated: “commitment organization, profession and form of employment”) (Felfe, Six, Schmook, & Knorz, 2014) of the organizational commitment questionnaire by Meyer, Allen, and Smith (1993) instead of the German short inventory (G-OCQ) (Maier & Woschée, 2014). The wording of the COBB items on calculative commitment seems more appropriate and easier to understand in the context of CW than the G-QCQ. The index comprises four items with five-point scales, one of which reads: “It would be very hard for me to leave my platform right now, even if I wanted to”. Again, only the word “organization” had to be replaced by “platform”.

The third dependent variable *weekly working hours* is measured metrically, with crowdworkers indicating their average working hours per week on the platform on which they were surveyed. This variable was chosen in particular to plausibly compare the quantitative performance of crowdworkers from different crowdworkers. Similar quantitative measures such as the number of tasks worked on in a week vary too much with the type of tasks. In addition, since better paid tasks typically require more time, especially when rework is required, it is also assumed that more hours per week increases the likelihood that a crowdworker will succeed on the platform.

To examine how a RBCS affects these dependent variables, special attention is given to the following two independent variables: *RBCS* is a binary variable that indicates whether the platform has implemented a RBCS or a non-reputational fixed task prize system, at least for their text creation tasks. *Rating level* is a categorical variable that is composed of the reported rating levels of the crowdworkers of the text creation platform. This platform is the only surveyed platform with a platform-wide RBCS. This variable consists especially of the observed rating levels 4, 4+, 4++ and 5 stars. At first glance the categories seem strange, but the two lowest rating levels on the text creation platform, namely 2 and 3 stars, are rather symbolic and are therefore rarely awarded and not observed in the data.

To control possible influences on the relationship between dependent and independent variables by the heterogeneity of the crowdworkers and the circumstances on the platform, seven control variables are considered. These variables refer to the platform-related information task availability and work record as well as personal circumstances of crowdworkers. To visualize resulting differences in the effects of the independent variables on the dependent variables and to determine whether the effects of the independent variables persist despite the control variables, the results section shows the models with and without control variables.

The *task availability* is measured with an index that summarizes the perceived availability of tasks on the platform that offered the questionnaire. Three items, each on a five-point scale, are included in the index. Crowdworkers were asked whether there were always tasks to work on, whether suitable tasks could be found quickly and easily, and finally, whether they rarely found interesting and exciting tasks. This index was used to determine whether an insufficient availability of tasks could be the reason for low performance levels or low commitment to the platform.

Variables related to a crowdworker's work record are also important to explain possible influences on a crowdworker's performance and commitment. Therefore, the *membership duration* and *additionally used platforms* are measured and included in the analyses.

In order to measure the *membership duration* of crowdworkers, they were asked to indicate their membership duration in months. This metric measure is used to control possible learning effects that may explain differences in performance or the development of an emotional commitment to a platform over time, independent of a RBCS. It should also be noted that a long membership is not necessarily the same as an active and committed membership. This means that although crowdworkers have been registered on a platform for years, may have performed no or only a few tasks in total. Since there are no direct costs involved, there are only few reasons to actively terminate an account except for possible data protection concerns. However, this variable cannot be regarded as uncritical. It can also be argued that a RBCS also promotes a longer membership duration, as it is assumed that RBCSs also promote commitment to the platform. Despite this concern, the variable was nevertheless considered to highlight the effect of a RBCS that is not based on time spent on a platform. To be able to meet some of the concerns, the differences between models with and without control variables are discussed.

The variable *additionally used platforms* was considered because a high number of additional platforms potentially results in lower performance on average, especially if these crowdworkers are also regularly employed. Further, if crowdworkers are working simultaneously on numerous platforms with access to a large pool of tasks, it is more likely that the work will be viewed in isolation from a particular platform. Therefore, a higher commitment with CW in general or with specific types of tasks seems more likely than a commitment to a specific platform. Due to this correlation with performance and commitment, it may reduce the effects of RBCSs and its rating levels on the dependent variables. Nevertheless, this variable was taken into account in the models that include control variables because crowdworkers that are active on multiple platforms are common and need to be considered. Furthermore, the concerns that the impact of the RBCS and its rating levels are over- or underestimated outweigh the concerns about including this variable.

The remaining four variables are intended to control possible influences by the personal circumstances of the surveyed crowdworkers. Among others, the variable *full-time crowdworker* was taken into account. It is binary and contains information whether CW is their main source of income or not. Crowdworkers were also asked about their current *employment status*. In doing so, their employment status was measured according to the following seven categories: self-employed, employed, self-employed and employed, not employed, retired, in study or training, or on parental leave. Multiple selections were possible. These two variables are included in the analyses because it can be assumed that both can significantly influence the performance and commitment of a crowdworker to a platform. Especially with regard to the available time outside of a regular employment and whether CW is considered a sideline. Because the work environment on CW platforms does not provide much job security and encourages competition among crowdworkers, gender and age differences in performance and commitment to the platform may occur. Therefore, the binary variable *Men* includes the identification of crowdworkers as male or female, while *age* was measured metrically. The *squared age* was also considered to control for inverse U-shaped effects.

For descriptive statistics of the analyzed variables and indices and for information on their distribution, see Table 1 and Table 2

Table 1. Descriptive statistics and correlations of analyzed dependent, independent and control variables

	Variable	Obs.	Mean	s.d.	1^{C}	2^{C}	3^{C}	4^{D}	5^{D}	6^{D}	7^{D}	8^{D}	9^{C}	10^{C}	11^{D}	12^{C}	13^{D}	14^{D}	15^{D}	16^{D}	17^{D}	18^{D}	19^{D}	20^{D}
1	Affective commitment^{C}	378	3.10	1.13	1.00																			
2	Calculative commitment^{C}	378	3.40	0.86	0.46*	1.00																		
3	Weekly working hours^{C}	378	10.16	10.61	0.52*	0.23*	1.00																	
4	RBCS^{D}	378	0.75	0.43	0.34*	0.07	0.46*	1.00																
	Rating levels																							
5	4 Stars^{D}	43	0.24	0.43	-0.11	-0.09	-0.13	.	1.00															
6	4+ Stars^{D}	74	0.41	0.49	-0.03	-0.12	-0.01	.	.	1.00														
7	4++ Stars^{D}	41	0.23	0.42	0.02	0.04	0.07	.	.	.	1.00													
8	5 Stars^{D}	21	0.12	0.32	0.17*	0.25*	0.09	.	.	.	.	1.00												
9	Task availability^{C}	378	3.40	0.90	0.28*	0.32*	0.37*	0.25*	0.06	-0.15*	0.01	0.13	1.00											
10	Membership duration^{C}	314	44.14	31.61	0.35*	0.15*	0.36*	0.30*	-0.08	-0.23*	0.15	0.27*	0.31*	1.00										
11	Full time crowdworker^{D}	366	0.29	0.46	0.37*	0.11*	0.53*	0.36*	-0.05	0.001	-0.03	0.12	0.22*	0.26*	1.00									
12	Additionally used platforms^{C}	378	1.14	1.83	-0.09	-0.10	-0.17*	-0.25*	0.14	-0.00	-0.17*	0.04	-0.11*	-0.04	0.05	1.00								
	Employment status																							
13	Employed^{D}	131	0.35	0.48	-0.29*	-0.01	-0.33*	-0.33*	-0.01	0.14	-0.07	-0.12	-0.14*	-0.24*	-0.39*	-0.04	1.00							
14	Self-employed^{D}	147	0.39	0.49	0.31*	0.004	0.40*	0.45*	-0.13	-0.10	0.09	0.19*	0.15*	0.32*	0.39*	0.06	.	1.00						
15	Employed and self-employed^{D}	26	0.07	0.25	0.01	-0.01	-0.02	0.06	-0.08	0.07	0.03	-0.05	0.06	0.02	0.004	0.01	.	.	1.00					
16	Study or training^{D}	33	0.09	0.28	-0.13*	-0.10	-0.19*	-0.16*	0.03	0.07	-0.07	-0.05	-0.12*	-0.20*	-0.07*	-0.07	-0.37*	-0.20*	-0.10*	1.00				
17	Parental leave^{D}	14	0.04	0.19	-0.04	0.03	-0.04	0.06	0.16*	0.01	-0.11	-0.07	0.09	-0.003	0.15*	-0.04	-0.17*	-0.10*	-0.05	-0.07	1.00			
18	Not employed^{D}	16	0.04	0.20	0.11*	0.06	0.08	0.01	0.13	-0.07	-0.03	-0.02	0.01	-0.02	0.14*	-0.04	.	.	.	-0.09*	.	1.00		
19	Retirement^{D}	11	0.03	0.17	0.04	0.09	0.06	0.01	0.11	-0.09	0.05	-0.07	-0.05	0.06	-0.01	0.14*	-0.15*	-0.08*	-0.04	-0.06	.	.	1.00	
20	Men^{D}	372	0.46	0.50	-0.26*	-0.15*	-0.17*	-0.25*	-0.11	-0.01	0.18*	-0.07	-0.19*	-0.06	-0.19*	-0.04	0.10*	-0.08*	-0.06	0.13*	-0.13*	-0.07*	-0.05	1.00
21	Age^{C}	375	39.80	12.08	0.10	0.10	0.33*	0.31*	0.10	-0.18*	0.03	0.10	0.22*	0.41*	0.13*	-0.01	-0.12*	0.21*	-0.04	-0.37*	-0.01	0.11*	0.29*	-0.19*

Notes: Correlations marked with * are significant at ≤ 0.05; D marks dummy or binary variables; C marks continuous variables; correlations between two continuous variables are calculated by Pearson's correlation coefficients; correlations between continuous and binary variables are calculated by point-biserial correlations; correlations between binary variables are calculated by Cramer's V and a Chi-square test; calculations with Stata 15.

Table 2. Distributions of control variables and the dependent variable weekly working hours

Crowdworkers doing text creation tasks on the…	**Text creation platform**	**Microtask platform**	**Mobile microtask platform**	**Testing platform**
	(n = 204)	(n = 79)	(n = 34)	(n = 61)
Employed	36 (17.65%)	38 (48.10%)	22 (64.71%)	35 (57.38%)
Self employed	116 (56.86%)	20 (25.32%)	4 (11.76%)	7 (11.48%)
Employed and self-employed	18 (8.82%)	4 (5.06%)	2 (5.88%)	2 (3.28%)
Study or training	7 (3.43%)	8 (10.13%)	3 (8.82%)	15 (24.59%)
Parental leave	7 (3.43%)	6 (7.59%)	.	1 (1.64%)
Not employed	13 (6.37%)	3 (3.80%)	.	.
Retirement	7 (3.43%)	.	3 (8.82%)	1 (1.64%)
	(n = 197)	(n = 75)	(n = 34)	(n = 60)
Full-time crowdworker	87 (44.16%)	13 (17.33%)	2 (5.88%)	5 (8.33%)
	(n = 169)	(n = 68)	(n = 32)	(n = 46)
Membership duration in month: mean (std. dev.)	58.11 (30.87)	27.62 (27.47)	33.88 (21.90)	23.43 (18.29)
	(n = 204)	(n = 79)	(n = 34)	(n = 61)
Mean task availability (5-point scale)	3.67	3.18	3.22	2.90
Additionally used platforms: mean (std. dev.)	0.85 (1.13)	0.94 (1.66)	3.21 (3.11)	1.23 (2.25)
Weekly working hours: mean (std. dev.)	15.72 (11.22)	5.97 (5.18)	2.5 (2.71)	1.28 (0.84)
Sex: female/male	125/76	43/36	11/23	22/36
	(n = 204)	(n = 79)	(n = 34)	(n = 58)
Age: mean (std. dev.)	43.76 (11.78)	37.24 (10.85)	34 (8.31)	32.78 (11.24)

Initial supporting findings for the hypotheses can already be seen in these distributions. As Table 3 shows, the three largest groups of surveyed crowdworkers on platforms with a RBCS have at least a moderate level of affective and calculative commitment to the platform, while the biggest group of crowdworkers on platform with non-reputational fixed task prize systems show a low affective and a moderate calculative commitment. This provides first evidence in favor of Hypotheses 1a and 1b, which posit positive links between a RBCS and affective as well as calculative commitment. In addition, a similar situation is evident for crowdworkers on the text creation platform. 45.10% of their crowdworkers report a high level of affective commitment and at least a moderate level of calculative commitment to the platform. Whether an implemented RBCS also affects the weekly working hours of crowdworkers, i.e. Hypothesis 1c, is difficult to derive from the distribution data in this section. Although the crowdworkers of the text creation platform and the microtask platform report higher weekly working hours compared to the other two platforms (see Table 2), it cannot be deduced from this information whether a RBCS is the reason for these differences.

Similar to Table 3, Table 4 shows first support for the hypotheses, namely Hypotheses 2a, 2b and 2c. It can be seen that with each higher rating level on the text creation platform, both the commitment facets and the weekly working hours gradually increase. Contrary to the assumed drop in performance at the highest possible rating level, Table 4 also shows an increase in performance for 5-star crowdworkers. Therefore, these distributions do not offer any first indications in favor of Hypothesis 2d. Further analyses will show to what extent the indications derived from the distributions find further support or prove to be unfounded.

Table 3. Distribution of crowdworkers' perceived affective commitment (AC) and calculative commitment (CC) to their platform

Crowdworkers on platforms **with** a RBCS (n = 283)			
	AC low	AC moderate	AC high
CC low	14 (4.95%)	7 (2.47%)	5 (1.77%)
CC moderate	36 (12.72%)	54 (19.08%)	35 (12.37%)
CC high	9 (3.18%)	37 (13.07%)	72 (25.44%)
Crowdworkers on platforms **without** a RBCS (n = 95)			
	AC low	AC moderate	AC high
CC low	8 (8.42%)	3 (3.16%)	-
CC moderate	27 (28.42%)	19 (20.00%)	2 (2.11%)
CC high	8 (8.42%)	12 (12.63%)	8 (8.42%)
Crowdworkers on the text creation platform (n = 204)			
	AC low	AC moderate	AC high
CC low	6 (2.94%)	5 (2.45%)	3 (1.47%)
CC moderate	22 (10.78%)	34 (16.67%)	31 (15.20%)
CC high	5 (2.45%)	24 (11.76%)	61 (29.90%)

Note: AC or CC low include values below 2.34; AC or CC moderate includes values between 2.34 and 3.66; AC or CC high include values above 3.66.

Table 4. Distribution of the dependent variables in relation to the observed rating levels on the text creation platform

Rating levels on the text creation platform	**4 Stars**	**4+ Stars**	**4++ Stars**	**5 Stars**
Frequency	43 (24.02%)	74 (41.34%)	41 (22.91%)	21 (11.73%)
Affective commitment: mean (std. dev.)	3.32 (1.18)	3.48 (1.06)	3.55 (1.09)	4.01 (0.64)
Calculative commitment: mean (std. dev.)	3.34 (1.01)	3.35 (0.87)	3.53 (0.72)	4.07 (0.40)
Weekly working hours: mean (std. dev.)	13.12 (10.08)	15.55 (12.19)	17.07 (11.03)	18.38 (10.77)

The empirical analysis consists of two methods and different steps. First, OLS regression analyses were used for the formal tests of the seven hypotheses. The first three hypotheses were tested by comparing crowdworkers from four CW platforms performing text creation tasks. Of these four platforms, two CW platforms offer a RBCS with a star rating for text creation tasks, while the other two platforms offer a non-reputational fixed task prize system. The tests of the remaining four hypotheses focused exclusively on differently rated crowdworkers of the text creation platform, as this is the only platform examined with a RBCS for each task they offer. Taking into account the heterogeneity of crowdworkers with regard to personal and platform-related circumstances, the OLS regression analysis was performed with linear regression models consisting of independent and dependent variables as well as multiple linear regression models that additionally include several control variables. It can therefore be checked whether the effects observed in the linear regression models are retained in the multiple regression models and whether there are changes in the estimates. Due to the different measurement scales, the regression analyses were calculated with standardized z-Scores of the corresponding variables. In addition to the OLS regression analysis, inverse-probability-weighted regression adjustments (IPWRA) were performed to test the hypotheses. In contrast to OLS regression analyses, IPWRA estimators use a three-step approach to estimate treatment effects (Imbens & Wooldridge, 2009; StataCorp. L.L.C., 2019). First, they estimate the parameters of the treatment model (i.e., the model used to predict the treatment status) and compute inverse-probability weights. Second, using the estimated inverse-probability weights, they adjust weighted regression models of the outcome for each treatment level and obtain the treatment-specific predicted outcomes for each subject. Third, they compute the means of the treatment-specific predicted outcomes. The contrasts of these averages provide the estimates of the average treatment effects (ATE). Therefore, IPWRAs accounting for the problem that each subject is observed in only one of the potential outcomes (Imbens & Wooldridge, 2009; StataCorp. L.L.C., 2019). Due to its double robustness property, a resulting estimator will be consistent if either the treatment or outcome model is correctly specified (Imbens & Wooldridge, 2009). In summary, OLS regression analysis and IPWRAs in combination can reveal possible variations of the methods and take the heterogeneity of the crowdworkers into account when testing the hypotheses.

5 Results

5.1 Results of the OLS regression analyses

Applying both the ordinary least squares (OLS) regression analyses and the inverse-probability-weighted regression adjustments (IPWRA) leads to essentially the same conclusions about the seven hypotheses and thus to mutual support for the results. The results show that crowdworkers on a platform with a RBCS report significantly higher affective commitment to the platform and work significantly more hours per week. The results are consistent with Hypotheses 1a and 1c. Contradicting Hypothesis 1b, there are no significant differences between platforms with a RBCS and platforms with a non-reputational fixed task prize system in terms of the calculative commitment of these crowdworkers. A possible reason for this could be the low explanatory power of the corresponding models to explain the variance in the calculative commitment of crowdworkers to the platform. The same models, on the other hand, were successfully used for crowdworkers' affective commitment and weekly working hours.

Nevertheless, it can be seen that the affective and calculative commitment of the crowdworkers of the text creation platform increases with each higher rating level, supporting Hypotheses 2a and 2b. While the weekly working hours of these crowdworkers increase until it stagnates at the second highest rating level and decreases strongly at the highest rating level. Thus, the decrease at the highest rating level supports Hypothesis 2d, while Hypothesis 2c receive only partial support, as it rather assumes a steady increase in weekly working hours with each higher rating. What further strengthens the partial support for Hypothesis 2c is that there is strong evidence that insufficient monetary incentives on the respective platform might be the reason for this performance trajectory. This is because the relative increase in pay per word increases steadily up to a 4++ star rating, while it stagnates or decreases slightly at a 4++ rating, and decreases sharply at the highest possible rating level. Thus, the trajectory of relative increases in pay per word resembles the observed trajectory of crowdworkers' performance.

All results from the tests of the seven hypotheses with OLS regression analyses are shown in Table 5 and Table 6, while Table 7 shows the results of the IPWRAs. More detailed statements about the hypotheses and comparison of the results between these two methods are discussed below.

Table 5. Regression findings of affective and calculative commitment and weekly working hours regressed on a binary variable for a RBCS

Independent variables	**Dependent variables** H1a: Affective commitment^{Z}		H1b: Calculative commitment^{Z}		H1c: Weekly working hours^{Z}	
	(1)	(2)	(3)	(4)	(5)	(6)
RBCS^{D}	0.785*** (0.113)	0.396* (0.137)*	0.159 (0.115)	-0.100 (0.158)	1.321*** (0.131)	0.590*** (0.154)
Full time crowdworker^{D}	.	0.368** (0.125)	.	0.085 (0.145)	.	1.065*** (0.141)
Membership duration^{Z}	.	0.249*** (0.055)	.	0.113 (0.064)	.	0.125* (0.062)
Additionally used platforms^{Z}	.	-0.077 (0.053)	.	-0.131* (0.062)	.	-0.192** (0.060)
Employment status^{CA}						
Self-employed vs. employed	.	0.355** (0.135)	.	0.100 (0.156)	.	0.393** (0.152)
Employed and self-employed vs. employed	.	0.146 (0.199)	.	0.043 (0.230)	.	0.105 (0.224)
Study or training vs. employed	.	-0.221 (0.199)	.	-0.460* (0.231)	.	-0.179 (0.225)
Parental leave vs. employed	.	-0.201 (0.301)	.	0.058 (0.348)	.	-0.605 (0.338)
Not employed vs. employed	.	0.634* (0.309)	.	0.229 (0.358)	.	0.162 (0.348)
Retirement vs. employed	.	0.676* (0.334)	.	0.857* (0.386)	.	0.599 (0.376)
Men^{D}	.	-0.332** (0.102)	.	-0.236* (0.119)	.	-0.044 (0.115)
Age^{Z}	.	-0.239** (0.076)	.	-0.105 (0.088)	.	0.030 (0.085)
Squared age^{Z}	.	0.034 (0.055)	.	0.039 (0.064)	.	0.106 (0.062)
Constant	-0.189 (0.097)	-0.073 (0.139)	0.081 (0.100)	0.242 (0.161)	-0.504*** (0.113)	-0.547** (0.156)
Observations	378	301	378	301	378	301
F-value	48.55***	10.69***	1.89	2.00*	101.99***	19.95***
adj. R-squared	0.112	0.296	0.002	0.041	0.211	0.451

*Note: Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1; D marks dummy or binary variables; CA marks the categorical variable; Z marks the z-Score; reference category for the categorical variable employment status is employed; calculations with Stata 15.*

Table 5 particularly illustrates the effects of an implemented RBCS on the two facets of commitment and weekly working hours of crowdworkers from four CW platforms working on text creation tasks. Further, it shows both models with and without control variables. These were used to control for crowdworkers' heterogeneity in personal and platform-specific circumstances and to highlight potential differences in effect sizes of the independent variables. Thus, conclusions can be drawn regarding the first three hypotheses.

In support of Hypothesis 1a, model (2) in Table 5 shows that an implemented RBCS on a CW platform is associated with a significant increase of 0.396 standard deviations above the mean affective commitment compared to a platform with a non-reputational fixed task prize system. Although the effect size in model (2) with control variables is smaller than in model (1) without control variables, the effect is still considerable and statistical significance remains. Also, the quality of model (2) can be considered acceptable with an explained variance of 0.296. For this reason, Hypothesis 1a can be retained.

Regarding Hypothesis 1b, the regression results of model (3) show a rather small and statistically insignificant effect of RBCSs on the calculative commitment of a crowdworker. Model (4), which includes control variables, even shows a negative effect. A possible reason for this unexpected result could be the low explained variance of the two models. Therefore, it can be assumed that the models (3) and (4) are not sufficiently suitable to explain whether RBCSs can influence the calculative commitment of a crowdworker to a platform. Given the limited explanatory power of both models, Hypothesis 1b cannot be supported, but neither can it be completely rejected. Rather, the hypothesis and models require further analysis.

In contrast, the OLS regression results are consistent with Hypothesis 1c. As shown in the models (5) and (6), an implemented RBCS is associated with a statistically significant effect on the working hours per week on a platform. The effect size in model (6) is considerable, with an increase of 0.590 standard deviation above the mean weekly working hours, despite included control variables. Moreover, the variance explained in model (6) of 0.451 is quite high compared to the other models. Overall, Hypothesis 1c can be retained.

In addition to the hypothesis tests, the regression results show several other effects of the control variables on the dependent variables. For example, affective commitment and weekly working hours are positively associated with a crowdworker's employment status, especially if he or she works full-time on the platform or is self-employed instead of being employed. Similar patterns can be observed for crowdworkers with a long membership duration. Non-employed or retired people as well as people with increasing age only show a significantly higher affective commitment to the platform. In contrast, a negative effect size on affective commitment is observed for men, and weekly working hours decrease with each additional CW platform used. Therefore, these models and the results show some indications that performance and commitment on CW platforms are determined by factors similar or at least comparable to those in regular employment.

Table 6. Regression findings of affective and calculative commitment regressed on observed rating levels of a RBCS

	Dependent variables					
Independent variables	H2a: Affective commitment^{Z}		H2b: Calculative commitment^{Z}		H2c & H2d: Working hours per week^{Z}	
	(1)	(2)	(3)	(4)	(5)	(6)
Rating level^{CA}						
4+ Stars vs. 4 Stars	0.139 (0.181)	0.236 (0.182)	0.008 (0.182)	0.218 (0.188)	0.285 (0.253)	0.495* (0.240)
4++ Stars vs. 4 Stars	0.209 (0.206)	0.350 (0.211)	0.213 (0.207)	0.496* (0.218)	0.463 (0.289)	0.451 (0.277)
5 Stars vs. 4 Stars	0.616* (0.251)	0.595* (0.246)	0.827*** (0.253)	0.983*** (0.254)	0.616 (0.352)	0.301 (0.324)
Full time crowdworker^{D}	.	0.448** (0.151)	.	-0.095 (0.156)	.	1.125*** (0.198)
Task availability^{Z}	.	0.256** (0.081)	.	0.354*** (0.084)	.	0.350** (0.107)
Additionally used platforms^{Z}	.	0.017 (0.126)	.	0.072 (0.130)	.	-0.216 (0.166)
Employment status^{CA}						
Self-employed vs. employed	.	0.128 (0.202)	.	-0.182 (0.208)	.	0.265 (0.265)
Employed and self-employed vs. employed	.	-0.042 (0.289)	.	-0.326 (0.298)	.	-0.139 (0.380)
Study or training vs. employed	.	0.278 (0.536)	.	-0.598 (0.553)	.	0.017 (0.705)
Parental leave vs. employed	.	0.213 (0.405)	.	0.315 (0.418)	.	-0.401 (0.532)
Not employed vs. employed	.	0.596 (0.320)	.	0.052 (0.330)	.	0.187 (0.420)
Retirement vs. employed	.	0.970* (0.447)	.	0.524 (0.462)	.	1.172* (0.588)
Men^{D}	.	-0.262 (0.147)	.	-0.347* (0.152)	.	-0.015 (0.193)
Age^{Z}	.	-0.244 (0.131)	.	-0.318* (0.135)	.	0.076 (0.172)
Squared Age^{Z}	.	0.052 (0.082)	.	0.156 (0.084)	.	0.074 (0.108)
Constant	0.592*** (0.144)	0.205 (0.231)	0.133 (0.145)	0.068 (0.239)	0.831*** (0.202)	-0.257 (0.304)

Table 6. Continued

Independent variables	**Dependent variables** H2a: Affective commitment^{Z}		H2b: Calculative commitment^{Z}		H2c & H2d: Working hours per week^{Z}	
	(1)	(2)	(3)	(4)	(5)	(6)
Observations	179	169	179	169	179	169
F-value	2.07	3.69***	4.52**	3.47***	1.35	5.97***
adj. R-squared	0.018	0.194	0.056	0.181	0.006	0.307

*Note: Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1; D marks dummy or binary variables; CA marks categorical variables; Z marks the respective z-Score; reference category for the categorical variable rating level is 4 Stars; reference category for the categorical variable employment status is employed; calculations with Stata 15.*

Table 6 summarizes the regression results of the tests of the remaining four hypotheses. These allow statements to be made about the effects of increasing rating levels on the two commitment facets and the weekly working hours of the crowdworkers on the text creation platform. Unlike the models in Table 5, the models with control variables in Table 6 do not consider the membership duration of crowdworkers, since a high rating level and the associated time-intensive performance thresholds can be associated with a long membership duration. Therefore, membership duration may bias the effect of rating levels on the dependent variables. Instead, task availability was included because low task availability could specifically explain low calculative commitment and low performance regardless of a crowdworker's rating level. According to Table 6, this assumption proved to be correct, as higher task availability was associated with significantly higher affective and calculative commitment, as well as more working hours per week.

Other control variables that have been shown to be important determinants of commitment and performance on this platform are whether the crowdworker is a full-time crowdworker, a pensioner or a man. According to the models (2) and (6) in Table 6, full-time crowdworkers and pensioners working on the text creation platform are associated with significantly higher affective commitment and more hours worked per week. The corresponding effect sizes are not only considerably high, but also statistically significant. For example, full-time crowdworkers are associated with an increase of 1.125 standard deviations above the mean weekly working hours, while pensioners are associated with an increase of 1.172 standard deviations above the mean weekly working hours compared to crowdworkers who are simultaneously engaged in regular employment. In addition to the positive effects on the affective facet and weekly working hours, being a man on the text creation platform is associated with a decrease of 0.347 standard deviation below the mean calculative commitment of a crowdworker. Therefore, these observed effects provide additional information about potential groups that are more likely to commit themselves or invest more working hours than others.

Similar to the positive effects of the control variables, the regression results in the columns (2) to (4) of Table 6 show that increasing rating levels are associated with a continuous increase in affective and calculative commitment. Thus, the regression results support Hypotheses 2a and 2b. However only the effect sizes of a 5-star rating on both commitment facets and additionally of 4++ stars on calculative commitment are statistically significant, whereby the differences in the effect sizes for the individual rating levels are considerable. For example, the coefficients in model (2) related to affective commitment show a steady increase from 0.236 for 4+ stars to 0.595 for 5 stars compared to 4-star crowdworkers. It should be noted that all effect sizes relating to rating level are always compared to 4-star crowdworkers, as this is the reference category of the categorical variable *Rating level*. Similarly, the coefficients in model (4) related to the calculative commitment also increase steadily from 0.218 for a 4+ star rating to 0.983 for 5 stars when compared to 4-star crowdworkers. Regardless of this small interpretive hurdle, the effective sizes display a clear positive and steady trend. In summary, Hypotheses 2a and 2b are supported by models of adequate quality and considerable effect sizes and should not be rashly rejected due to a lack of statistical significance for each rating level.

The regression results in the models (5) and (6) in Table 6 show mixed results and are therefore not fully consistent with Hypothesis 2c, but support Hypothesis 2d. It can be seen that only the basic model (5) without control variables shows the expected steady increase in weekly working hours with each higher rating level, while model (6) that includes control variables shows only a steady increase up to a 4+ rating level, stagnation at a 4++ rating level and a decrease at the highest rating level. Based on the mixed results in the two models, it appears that the dependent variable weekly working hours is more sensitive to control variables than the two commitment facets. However, the size of the corresponding coefficients in both models is considerable, but only the model quality of 0.307 for model (6) is sufficiently high compared to 0.006 for model (5). Therefore, the statements for Hypotheses 2c and 2d are based on model (6). Accordingly, Hypothesis 2c is only partially supported because a stagnant or slightly reduced performance level was observed for the second highest rating level, i.e., 4++ stars. Whether this is actually due to insufficient incentives cannot be clearly stated from the observations in the dataset. By additionally considering the pay-per-word on the platform's website, it is noticeable that the trajectory of performance essentially mirrors the trajectory of the relative increases in pay-per-word for the respective rating levels. That is, the relative increase in pay-per-word is highest for crowdworkers rated 4+ stars, 60% more per word, compared to 4-star crowdworkers. Crowdworkers with 4++ stars receive only a 54% premium over a crowdworker with 4+ stars, while 5-star crowdworkers receive 27% more than a crowdworker with 4++ stars. Thus, the relative increases in pay per word stagnate for crowdworkers with 4 ++ stars and decline sharply for crowdworkers with 5 stars. Insufficient incentives, at least monetary ones, thus seem to be a plausible reason for the decline in weekly working hours. Since this does not completely rule out other reasons for the variation in performance, it still provides partial but strong support for Hypothesis 2c. In contrast, consistent with Hypothesis 2d is the stagnation of weekly working hours and the decline at the highest possible rating level. Again, the Hypothesis 2c and 2d are supported by a model of adequate quality and considerable effect sizes and should therefore not be rashly rejected due to a lack of statistical significance for each rating level.

Concluding on the results in Table 6, it can be seen that the RBCS in the dataset gradually increases both commitment facets of crowdworkers to the platform up to the highest possible rating of 5 stars, while the weekly working hours by crowdworkers already peak at a rating of 4+, the third highest rating on the platform. This could additionally indicate that the material and immaterial incentives in the RBCSs studied are sufficient in their current state to steadily increase crowdworkers' commitment to the platform during their "platform career," but not to gradually promote their performance. This opens up new research questions in particular that could specifically address the underlying mechanisms of commitment and performance on a platform to better understand and indeed create sufficient incentives.

Due to some mixed results from the models with and without control variables, and to see if these results can be replicated using another method, the following section presents and discusses the results of inverse-probability-weighted regression adjustments (IPWRA).

5.2 Results of additionally applied IPWRAs

With slight differences in effect sizes and statistical significance between the results of the IPWRAs and the OLS regression analyses, the statements on the hypotheses are essentially reflected in the IPWRA results. That is, according to the IPWRAS results, an implemented RBCS is still associated with higher affective commitment of crowdworkers to the platform and more hours worked per week compared to platforms with a non-reputational fixed task prize system, but still no significant effects on the calculative commitment were observed. Therefore, there is additional support for Hypotheses 1a and 1c, but not for Hypothesis 1b. Furthermore, the IPWRA results show that each higher rating level is associated with a steady increase in both commitment facets and weekly working hours. However, as with the OLS regression results, the weekly working hours stagnate at the second highest rating level and strongly decrease at the highest rating level, consistent with Hypothesis 2d. Therefore, the IPWRAs also support Hypotheses 2a, 2b, and 2d, while Hypothesis 2c is again only partially supported, as weekly working hours do not increase steadily with each higher rating level. An illustration of the results of the IPWRAs is shown in Table 7 and is compared in detail with the OLS regression results below.

Table 7. IPWRA findings of affective and calculative commitment, and working hours per week examined on the variables for a RBCS and its rating levels

		Dependent variables		
		Affective commitment (z-Score)	Calculative commitment (z-Score)	Working hours per week (z-Score)
		(1)	(2)	(3)
Average treatment effect RBCS (1 vs 0)	(1)	0.744***	-0.106	1.210***
		(0.205)	(0.194)	(0.085)
Potential outcome mean RBCS (0)		-0.202	0.283	-0.534***
		(0.196)	(0.183)	(0.030)
Observations		301	301	301
Average treatment effect 4+ Stars vs. 4 Stars	(2)	0.253	0.384	0.738**
		(0.162)	(0.203)	(0.226)
4++ Stars vs. 4 Stars		0.431*	0.489*	0.723**
		(0.189)	(0.213)	(0.245)
5 Stars vs. 4 Stars		0.624***	1.174***	0.589*
		(0.176)	(0.197)	(0.241)
Potential outcome mean 4 Stars		0.504***	-0.167	0.616***
		(0.133)	(0.180)	(0.155)
Observations		169	169	169
Estimator		Inverse-probability-weighted regression adjustment		
Outcome model		Linear		
Treatment model		Multinomial logit		
Outcome/Treatment model: control variables		Full time crowdworker^{D}, additionally used platforms^{Z}, men^{D}, age^{Z}, and age squared^{Z}. Row (1) additionally includes membership duration^{Z} Row (2) additionally includes task availability^{Z}		

*Note: Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1; D marks dummy or binary variables; Z marks calculated z-Scores; calculations with Stata 15.*

As can be seen, Table 7 summarizes the IPWRA results of testing all seven hypotheses. Row (1) of Table 7 shows the effects of an implemented RBCS on the two commitment facets and the weekly working hours of crowdworkers in the platform comparison, and row (2) shows the effects of the different rating levels on the dependent variables within the text creation platform.

With the exception of the control variable employment status, the IPWRAs models are composed of the same dependent, independent, and control variables as the respective OLS regression models. The categorical variable employment status was excluded for methodological reasons, as it results in very low propensity scores that cannot be used to run IPWRAs. Since the other control variables previously used were still included to create comparable groups of crowdworkers, the exclusion of employment status in the IPWRAs was considered acceptable. To highlight the core results of the hypothesis tests, the coefficients for each control variable are not included in Table 7. Conveniently, the interpretation of the coefficients is the same as for OLS regression. After comparing the IPWRA results in Table 7 with the OLS regression results in Table 5 and Table 6, it becomes clear that there are only minor differences. These differences can probably also be explained to some extent by the excluded control variable employment status. This shows not only that the exclusion was not that problematic, but also that the IPWRAs provide additional support for the hypotheses.

Comparing Table 7 column-by-column for simplicity, it appears that the two models in column (1) related to crowdworkers' affective commitment have very similar effect sizes and statistical significances for both an implemented RBCS and increasing rating levels compared to the OLS regression models in column (2) of Table 5 and Table 6. For example, in the OLS regression analyses, the effect sizes discussed previously on affective commitment by increasing rating levels are 0.236 for crowdworkers with 4+ stars, 0.350 for crowdworkers with 4++ stars, and 0.595 for crowdworkers with 5 stars, while IPWRAs estimate similar coefficients of 0.253, 0.431, and 0.624. Unlike the OLS regression results, the IPWRAs results for comparing crowdworkers with different rating levels show that statistical significance is not only evident for a 5-star rating, but also for a 4++ star rating. Despite these small improvements, an implemented RBCS on the CW platform can still be associated with a significant increase in affective commitment compared to a platform with a non-reputational fixed task prize system, supporting Hypothesis 1a. Furthermore, each higher rating level within a platform can be associated with a gradual increase in affective commitment to the platform, supporting Hypothesis 2a. Overall, thanks to the double-robustness characteristics of an IPWRA, the results provide additional robust support for Hypotheses 1a and 2a, assuming that either the outcome model or the treatment model of IPWRAs are correctly specified.

Similar results can be observed for the effects of a RBCS and increasing rating levels on the calculative commitment of crowdworkers. The results of the IPWRAs in column (2) in Table 7 also largely reflect the results of OLS regression models in column (4) in Table 5 and Table 6.

Thus, the results of the IPWRAs continue to show no significant or statistically significant differences in the calculative commitment of crowdworkers to platforms with a RBCS and platforms with non-reputational fixed task prize systems, which continues to reject Hypothesis 1b. However, the IPWRAs continue to support that each higher rating level is associated with a gradual increase in calculative commitment to the text creation platform, which is associated with Hypothesis 2b. For this reason, Hypothesis 1b remains rejected and Hypothesis 2b receives additional robust support under the assumption that at least one of the IPWRA models is correctly specified.

For the effects of an implemented RBCS and increasing rating levels on crowdworkers' weekly working hours, the IPWRAs results in column (3) in Table 7 actually show stronger support than the OLS regression models in column (6) in Table 5 and Table 6. This stronger support is based on both higher effect sizes and statistical significance for all independent variables in the models. For example, according to the OLS regression results, an implemented RBCS is associated with a statistically significant increase of 0.590 standard deviation above the mean weekly working hours compared to platforms with a non-reputational fixed task prize system, while the IPWRAs have a much stronger and also statistically significant effect size of 1.210. Further increases are observed for the estimated effect sizes of the individual rating levels, but unlike the results of the OLS regressions, the effect sizes of all rating levels are statistically significant at least at a 5% level. It is possible that this can be explained to some extent by the exclusion of the control variable employment status. However, despite the higher effect sizes and their improved statistical significance, the previous statements regarding Hypotheses 1c, 2c, and 2d remain essentially unchanged. Specifically, an implemented RBCS is associated with crowdworkers working significantly more hours per week, supporting Hypothesis 1c, while the weekly working hours of crowdworkers on the text creation platform increases with the rating level until it stagnates at the second highest rating level and significantly decreases at the highest rating level. Thus, there is only partial support for Hypothesis 2c, which assumes a steady increase for each higher rating level if sufficient incentives are in place, and additional support for Hypothesis 2d, which assumes a significant decrease at the highest rating level at the latest.

In summary, the OLS regression analyses and IPWRAs support most of the seven hypotheses, with the exception of Hypothesis 1b and Hypothesis 2c. However, only Hypothesis 1b is currently rejected due to the low explanatory power of the models, while Hypothesis 2c receives strong partial support. Nevertheless, the conclusions reached appear robust for all hypotheses.

In particular, there are slight differences in effect sizes and statistical significance between the linear and multiple linear regression models as well as between the results of the OLS regression analyses and the IPWRAs. Therefore, RBCSs on CW platforms seem to be a suitable tool to motivate and commit crowdworkers.

6 Discussion

6.1 Implications

This paper emphasizes the positive effects of RBCSs on crowdworkers' motivation and commitment to a platform by comparing crowdworkers from platforms with different compensation systems as well as differently rated crowdworkers within a particular platform. The intention was to examine this particular long-term compensation system on CW platforms and employment relationships in CW, as RBCSs are not the subject of current CW research despite its active use on platforms and the rather minor role of employment relationships in CW research. This focus may seem paradoxical in the highly flexible work setting of CW without employment contracts and frequently one-time exchanges (Schulte et al., 2020). This work contributes to the resolution of this apparent paradox by addressing both the lack of theoretically and empirically based knowledge on these topics.

The theoretical contribution of this work is characterized by reconstructing the logic of RBCSs through their similarities to internal labor markets and the forms of deferred compensation. Both concepts have been developed for regular employment relationships to motivate and retain employees and are now applied to the CW context. Locke and Latham's (2002) goal-setting theory provides the theoretical basis for this argument, explaining how an internal platform hierarchy and associated rewards for each "promotion" set goals for crowdworkers and why these goals can motivate and commit them in the long run. In total, seven hypotheses were derived that address the effectiveness of RBCS to motivate crowdworkers and commit them to the platform.

The quantitative empirical approach to test these hypotheses supports the argument that RBCSs are indeed a promising tool to motivate and commit crowdworkers. This is especially true with regard to crowdworkers' emotional commitment to the platform. Crowdworkers not only report significantly higher affective commitment to platforms with a RBCS, but also a steady increase in affective commitment with each higher rating level. The analysis on calculative commitment, on the other hand, do not show a significant difference in calculative commitment between platforms with a RBCS and platforms with a non-reputational fixed task prize system. A rather simple methodological reason for this could be the low explanatory power of the models used to explain the variance of a crowdworker's calculative commitment and that a different model is needed accordingly. It is also conceivable that some crowdworkers feel committed to a CW platform on a calculative level because they simply lack other income alternatives and are therefore even willing to accept unbalanced working conditions and compensation systems. However, similar to the affective commitment, the calculative commitment to a platform with a RBCS increases significantly up to the highest possible rating level on the platform. Thus, an implemented RBCS enables platforms to gradually and significantly increase crowdworkers' emotional and rational commitment to the platform, while the platform comparison only shows the effect on affective commitment. The effect on crowdworkers' invested weekly working hours on a platform with a RBCS is similarly promising, but the results show mixed effects. Crowdworkers on a platform with a RBCS, put in significantly more hours of work per week than crowdworkers on platforms without such a compensation system. However, comparing differently rated crowdworkers, the invested working hours increase steadily with each higher rating level, but stagnate at the second highest rating and strongly decrease at the highest rating. One reason for this could be insufficient incentives, especially monetary incentives, for steadily increasing the weekly working hours. Further research on the underlying mechanisms that explain commitment and performance on the CW platform with and without a RBCS is therefore needed to better understand and create more incentives that are perceived as sufficient by crowdworkers.

By expanding CW research with extant work on employment relationships and providing empirical evidence on the effects of RBCSs, this work contributes to CW research in several ways. First, it theoretically and empirically introduces the under-researched RBCS as a long-term form of compensation system to CW research. This not only extends the type of compensation systems studied in CW, but also differs from the short-term view in the literature. There are already some studies that focus on the impact of non-monetary rating systems on crowdworkers' performance and participation (Basili & Rossi, 2020; Feng et al., 2018; Goes et al., 2016; Goh et al., 2017; Jian et al., 2019; Peer et al., 2014; Riedl & Seidel, 2018; Schörpf et al., 2017). However, none of them studied the effects of a rating system that awards both material and immaterial rewards. Only Schörpf et al. (2017) conceptually point out that rating systems on a CW platform are a form of control feature and thus an essential part of the platform design. Moreover, they emphasize that ratings also have a particularly strong impact on crowdworkers' time use, income, and creativity, and thus on their working and living conditions. Second, this work emphasizes the application of Meyer and Allen's (1991) organizational commitment in CW. This is done by distinguishing affective and calculative commitment of crowdworkers as desired outcomes. Up to this point, the relevant CW literature has rarely considered the different facets of organizational commitment to a platform or the influence of a platform's compensation system on commitment in general. Only Schulten and Schaefer (2015) show that affective commitment to a platform can be fostered by satisfaction with the crowdsourcing process and digital communities. Note that this study focus on crowdsourcing and not on CW, therefore it does not necessarily include financially rewarded work on platforms. Further research is needed to examine the extent to which the findings apply to CW platforms as well. In contrast, the facets of calculative commitment and normative commitment have not been directly considered in the literature. Only Ye and Kankanhalli (2017) provide indications on the calculative commitment of crowdworkers to a platform. Their study on the impact of crowdworkers' cost and benefit considerations on their participation at least allows for the plausible assumption that advantageous situations on the platform may also have a positive impact on crowdworkers' calculative commitment. Third, this paper emphasizes the employment relationship between platforms and their registered crowdworkers. The consideration of CW as an employment relationship is particularly supported by the focus on RBCS, a long-term compensation system that has been implemented on CW platforms for years and serves to motivate and retain valuable crowdworkers. Therefore, unlike the usual one-off and sometimes anonymous exchanges between clients and crowdworkers (Brabham, 2008; Zheng et al., 2011), the frequent interactions between a platform with a RBCS and its crowdworkers over years can be interpreted as an employment relationship in a broader sense. While there is work that addresses this view and the extent to which platform providers act as employers (Bracha & Burke, 2016; Prassl & Risak, 2015; Stefano, 2016), these tend to focus on the legal aspects of these relationships. Overall, this paper usefully extends CW research by theoretically introducing RBCSs using existing research on employment relationships from personnel economics and psychological perspectives, and empirically underpinning it through a quantitative empirical approach. In doing so, it also emphasizes organizational commitment and employment relationships in CW.

Based on these contributions, implications for CW platforms can be derived. A more obvious implication is that CW platforms, especially for challenging tasks, with commitment and incentive issues might consider implementing a RBCS as a tool on their platform. Therefore, these platforms could offer more advantageous working conditions that can be actively improved by crowdworkers' through their reputation and qualifications (Hemsen, 2021). This in turn may have a strong impact not only on the working and possibly living conditions of crowdworkers through their income and time use, but also on their performance and behavior on the platform. Building on this, platforms that actively seek mutually beneficial exchanges between themselves and their crowdworkers could also face up to their responsibilities towards their registered crowdworkers. Despite their lack of status as regular employers, intermediary platforms are the central element of CW by setting the framework of rules for exchanges between all parties (Buettner, 2015; Giard et al., 2019; Zhao & Zhu, 2014). Crowdworkers or users, in particular, are the party with the least bargaining power and are therefore primarily subject to the requirements of internet platforms (Prassl & Risak, 2015). Therefore, it would be beneficial for a long-term relationship with crowdworkers if platform providers accepted this responsibility and acted accordingly.

This also has further implications for the compensation system of CW platforms. It is of importance to consider the economic and social needs of different groups of crowdworkers when designing a motivating and committing compensation system on a platform (Zheng et al., 2011). Since not all incentives are equally important to every crowdworker, some crowdworkers may be motivated by additional income, while others may simply want to pass the time productively (Brabham, 2010; Zheng et al., 2011). Therefore, in order to adapt a suitable compensation system for the platform and commit valuable crowdworkers, it is necessary to identify relevant target groups and their motives for engaging in CW. A more person-centric approach in incentive design on CW platforms therefore appears essential to foster better working conditions for crowdworkers and lasting valuable working relationships for CW platforms.

Based on the view of CW as an employment relationship, this also implies that researchers focusing on CW should also pay more attention to transferring the extant body of research on employment relationships and other related topics to the CW context. This would usefully enrich CW research, as this field tends to be exploratory and often lacks a theoretical foundation (Hemsen, Schulte, Schlicher, & Schneider, 2021; Zhao & Zhu, 2014). Moreover, it is likely that previous work on employment relationships from different disciplines shows that processes and information flows in CW are already known and well researched (Puranam, Alexy, & Reitzig, 2014). This, in turn, underlines that the novelty of CW can essentially be explained by the rearrangement of known processes and information flows compared to regular employments. This is not to downplay CW, but only to increase further understanding of this type of work.

6.2 Limitations and future research

Like other empirical studies, this work must face various limitations. The survey for this work was conducted as a paid task on four German-speaking CW platforms. The platforms informed their crowdworkers by mail and made the task available to every registered crowdworker. Each crowdworker could decide whether they wanted to participate or not. The survey was stopped on a platform once 200 crowdworkers had completed the questionnaire. This convenient way of collecting data on CW platforms has also been widely used by other researchers (Buettner, 2015; Ghezzi, Gabelloni, Martini, & Natalicchio, 2018; Pedersen et al., 2013; Zhao & Zhu, 2014). However, the approach implies that the total sample of initially 803 crowdworkers on four different CW platforms is not random. This raises questions about whether the observed effects are selection effects, incentive effects, or both. Moreover, it raises questions as to whether this work deals with the lack of representative data from the respective platforms, and how to assess the quality of the data.

Since both selection and incentive effects are part of the logic of RBCSs and RBCSs are studied by analyzing a cross-sectional data set, it is difficult to clearly separate these effects. Thus, it cannot be claimed that the results of this work actually show causality either. Nevertheless, this work shows that the observed effects of RBCS are not only driven by selection effects, but also by incentive effects. This is because the effects of RBCSs remain robust despite the inclusion of several person- and platform-specific control variables in the OLS regression models and the inclusion of the contrast of the mean treatment-specific predicted outcomes in the IPWRAs. Thus, the IPWRAs additionally account for the problem that each crowdworker is observed in only one of the potential outcomes. Overall, the robust results and patterns that are strongly consistent with the theory, including the flattening of the incentive effects at higher rating levels, suggest that the effects are not pure artifacts.

A dataset that does not fully represent the crowd of participating CW platforms and results that cannot be generalized are further limitations of this work, as only four CW platforms offered the questionnaire to their crowd and only a limited number of crowdworkers who were online in a given time frame were able to participate. Despite these limitations, this was not considered problematic due to the hypotheses formulated in this work. Specifically, the tested hypotheses aim to show that RBCSs and their individual rating levels can be associated with positive effects on crowdworkers' commitment and performance on these platforms, which is indeed observable. The extent to which these findings are transferable to other types of platforms or CW in general needs to be clarified through future research. Although this work lacks representative data for CW platforms and cannot draw causal conclusions, it nevertheless provides meaningful insights on RBCSs and CW as an employment relationship for future research on CW.

Derived from the lack of representativeness of the data, the heterogeneity of crowdworkers, on the other hand, posed a bigger problem. The heterogeneity and thus the comparability of crowdworkers was especially a problem when crowdworkers from different CW platforms with different specialization in task types were compared. For this reason, of the original 803 crowdworkers surveyed, only 378 crowdworkers engaged in text creation tasks were compared across platforms. These crowdworkers were selected because the text creation task was the only task type that occurred in significant numbers across all four platforms. Therefore, it was assumed that these crowdworkers had at least a minimum of comparable skills. In addition, the analyses considered several variables that control for crowdworkers' personal circumstances, such as employment status, whether CW is their full-time job, gender, age, and platform-related information such as length of membership, number of additionally used platforms, and task availability. The participation of only Germans in the survey may also reduce variation from different cultural backgrounds in an otherwise global market. Finally, IPWRAs were also applied. In contrast, comparability of crowdworkers within the text creation platform was less of an issue because crowdworkers worked on the same platform and on the same types of tasks. Despite better comparability, personal circumstances and platform-related factors were also considered and IPWRAs were applied. Overall, several factors were taken into account to control for the heterogeneity of crowdworkers and improve their comparability when analyzing the dataset.

The overall quality of this self-reported and cross-sectional questionnaire dataset was found to be good (Giard et al., 2019). The quality of the dataset was further improved by including an attention check for crowdworkers in the questionnaire in advance and a subsequent analysis of conspicuous response patterns, particularly for crowdworkers who take less than 10 minutes to complete the questionnaire with its 71 questions. Despite the thorough review of the data, only 9 crowdworkers were excluded from the dataset due to a failed attention check, which also speaks to the good quality of the dataset.

Due to these limitations, future research should focus on longitudinal datasets and randomized samples, as current research on CW tends to be exploratory and mostly based on cross-sectional data (Hemsen et al., 2021; Zhao & Zhu, 2014). In addition, this would potentially allow for more generalized conclusions about RBCSs and employment relationships in CW. Not part of this work, but of relevance, would also be research focusing on the ethical concerns arising from crowdworkers' reliance on the CW platform and the means they use, such as RBCS, as well as the additional pressures imposed by RBCS. The importance arises in particular from the fact that the number of people who find employment on a platform work there for years or even full-time. Accordingly, more research is needed in these areas to actively shape the path for an adequate digital work environment that is much more accessible to a large number of people than the regular labor market.

7 References

Acar, O. A. 2018. Harnessing the creative potential of consumers: money, participation, and creativity in idea crowdsourcing. ***Marketing Letters***, 29(2): 177–188.

Archak, N. 2010. Money, Glory and Cheap Talk: Analyzing Strategic Behavior of Contestants in Simultaneous Crowdsourcing Contests on TopCoder.com. ***Proceedings of the 19th International Conference on World Wide Web***, 1: 21–30.

Atchison, T. J. 1991. The employment relationship: un-tied or re-tied? ***Academy of Management Perspectives***, 5(4): 52–62.

Bandura, A., Freeman, W. H., & Lightsey, R. 1997. ***Self-efficacy: The exercise of control***. New York: Freeman.

Basili, M., & Rossi, M. A. 2020. Platform-mediated reputation systems in the sharing economy and incentives to provide service quality: The case of ridesharing services. ***Electronic Commerce Research and Applications***, 39: 100835.

Boons, M., Stam, D., & Barkema, H. G. 2015. Feelings of Pride and Respect as Drivers of Ongoing Member Activity on Crowdsourcing Platforms. ***Journal of Management Studies***, 52(6): 717–741.

Brabham, D. C. 2008. Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. ***Convergence***, 14(1): 75–90.

Brabham, D. C. 2010. Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. ***Information, Communication & Society***, 13(8): 1122–1145.

Bracha, A., & Burke, M. A. 2016. Who Counts as Employed? Informal Work, Employment Status, and Labor Market Slack. ***Working Papers Federal Reserve Bank of Boston***, No. 16-29: 1–42.

Buettner, R. 2015. A Systematic Literature Review of Crowdsourcing Research from a Human Resource Management Perspective. ***Proceedings of the 48th Hawaii International Conference on System Sciences***, 1: 4609–4618.

Chittilappilly, A. I., Chen, L., & Amer-Yahia, S. 2016. A Survey of General-Purpose Crowdsourcing Techniques. ***IEEE Transactions on Knowledge and Data Engineering***, 28(9): 2246–2266.

Dalle, J.-M., et al. 2017. Microwork platforms as enablers to new ecosystems and business models: The challenge of managing difficult tasks. ***International Journal of Technology Management***, 75(1-4): 55–72.

DiPalantino, D., & Vojnovic, M. 2009. Crowdsourcing and All-Pay Auctions, ***Proceedings of the 10th ACM conference on Electronic commerce***: 119–128.

Doeringer, P. B., & Piore, M. J. 1985. ***Internal Labor Markets and Manpower Analysis*** (1st ed.). Armonk, New York, USA: Me Sharpe.

Eickhoff, C., et al. 2012. Quality through Flow and Immersion: Gamifying Crowdsourced Relevance Assessments. ***Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval***: 871–880.

Felfe, J., et al. 2014. Commitment Organisation, Beruf und Beschäftigungsform (COBB). ***Zusammenstellung sozialwissenschaftlicher Items und Skalen***, https://doi.org/10.6102/zis9.

Feng, Y., et al. 2018. Gamification artifacts and crowdsourcing participation: Examining the mediating role of intrinsic motivations. ***Computers in Human Behavior***, 81: 124–136.

Franke, N., Keinz, P., & Klausberger, K. 2013. "Does This Sound Like a Fair Deal?": Antecedents and Consequences of Fairness Expectations in the Individual's Decision to Participate in Firm Innovation. ***Organization Science***, 24(5): 1495–1516.

Frey, K., Lüthje, C., & Haag, S. 2011. Whom Should Firms Attract to Open Innovation Platforms? The Role of Knowledge Diversity and Motivation. ***Long Range Planning***, 44(5-6): 397–420.

Gadiraju, U., et al. 2017. Using Worker Self-Assessments for Competence-based Pre-Selection in Crowdsourcing Microtasks. ***ACM Transactions on Computer-Human Interaction***, 24(4): 1–26.

Ghezzi, A., et al. 2018. Crowdsourcing: A Review and Suggestions for Future Research. ***International Journal of Management Reviews***, 20(2): 343–363.

Giard, N., et al. 2019. ***Interdisziplinäre Befragung von Crowdworkern***: ***Technical Report.*** Research Institute for Cognition and Robotics. Bielefeld.

Goes, P. B., Guo, C., & Lin, M. 2016. Do Incentive Hierarchies Induce User Effort? Evidence from an Online Knowledge Exchange. ***Information Systems Research***, 27(3): 497–516.

Goh, D. H.-L., Pe-Than, E. P. P., & Lee, C. S. 2017. Perceptions of virtual reward systems in crowdsourcing games. ***Computers in Human Behavior***, 70: 365–374.

Goncalves, J., et al. 2015. Motivating participation and improving quality of contribution in ubiquitous crowdsourcing. ***Computer Networks***, 90: 34–48.

Gong, Y., et al. 2009. Human Resources Management and Firm Performance: The Differential Role of Managerial Affective and Continuance Commitment. ***Journal of Applied Psychology***, 94(1): 263–275.

Hemsen, P., et al. 2021. Crowdsourcing for paid work as a new form of employment relationship: A content analysis. ***Working Paper***.

Hemsen, P. 2021. How do rating-based compensation systems on crowdworking platforms work? Providing long-term and reputational compensations for expert crowdworkers. ***Working Paper***.

Ho, C.-J., et al. 2015. Incentivizing High Quality Crowdwork. ***Proceedings of the 24th International Conference on World Wide Web***: 419–429.

Horton, J. J., & Chilton, L. B. 2010. The Labor Economics of Paid Crowdsourcing. ***Proceedings of the 11th ACM conference on Electronic commerce***: 209–218.

Hsieh, G., & Kocielnik, R. 2016. You Get Who You Pay for: The Impact of Incentives on Participation Bias. ***Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing***: 823–835.

Imbens, G. W., & Wooldridge, J. M. 2009. Recent developments in the econometrics of program evaluation. ***Journal of economic literature***, 47(1): 5–86.

Jian, L., et al. 2019. Managing the Crowds: The Effect of Prize Guarantees and In-Process Feedback on Participation in Crowdsourcing Contests. ***MIS Quarterly***, 43(1): 97–112.

Kawajiri, R., Shimosaka, M., & Kahima, H. 2014. Steered Crowdsensing: Incentive Design towards Quality-Oriented Place-Centric Crowdsensing. ***Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing***: 691–701.

Lakhani, K. R., & Wolf, R. G. 2003. Why Hackers Do What They Do: Understanding Motivation Effort in Free/Open Source Software Projects. ***MIT Sloan School of Management: Working Paper***, 4425(3): 1–27.

Lawler III, E. E. 1986. ***High-Involvement Management***: ***Participative Strategies for Improving Organizational Performance*** (1st ed.). 350 Sansome Street, San Francisco, CA 49104, USA: ERIC.

Lazear, E. P. 1990. Pensions and Deferred Benefits as Strategic Compensation. ***Industrial Relations: A Journal of Economy and Society***, 29(2): 263–280.

Lazear, E. P. 2004. The Peter Principle: A Theory of Decline. ***Journal of Political Economy***, 112(1): 141-163.

Lazear, E. P., & Gibbs, M. 2014. ***Personnel economics in practice:*** John Wiley & Sons.

Leimeister, J. M., et al. 2009. Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition. ***Journal of Management Information Systems***, 26(1): 197–224.

Liu, T. X., et al. 2014. Crowdsourcing with All-Pay Auctions: A Field Experiment on Taskcn. ***Management Science***, 60(8): 2020–2037.

Locke, E. A., & Latham, G. P. 2002. Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey. ***American Psychologist***, 57(9): 705–717.

Maier, G. W., & Woschée, R.-M. 2014. Deutsche Fassung des Organizational Commitment Questionnaire (OCQ-G). ***Zusammenstellung sozialwissenschaftlicher Items und Skalen***, doi: 10.6102/ZIS10.

Meyer, J. P., & Allen, N. J. 1991. A three-component conceptualization of organizational commitment. ***Human Resource Management Review***, 1(1): 61–89.

Meyer, J. P., Allen, N. J., & Smith, C. A. 1993. Commitment to Organizations and Occupations: Extension and Test of a Three-Component Conceptualization. ***Journal of Applied Psychology***, 78(4): 538–551.

Meyer, J. P., & Herscovitch, L. 2001. Commitment in the workplace: Toward a general model. ***Human Resource Management Review***, 11(3): 299–326.

Meyer, J. P., Stanley, L. J., & Parfyonova, N. M. 2012. Employee commitment in context: The nature and implication of commitment profiles. ***Journal of Vocational Behavior***, 80(1): 1–16.

Mowday, R. T., Porter, L. W., & Steers, R. M. 1982. ***Employee - Organization Linkages***: ***The Psychology of Commitment, Absenteeism and Turnover*** (1st ed.). New York: Academic Press.

Osterman, P. 1988. ***Employment futures***: ***Reorganization, dislocation, and public policy:*** Oxford University Press, USA.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. 2010. Running Experiments on Amazon Mechanical Turk. ***Judgment and Decision making***, 5(5): 411–419.

Pedersen, J., et al. 2013. Conceptual Foundations of Crowdsourcing: A Review of IS Research. ***IEEE 2013 – 46th Hawaii International Conference on System Sciences***: 579–588.

Pee, L. G., Koh, E., & Goh, M. 2018. Trait Motivations of Crowdsourcing and Task Choice: A Distal-Proximal Perspective. ***International Journal of Information Management***, 40: 28–41.

Peer, E., Vosgerau, J., & Acquisti, A. 2014. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. ***Behavior research methods***, 46(4): 1023–1031.

Prassl, J., & Risak, M. 2015. Uber, Taskrabbit, & Co: Platforms as Employers? Rethinking the Legal Analysis of Crowdwork. ***Comparative Labour Law & Policy Journal***, 37: 619–649.

Puranam, P., Alexy, O., & Reitzig, M. 2014. What's "new" about new forms of organizing? ***Academy of Management Review***, 39(2): 162–180.

Riedl, C., & Seidel, V. P. 2018. Learning from Mixed Signals in Online Innovation Communities. ***Organization Science***, 29(6): 1010–1032.

Schörpf, P., et al. 2017. Triangular love–hate: Management and control in creative crowdworking. ***New Technology, Work and Employment***, 32(1): 43–58.

Schulte, J., Schlicher, K. D., & Maier, G. W. 2020. Working everywhere and every time? Chances and risks in crowdworking and crowdsourcing work design. ***Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie (GIO)***, 51(1): 59–69.

Schulten, M. B., & Schaefer, F. 2015. Affective commitment and customer loyalty in crowdsourcing: antecedents, interdependencies, and practical implications. ***The International Review of Retail, Distribution and Consumer Research***, 25(5): 516–528.

Shore, L. M., et al. 2006. Social and Economic Exchange: Construct Development and Validation. ***Journal of Applied Social Psychology***, 36(4): 837–867.

Standing, S., & Standing, C. 2018. The ethical use of crowdsourcing. ***Business Ethics: A European Review***, 27(1): 72–80.

StataCorp. L.L.C. 2019. Stata: Release 16. Statistical Software. ***College Station, TX: StataCorp LLC.***

Stefano, V. de 2016. Introduction: Crowdsourcing, the Gig-Economy and the Law. ***Comparative Labour Law & Policy Journal***, 37(3): 1–10.

Stewart, O., Lubensky, D., & Huerta, J. M. 2010. Crowdsourcing Participation Inequality: A SCOUT Model for the Enterprise Domain. ***Proceedings of the ACM SIGKDD Workshop on Human Computation***: 30–33.

Whitener, E. M., & Walz, P. M. 1993. Exchange Theory Determinants of Affective and Continuance Commitment and Turnover. ***Journal of Vocational Behavior***, 42(3): 265–281.

Wiener, Y. 1982. Commitment in Organizations: A Normative View. ***Academy of Management Review***, 7(3): 418–428.

Ye, H. J., & Kankanhalli, A. 2017. Solvers' participation in crowdsourcing platforms: Examining the impacts of trust, and benefit and cost factors. ***Journal of Strategic Information Systems***, 26(2): 101–117.

Zhao, Y., & Zhu, Q. 2014. Evaluation on crowdsourcing research: Current status and future direction. ***Information Systems Frontiers***, 16(3): 417–434.

Zheng, H., Li, D., & Hou, W. 2011. Task Design, Motivation, and Participation in Crowdsourcing Contests. ***International Journal of Electronic Commerce***, 15(4): 57–88.

Zhong, Q. Y., Wang, Y. J., & Qiu, J. N. 2011. Empirical research on uses' continuous participation behavior in crowdsourcing community. ***Journal of Dalian University of Technology (Social Sciences)***, 32(1): 1–6.