

How do rating-based compensation systems on crowdworking platforms work? Providing long-term and reputational compensation for expert crowdworkers¹

Paul Hensen

University of Paderborn, chair of personnel economics, Paderborn, Germany
paul.hensen@uni-paderborn.de

Abstract Rating-based compensation systems (RBCS) are a promising form of long-term compensation on intermediary internet platforms for paid digital work, namely crowdworking (CW) platforms: Crowdworkers receive an individual rating, e.g. from one to five stars, based on past performance and subjective evaluation criteria, which also then grants access to more interesting and challenging tasks and better pay for a given amount of work. Although a number of CW platforms have implemented a RBCS for years, RBCSs that potentially promoting employment relationships between CW platforms and crowdworkers play a rather minor role in CW research. This paper therefore creates the term RBCS and introduces it into the CW literature by explaining its composition, providing an overview of its distribution on German-speaking CW platforms, identifying characteristics of platforms with such a system, and reconstructing the logic of RBCSs. Evidence about RBCSs comes from a comprehensive review of 32 German-speaking CW platforms with and without a RBCS. Basically, each RBCS consists of variations of three components, namely performance and behavior thresholds, status visualizations, and incentives. Despite possible variations of RBCSs, CW platforms with a similar specialization in a task type use similar combinations of these components. Especially CW platforms for sophisticated task types such as text creation and graphic design have implemented such systems. What also stands out about these platforms is that they have a very competitive working environment and rely on skilled crowdworkers with different levels of qualification to meet clients' requirements. Furthermore, this paper also reconstructs the logic of how a RBCS motivates and commits crowdworkers by drawing on extant work on regular employment from a primarily personnel economics perspective. It shows that the hierarchy of rating levels is similar to internal labor markets and that the incentives for moving up the hierarchy can be interpreted as a form of deferred compensation. The goal-setting theory

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].

provides the theoretical foundation that explains how an internal platform hierarchy, the aspiration to and achievement of its rating levels, and the associated incentives drive crowdworkers' motivation and commitment to the platform over the long term. Overall, this paper presents RBCS on platforms as a long-term compensation system contrary to the more short-term view in extant CW research.

1 Introduction

An increasing number of people with a certain level of expertise, so-called crowdworkers, regularly perform paid digital work on intermediary internet platforms (Kuek, Paradi-Guilford, Fayomi, Imaizumi, Ipeirotis, Pina, & Singh, 2015) or crowdworking (CW) platforms for short (Schulte, Schlicher, & Maier, 2020). In doing so, these crowdworkers typically accept lower pay than for comparable work in conventional employment (Brabham, 2012; Howe, 2006; Prassl & Risak, 2015), but also tend to compensate by seeking better-paying or more interesting tasks on multiple platforms (Giard et al., 2019; Hemsén, Schulte, Schlicher, & Schneider, 2021). Since attrition rates on CW platforms tend to be high (Schulten & Schaefer, 2015), this could be one of the reasons. Although CW platforms in particular, which mediate sophisticated task types such as design or text creation, rely on continuous participation of suitably skilled crowdworkers (Boons, Stam, & Barkema, 2015; Schulten & Schaefer, 2015), the predominant form of compensation on CW platforms tends to be short-term oriented. The platform offers a predefined task prize system without fully considering the reputation or qualifications of crowdworkers, which means that working conditions and pay rarely if ever improve over time (Fieseler, Bucher, & Hoffmann, 2019). This raises the question of compensation systems on CW platforms that better match the required qualification level while motivating and committing crowdworkers in the long run. This paper is therefore devoted to what the author calls a "rating-based compensation system" (RBCS), a long-term but under-researched compensation system that has been used on some CW platforms for years.

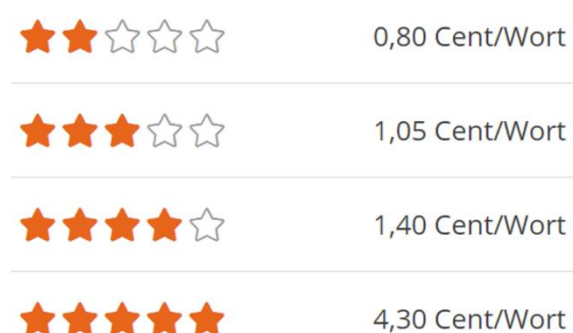
Besides the conceptual introduction of RBCS in CW research, a debate has also arisen about the definition of the more general term crowdsourcing (Estellés-Arolas & González-Ladrón-de-Guevara, 2012) and its financially compensated variant CW (Schulte et al., 2020), which is addressed in advance for better understanding. In both crowdsourcing and CW, a trilateral exchange takes place between an intermediary internet platform, a client, and suitable qualified persons performing online activities (Estellés-Arolas & González-Ladrón-de-Guevara, 2012; Howe, 2006; Schulte et al., 2020), who are referred to as crowdworkers in this paper for simplicity. It all starts with clients, who can be individuals, groups, or organizations that

outsource a digital task with a clearly defined goal to an intermediary platform (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). The platform displays these tasks online by a call to a specified crowd, usually their registered crowdworkers (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). Respective calls include the description of the task and information about the benefits for each actor involved (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). Since there are no employment contracts, crowdworkers take on these tasks voluntarily (Estellés-Arolas & González-Ladrón-de-Guevara, 2012) and therefore are not obliged to complete them. They provide their resources, such as time, money, effort, or expertise, and receive a whole range of benefits in return, such as intrinsic enjoyment of doing the task or financial compensation (Estellés-Arolas & González-Ladrón-de-Guevara, 2012; Schulte et al., 2020). The use of such platforms is not directly related to costs for crowdworkers, as the platform keeps a share of the task prize paid by clients for providing the work environment and acting as a trustee. The main difference between crowdsourcing and CW is that crowdworkers on CW platforms are always financially compensated for a satisfactory solution, which is not always the case with crowdsourcing (Schulte et al., 2020). Therefore, CW can be considered as gainful work.

Many CW platforms use non-reputational, predefined task prize systems to financially compensate their crowdworkers. Not only are the associated task prizes typically lower than payments for equivalent tasks in regular employment, but crowdworkers rarely have the opportunity to significantly improve their compensation or overall working conditions (Berg, 2016). That is the case unless they familiarize themselves with platform and client practices to speed up the workflow, select better-paying tasks when possible, and specialize in certain types of tasks. Improving crowdworkers' working conditions through gained reputation or validated qualifications are usually not part of these short-term compensation systems (Chittilappilly, Chen, & Amer-Yahia, 2016). Accordingly, CW research focusing on short-term incentive design predominates (Hemsen et al., 2021).

However, there are also CW platforms that use a RBCS as a long-term form of compensation system, establishing a committed relationship with valuable crowdworkers. Unlike regular task prize systems, platforms with a RBCS rely on crowdworkers' status on the platform to govern their activities, as they cannot fully rely on crowdworkers' self-reported information about their qualities (Gadiraju, Fetahu, Kawase, Siehndel, & Dietze, 2017). As Figure 1 shows, registered crowdworkers receive an individual rating, ranging for instance from one to five stars. The rating is based on the crowdworker's past performance, typically in terms of quantitative outputs such as task completions and qualitative outputs such as the platform's and clients'

evaluations. Typically, with each higher rating level, crowdworkers gain access to an expanded task pool, are paid better for a given amount of work, and receive other privileges. Since a crowdworker's rating is visualized through stars, badges, or titles, it is likely that these gamified elements also accommodate needs for self-presentation, self-efficacy, social bonds and playfulness (Feng, Jonathan Ye, Yu, Yang, & Cui, 2018; Goes, Guo, & Lin, 2016; Ryan & Deci, 2000). Moreover, it can also motivate crowdworkers who want to ascend the platform-exclusive "incentive hierarchy" (Goes et al., 2016) and generate long-term commitment to the platform, as crowdworkers may feel respected and valued (Boons et al., 2015; Schulten & Schaefer, 2015).



Screenshot: <https://www.textbroker.de/autoren-verguetung-konditionen> (15.01.2021)

Note: This example of a RBCS from the German text creation platform Textbroker.de shows a frequently represented type of status visualization, namely a star rating. The core element of this RBCS is that the payment per word increases with each additional star.

Figure 1. Example of a rating-based compensation system on a German-speaking text creation platform

It is not surprising to see the emergence of RBCSs on CW platforms. Despite highly flexible working conditions, platforms for challenging task types rely on ongoing relationships with experienced crowdworkers without offering an employment contract (Boons et al., 2015; Prassl & Risak, 2015; Schulten & Schaefer, 2015). It can therefore be assumed that these highly specialized platforms have to deal with incentive and commitment issues. Especially since designers, programmers, or writers are more difficult to replace than crowdworkers, who perform simple and repetitive tasks such as tagging photos and thus requiring only basic technical knowledge. In the absence of employment contracts, appropriate incentives are particularly important to attract, motivate, and retain suitable crowdworkers (Boons et al., 2015; Schulten & Schaefer, 2015). Some CW platforms train crowdworkers, but the constant risk of attrition seems to lead to trainings that mostly convey basic knowledge about a platform's

procedures and common task types. Therefore, it is more promising to directly address experienced crowdworkers and improve their working conditions to increase the attractiveness of CW (Schulten & Schaefer, 2015). This paper argues that a RBCS is a compensation system to achieve these goals.

As the RBCS is not, to the author's knowledge, the subject of current CW research, this paper addresses three previously unanswered questions: How are RBCSs composed in terms of used performance and behavior evaluation criteria or thresholds, visualizations of status and designated incentives? What are characteristics of CW platforms with a RBCS? How may a RBCS matters for crowdworker's performance on the platform and their commitment to the platform?

The questions are addressed by examining 32 German-speaking CW platforms using an exploratory approach and by applying extant work on employment relationships to CW. This conceptual paper yields three main findings: First, despite possible numerous variations in performance and behavior evaluation criteria for crowdworkers, status visualization, and incentives, it could be observed that similar components and combinations of them are used depending on the specialization of a platform. These variations seem to be necessary to meet the different requirements of specific task types. Second, it is noticeable among CW platforms with a RBCS that they have a highly competitive work environment and rely on crowdworkers with different levels of qualification, as they apply qualification-based selection. Further, specialized platforms limit access to tasks determined by a crowdworker's reputation and qualifications. This is especially necessary because CW platforms cannot fully rely on crowdworkers' self-reporting of their reputation and qualifications (Gadiraju et al., 2017). Third, the logic of RBCSs to motivate and commit crowdworkers to a CW platform can be reconstructed by drawing on extant work on employment relationships from a mainly personnel economics perspective. According to this work, the internal rating hierarchies on the CW platform have several similarities to internal labor markets in regular employment relationships, and the gradual improvement of pay and working conditions over time resembles forms of deferred compensation. On a more abstract level, Locke and Latham's (2002) goal-setting theory additionally provides the theoretical foundation that can explain why such hierarchies, striving for and achieving higher rating levels, and incentives can increase crowdworkers' motivation and commitment to the platform in the long run.

Overall, the long-term approach of RBCSs and the emphasis on employment relationships in CW adds a useful new perspective to the primarily short-term view in CW research.

Understanding RBCS is important because they are already part of CW platforms and enable a more mutually beneficial relationship between platforms and crowdworkers by empowering crowdworkers to actively improve their working conditions on the platform.

2 Rating-based compensation systems

As there is no detailed information about RBCSs on CW platforms in the literature, more information is needed for a better understanding. Therefore, the basic composition of a RBCS, the distribution of such a system and characteristics of platforms that have been implemented are explained below. First evidence comes from reviewing 32 German-speaking CW platforms of different task types.

2.1 Developing the overview

Three platform selection criteria were applied in the literature search and web search for suitable CW platforms. First, the CW platform is an intermediary that mediates via an open call and can intervene digital work between their registered crowdworkers and their clients (Schulte et al., 2020). This criterion ensures that no crowdworker has an employment contract with an involved party as well as that the platform can intervene in the working conditions. Second, the platform compensates monetarily satisfactory results that ensures only platforms for digital gainful work. Third, the platform has a German-speaking web interface to improve the comparability of these platforms, since they address crowdworkers with similar expectations regarding working conditions and pay. The German CW market was additionally of particular interest as this research was funded by the Ministry of Culture and Science of the German State of North Rhine-Westphalia as part of the interdisciplinary research program "Digital Future". This is a collaboration between the University of Paderborn and the University of Bielefeld with researchers from the fields of business administration, computer science, sociology, psychology and engineering to research the topics of CW and data security and privacy in digitalized work processes. In total, 32 German-speaking CW platforms were identified.

Of the 32 platforms, 15 were identified during a literature research focusing on CW platforms as an employment relationship (e.g. in Brabham, 2010; Hensen et al., 2021; Leimeister, Zogaj, Durward, & Blohm, 2016). Additionally, the literature research provided information on relevant terms within the context of CW platforms, which was continuously supplemented during the web research. The web research with the searching terms “crowdworking platform”, “crowdsourcing platform”, “crowd* platform*”, “design platform”, “testing platform”, “microtask platform”, “innovation platform”, “market platform”, “text creation platform”

(searching terms were translated to English) added 17 CW platforms.

These CW platforms were codified by 17 characteristics from six categories. These distinctions were identified through a review of the literature on CW as an employment relationship and expanded to include general characteristics such as age of the platform, size of the community, headquarters, year of foundation, and characteristics of RBCSs. The following characteristics and categories are considered (see Table 1 for more details):

General includes characteristics about the (1) platform type, (2) the size of the community, (3) the headquarter and (4) the founding year. In terms of platform type, this characteristic distinguishes platforms, whether they are platforms for simple and very short tasks, so-called microtask platforms, or platforms for more sophisticated tasks, such as software and web interface testing platforms, text creation platforms, market platforms for a broader range of freelance tasks, design platforms, or platforms for innovation generation (Leimeister et al., 2016).

Task includes the (5) task granularity that distinguished between stand-alone projects or small fragments of it, whereas (6) the task complexity describes whether or not special qualifications beyond basic technical understanding such as programming, design or text creation knowledge are required on the platform (Buettner, 2015; Durward, Blohm, & Leimeister, 2016; Ghezzi, Gabelloni, Martini, & Natalicchio, 2018)

Crowdworker selection summarizes the characteristics of whether the platform uses (7) qualification-based crowdworker selection or (8) context-specific crowdworker selection, e.g., by gender, age, or available technical equipment, or both (Durward et al., 2016; Gadiraju et al., 2017).

Work situation distinguishes (9) contests for the first best result with crowdworkers working exclusively on a solution (i.e. task-oriented contests), (10) contests for the best results with crowdworkers working simultaneously on a solution (i.e. result-oriented contests), or (11) a cooperative work situation for a cooperative solution (Chittilappilly et al., 2016; Durward et al., 2016).

Solution contains information whether the platform focuses on (12) selective solutions, which describe tasks with one or only a few acceptable solutions, or on (13) integrative solutions, which describe the aggregation of many partial solutions into an overall solution (Durward et al., 2016; Pedersen et al., 2013). In addition, the (14) visibility of the solution by other parties was also taken into account (Durward et al., 2016).

Table 1. Overview of 32 German-speaking crowdworking platforms differentiated by 17 characteristics

General			Task		Worker selection		Work situation		Solution		Compensation	
Name	Platform type	Headquarter Community size Founding year	Granularity: (1) High (0) Low	Complexity: (1) High (0) Low	Qualification- based selection: (1) Yes (0) No	Context-based selection: (1) Yes (0) No	Result-oriented context: (1) Yes (0) No	Task-oriented context: (1) Yes (0) No	Solution visibility: (1) Yes (0) No	Solution: (1) Selective (0) Integrative	Rating-based compensation: (1) Yes (0) No	Number of rating levels
99designs	Designing platform	Australia 900,000 2008	0	1	1	0	1	0	1	1	1	3
Brandsupply	Designing platform	Netherlands 26,530 2012	0	1	1	0	1	0	1	1	0	0
Crowdsite	Designing platform	Netherlands 63,368 2009	0	1	1	0	1	0	1	1	0	0
DesignCrowd	Designing platform	Australia 562,444 2008	0	1	1	0	1	0	1	1	0	0
Designertlassen	Designing platform	Germany 76,300 2008	0	1	1	0	1	0	1	1	1	2
jovoto	Designing platform	Germany 80,000 2007	0	1	1	0	1	0	1	1	0	0
Logoarena	Designing platform	Canada 100 2010	0	1	1	0	1	0	1	1	0	0
Talenthouse	Designing platform	USA 750,000 2009	0	1	0	0	1	0	0	1	0	0
Bluepatent	Market platform	Germany - 2011	0	1	1	0	0	1	0	1	0	0
Expertcloud	Market platform	Germany 250 2010	0	0	1	0	0	1	0	1	0	0
Fiverr	Market platform	Israel 100,000 2010	0	1	0	0	1	0	0	1	0	0
Freelancer	Market platform	Australia 22,851,084 2004	0	1	1	0	0	1	0	1	0	0
Twago	Market platform	Germany 627,897 2009	0	1	1	0	0	1	0	1	0	0
VoiceBunny	Market platform	USA 100,000 2012	1	1	1	1	0	1	0	1	0	0
Appjobber	Microtask platform	Germany 1,000,000 2010	1	0	0	1	0	1	0	1	0	0
Clickworker	Microtask platform	Germany 800,000 2005	1	0	1	0	0	1	0	1	1	5
Crowdguru	Microtask platform	Germany 50,000 2008	1	0	1	0	0	1	0	1	0	0
Gprofit	Microtask platform	Germany - 2016	1	0	0	0	0	1	0	0	0	0
Streetspotr	Microtask platform	Germany 445,000 2011	1	0	0	1	0	1	0	1	0	0
Veuro	Microtask platform	Germany - 2010	1	0	0	0	0	1	0	0	0	0
WorkGenius	Microtask platform	Germany 250,000 2011	1	0	1	0	0	1	0	1	1	4
Appraise	Testing platform	USA 250,000 2007	0	1	1	1	0	1	0	0	1	6
Rapiduser test	Testing platform	Germany 20,000 2011	0	1	0	1	0	1	0	0	0	0
test.io	Testing platform	Germany 20,000 2011	0	1	1	1	0	1	0	0	1	2
testbee	Testing platform	Germany 5,000 2018	0	1	0	1	0	1	0	0	0	0
Testbirds	Testing platform	Germany 200,000 2011	0	1	1	1	0	1	0	0	0	0
Testemitt.de	Testing platform	Germany - -	0	1	0	1	0	1	0	0	0	0
Uninspect	Testing platform	Germany - 2012	0	1	0	1	0	1	0	0	0	0
Content	Text creation platform	Germany 6,500 2010	0	1	1	0	0	1	0	1	1	6
Textbroker	Text creation platform	Germany 100,001 2005	0	1	1	0	0	1	0	1	1	4
Textmaster	Text creation platform	Belgium - 2011	0	1	1	0	0	1	0	1	0	0
Tipsofbytes	Text creation platform	Germany 12,000 2007	0	1	1	1	0	1	0	1	0	0

Finally, *compensation* summarizes characteristics describing whether the platform has implemented a (15) short-term, non-reputational, predefined task prize system or (16) a RBCS, as well as (17) the number of implemented rating levels on the respective CW platform.

In the following subsections, the platform overview and its characteristics are used to explain the basic composition of observed RBCSs and to distinguish observed platforms into groups with and without a RBCS to determine their characteristics.

2.2 Composition of rating-based compensation systems

Based on reviewed CW platforms, RBCSs are composed of evaluation criteria forming rating-based performance and behavior thresholds, visualizations of crowdworkers' status and assigned incentives types. Why these components are used, how they can be varied or combined, and what similarities RBCSs of certain platform types show, is part of the following explanations.

Each rating level relies on predefined thresholds for the performance and behavior of crowdworkers on the CW platform. Relevant measurements are very context-specific, but can be basically distinguished between quantitative and qualitative measurements. Starting with quantitative measures coming from platforms and their clients, these usually focus on essential elements of tasks, such as amount of written words, detected software bugs, number of solved tasks; or participation measures such as frequency or recency of participation. Qualitative measures are the result of subjective evaluations through the platform, clients or even other peers. Examples of measures are the fulfillment of task requirements, following task-specific guidelines, for instance for language, design or software test cases, or the behavior, communication or reliability of crowdworkers. Comparing these measures, quantitative thresholds for rating levels are more comprehensible and easier to set than qualitative thresholds. In order to define comparable and comprehensible qualitative thresholds, qualitative criteria for subjective evaluations are usually quantified using scales, such as 5-point Likert scales. Overall, this results in a set of predefined and mostly open communicated thresholds of quantitative and qualitative measures for each rating level, which potentially affect the performance and behavior of crowdworkers on respective CW platforms (Schörpf, Flecker, Schönauer, & Hubert, 2017). By surpassing predefined thresholds, platforms are willing to reward valuable and cooperative crowdworkers and thus work towards an ongoing relationship that benefits both sides.

Granted that the platform is aware about required thresholds for each rating level, the chosen visualization of a crowdworker's rating level communicate to others, namely the platform,

clients' or crowdworkers' peers, the progress of a crowdworker's status development on the CW platform (Basili & Rossi, 2020; Goes et al., 2016; Peer, Vosgerau, & Acquisti, 2014). The most common visualizations of ratings are typical 5-star ratings (e.g. content.de and textbroker.de); badges representing a specific rating level such as rated, proven, bronze, silver or gold badges (e.g. uTest by applause.com); or hierarchically ordered titles for each rating level such as Junior, Mid-Level, Senior and Genius (e.g. workgenius.de). Further or mixed forms of visualizations are possible and can be modified as desired, as long as they can communicate status differences between crowdworkers in a comprehensible way and put them in a hierarchical or specific order. A clear communication of status differences through rating levels seems essential to establish an accepted and comprehensible set of varying incentives for different groups of crowdworkers on the same platform.

Observed incentive types on German-speaking CW platforms that are associated with a RBCS can usually be divided into three groups: minimum pay rates or payout raises; (2) "career" options (3) and varying task availability. Minimum pay rates and payout raises are extensions of the core concept of any CW platform, namely predefined pay rates per task or task prizes (Ghezzi et al., 2018). Minimum pay rates ensure payments above a rating-based minimum level. This type of payment seems to be more common on platforms with clear and measurable solution requirements such as a text with a predefined word count. Thus, platforms that offer text creation tasks in particular tend to establish a RBCS with a minimum pay. This makes both the outcome and the payment for a satisfactory solution potentially more predictable for the platform and the crowdworkers. Whereas CW platforms that offer tasks with vague requirements tend to establish RBCSs with payout raises. These payout raises add a relative or an absolute share on the pay rate for a satisfactory result, depending on a crowdworker's rating. Payout raises seem to be more beneficial for less specified tasks such as software tests or design, since it is hardly possible to define how many errors each tester should find or which design in a competition with multiple designers a client tends to choose. Under these circumstances, additional payout raises can reduce the risk of uncompensated work due to unpredictable task outcomes.

Although CW platforms with a RBCS provide possibilities for status development, there are also platforms that offer in some way minor career opportunities. This is especially a phenomenon on testing platforms with a RBCS. The reason may lie in the necessity for an integrative solution for software test cases, where crowdworkers work on a test case simultaneously but independently and their results are summarized in a final report. This coordination requires experience and reliable community moderators in chats or test cases

(Zogaj, Bretschneider, & Leimeister, 2014). It is therefore more likely that high ranked and reliable testers will become official chat moderators or team leaders for test cases. Some platforms extend this approach by highly individualized development opportunities on the platform through ratings for each test type available on the platform, such as accessibility tests, security tests, or others (e.g. uTest by Applause).

Not only are rating levels used as a steering element for incentives, status and other development opportunities, but also to restrict access to particular tasks. By segmenting crowdworkers based on their past performance and behavior, CW platforms can assign tasks with particular requirements to suitable groups of crowdworkers (Gadiraju et al., 2017). It appears to be very restrictive, but crowdworkers are often unaware of their true skill level and tend to gradually overestimate their own skills and performance (Gadiraju et al., 2017). This in turn potentially increases the risk of unsatisfactory task completions and thus the risk of dissatisfied clients and disappointed crowdworkers. Therefore, it is less surprising that gradually increasing access to tasks is an incentive that all platforms reviewed with a RBCS have created, regardless of their specialization.

In summary, a RBCS allows for many variations in thresholds, visualizations, and incentives, making them applicable to different types of CW platforms, especially those that mediate sophisticated tasks. Despite its possibilities, RBCSs are not that widespread on the German-speaking CW market. Of 32 surveyed CW platforms, only 8 implemented a RBCS on their platform. Possible reasons may be the additional effort and expenses for implementing and controlling such a compensation system as well as effort and expenses that might not be profitable for every type of CW platform.

2.3 Platform characteristics

Depending on the type of a CW platform and the associated platform characteristics, the components of RBCSs - thresholds, visualizations, and incentives - vary in a particular way. These specific platform characteristics also provide further explanation below why some CW platforms with particular task specialization and combination of characteristics tend to implement a RBCS.

An exploratory approach similar to the preparation of a fuzzy set Qualitative Comparative Analysis (fs/QCA) was adopted to highlight certain combinations of platform characteristics that could be possible indications for the implementation of a RBCS. In general, fs/QCA is a method based on Boolean logic and set theory (Ragin, 2000, 2008; Schneider & Wagemann, 2012). Particular combinations of necessary and sufficient conditions form a mechanism that

explains an outcome in causally complex situations (Gerrits & Verweij, 2013). However, a full fs/QCA was not applied because the available platform characteristics as conditions would result in too many conflicting groups covering both platforms with a RBCS and platforms with a non-reputational fixed task prize system, thus provide no clear fs/QCA solutions. One possible reason could be missing characteristics that were not included in the dataset because they were not considered relevant, present, or available. Therefore, an exploratory approach was chosen that still included steps from the fs/QCA. In this approach, the software "fsQCA" (Ragin & Davey, 2016) was used to select those characteristics that minimize the number of platforms in groups that include both platforms with a RBCS and platforms with a non-reputational fixed task prize system. Thus, to the extent possible, this approach generated platform groups with certain characteristics that included only platforms with the same type of compensation system. As a result, seven characteristics stand out from the original 17 characteristics in the platform overview. These seven characteristics describe whether the platform has established task-oriented or result-oriented contests, uses qualification-based selection, focuses on one or a few best solutions (i.e., selective solution) or combines many individual solutions into one report (i.e., integrative solution), was founded later than 2008 or not, and has more than 100,000 registered crowdworkers. The selected thresholds for platform age and number of users were initially derived from trends in platform survey data and further specified by the fs/QCA software.

As can be seen in Table 2, this subset of seven characteristics was then used as the basis for interpretation about commonly shared characteristics and circumstances of the studied platforms with and without a RBCS. What stands out is that each surveyed platform with a RBCS (highlighted in green and yellow in Table 2) has a competitive work environment and has implemented a qualification-based crowdworker selection for each mediated task. This means that the crowdworkers on these specific CW platforms work in a competitive work environment that is dominated by contests for the first satisfactory or the best task completion. In addition to this competitive environment, the access to particular tasks and its contests is restricted by minimum requirements for crowdworkers' reputation and qualifications. Reputation and validated qualifications achieved by crowdworkers are reflected and displayed on the platform by their individual rating level. Therefore, the individual rating levels of crowdworkers are used by the platform as a steering element to coordinate their varying needs for qualification levels in mediated contests for platforms' clients.

Table 2. Overview of German-speaking crowdworking platforms distinguished by seven platform characteristics

<div> <div>Rating (R)</div> <div>No rating (NR)</div> <div>Conflict (R / NR)</div> <div>No observations</div> </div>		Task-oriented contest				Result-oriented contest			
		Qualification-based selection		No qualification-based selection		Qualification-based selection		No qualification-based selection	
		Selective solution	Integrative solution	Selective solution	Integrative solution	Selective solution	Integrative solution	Selective solution	Integrative solution
User (>100.000)	Platform age (> 2008)	M^R R (1) / NR (2)	T NR (1)	MI NR (2)				M NR (2)	
	Platform age (≤ 2008)	TXT^R, M R (2) / NR (1)	T R (1)			D R (1) / NR (1)			
User (≤ 100.000)	Platform age (> 2008)	TXT^R, M R (1) / NR (3)	T R (1)		MI, T NR (6)	D NR (3)			
	Platform age (≤ 2008)	MI, TXT NR (2)				D R(1) / NR (1)			

Note:

- Platform specialization: **MI**: platform for easy and very short tasks or microtasks; **T**: platform for testing tasks; **M**: Market platform for a broad spectrum of tasks; **TXT**: platform for text creation tasks; **D**: platform for graphic designing tasks
- ^R: Marks the platform type with a rating-based compensation system
- Number of observed platforms in parentheses

Building on that, Table 2 also shows that all CW platforms which were reviewed exhibited a RBCS deal with challenging task types with different qualification requirements such as design, software or web interface testing, text creation, or broader defined freelance tasks. In contrast to platforms with a RBCS, a characteristic of platforms without such a compensation system are short-term fixed pay rates per task or fixed task prizes that do not fully take into account crowdworkers reputation or qualifications. This further suggests that RBCSs may be a more appropriate tool for highly specialized platforms that offer challenging tasks. An indicator for the long-term nature of RBCSs may be the number of rating levels on a CW platform. As Table 1 shows, the number of rating levels on each CW platform varies between 2 and 6 attainable levels. Since each rating level is coupled with a set of challenging and time-consuming thresholds for crowdworkers, it can be assumed that platforms are not willing to compensate crowdworkers with unknown performance levels and behaviors at the highest compensation rate after only a short period of time. Long-term or even commitment intentions with valuable crowdworkers seem plausible.

In summary, a RBCS has been implemented primarily by CW platforms, which specialize in specific and demanding task types, such as designing, testing or task creation. The platforms examined show that many variations in performance and behavior thresholds, visualizations and incentives are possible. These variations seem to be necessary to meet the different

requirements of particular demanding task types. However, it was observed that depending on the specialization of a platform, similar RBCS components and combinations thereof are used. For example, CW platforms for text creation usually offer minimum pay rates per word and access to particular task pools depending on a crowdworker's star rating, while platforms for software and web interface testing offer pay out raises for completed test cases and access to further test cases depending on a crowdworker's rating for a specific test types, which is usually visualized by badges. These CW platforms are also characterized as having a highly competitive work environment and need to select crowdworkers based on their reputation and qualifications. Unfortunately, the platforms cannot fully rely on external or self-reported information from crowdworkers (Gadiraju et al., 2017). Therefore, CW platforms must collect information themselves, assign individual rating levels to each registered crowdworker and use these ratings as steering elements to coordinate their registered crowdworkers. This further underlines the long-term nature of RBCSs due to demanding thresholds for each rating level and the time required to gather reliable information.

However, RBCSs were not implemented by all examined platforms that are suitable for them. Brandsupply, Testbirds, TripsbyTips and Twago are exemplary CW platforms for their type of specialization. Although they specialize in challenging CW tasks and are comparable platforms that already use RBCSs, they have not themselves implemented such a compensation system. Possible reasons may be a lack of awareness that a RBCS would meet their requirements, or the effort and cost of implementing such a compensation system are considered too high. Nevertheless, RBCSs and its variations have already been observed on several CW platforms, and although they could potentially address the incentive and commitment problems of the platforms, they are not very widespread, at least in the German-speaking CW market.

3 Theoretical foundations of rating-based compensation systems

3.1 Related work

Despite the active use of RBCSs on CW platforms, long-term compensation and ongoing employment relationships between platforms and crowdworkers play a rather subordinate role in the CW research (Hemsen et al., 2021). CW research has strongly focused on the optimization of CW processes, for instance, by providing uniform short-term compensation for crowdworkers, namely with predefined task prizes despite their different reputation or qualifications (Hemsen et al., 2021). However, the research on compensation falls mainly into two areas: monetary and non-monetary compensation.

The research on monetary compensation is dedicated to extrinsically motivated crowdworkers on CW platforms through financial compensation. In particular, corresponding studies vary task prizes (e.g. Hsieh & Kocielnik, 2016; Lee, Chan, Ho, Choy, & Ip, 2015; Liu, Yang, Adamic, & Chen, 2014), pay additional bonuses (e.g. Ho, Slivkins, Suri, & Vaughan, 2015; Ming, Yiling, & Yu-An, 2013, 2014), use different payout intervals for the earned income (e.g. Ho et al., 2015; Ikeda & Bernstein, 2016) or even shift from individual payouts to group incentives (e.g. Riedl & Wooley, 2017). Overall, these studies show that higher pay rates or task prizes are associated with a significantly higher participation (e.g. Hsieh & Kocielnik, 2016; Lee et al., 2015), qualitative performance in form of accuracy and novelty (e.g. Liu et al., 2014), quantitative performance (e.g. Füller, Hutter, & Fries, 2012; Yang, 2012), and a higher job satisfaction (Brawley & Pury, 2016). Despite the different positive effects of monetary incentives, the vast majority of the respective studies focus only on a short-term aspect of performance and do not consider its development over time.

Research into non-monetary compensation on CW platforms frequently refer to non-cash awards, which are given in recognition of a high level of accomplishment or performance (Rose, 1998). Corresponding studies analyze the recognition of accomplishments or performance through gamified elements such as points, ratings, or rankings as forms of non-monetary compensation (e.g. Feng et al., 2018; Goes et al., 2016; Goh, Pe-Than, & Lee, 2017). These gamified elements are widespread on CW platforms and often visible to others, allowing the status of a crowdworker to be clearly communicated (Goes et al., 2016). Summarized, non-monetary incentives in form of gamified elements are mostly associated with positive influences on crowdworkers' participation (Feng et al., 2018; Franke, Keinz, & Klausberger, 2013), quantitative (Yang, 2011) or qualitative performance (Dalle, den Besten, Martínez, & Maraut, 2017; Goh et al., 2017; Lee et al., 2015). Accordingly, crowdworkers are more willing to submit their results (Franke et al., 2013), participate in contests (Feng et al., 2018), work more accurately and punctually (Goh et al., 2017), and perform more tasks (Yang, 2011) when non-monetary compensation is offered. Although there are differences between the different types of gamified elements, their effects remain largely positive (Hemsen et al., 2021). Moreover, non-monetary compensation also target several important intrinsic motivations (Goh et al., 2017; Ryan & Deci, 2000): presenting oneself in a preferred manner to others (i.e. self-presentation), proving one's own capabilities (i.e. self-efficacy), to communicate, interact or compete with other people (i.e. social bonds), or simply to enjoy the process of participation (i.e. playfulness).

Overall, a whole range of components is available to design a monetary or non-monetary

compensation system on CW platforms. Despite the research focus on monetary compensation and even gamified elements such as ratings on CW platforms, to my knowledge these two aspects have not been neither combined nor studied in detail in the context of CW. As a result, it is largely unclear how RBCSs affect crowdworkers' motivation and commitment to the platform when they have the opportunity to improve their pay or working conditions through their past performance and behavior, or how RBCSs could be used to build ongoing employment relationships between platforms and crowdworkers. Therefore, research on RBCSs as a long-term compensation system on platforms would be a reasonable extension for CW research.

There have been few attempts to examine the varying effects of non-monetary rating systems on CW platforms. Despite differing views on non-monetary rating systems, relevant studies focus primarily on the performance or participation of crowdworkers in contests or tasks as desired outcomes. Studies that emphasize the performance of crowdworkers consider non-monetary ratings as a visualized form of a crowdworker's reputation on the platform (Basili & Rossi, 2020; Goes et al., 2016; Peer et al., 2014) or as a signal for crowdworkers to adapt their behavior to the requirements of the platform and their clients (Riedl & Seidel, 2018). With crowdworkers' participation as a desired outcome, studies consider non-monetary ratings as a virtual reward system with playful 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 the ongoing work process (Jian, Yang, Ba, Lu, & Jiang, 2019). Schörpf et al. (2017) offer a more holistic view of ratings. They point out that ratings are a form of a control feature and thus part of the platform design, which have a particularly strong impact on the time use, income and creativity of crowdworkers and thus on their working and living conditions. Although several perspectives are taken when considering non-monetary ratings, providing evidence of possible positive effects of at least the non-monetary component of a RBCS, the lack of theoretical and empirical evidence on the combined effects of monetary and non-monetary incentives of RBCSs remains.

3.2 Theory

Since CW platforms for sophisticated tasks, which rely on expert crowdworkers with different qualification levels and their frequent participation, decided to implement a RBCS, it is reasonable to assume that these platforms have been attempting to address and solve incentive and commitment issues. How RBCSs motivate and commit crowdworkers to a CW platform is reconstructed by drawing on concepts developed for regular employment. In particular, the

hierarchy of rating levels resembles internal labor markets, and the incentive raises connected with higher ratings can be interpreted as a form of deferred compensation. The theoretical basis is provided by the goal-setting theory of Locke and Latham (2002). Accordingly, aiming for and achieving higher rating levels and the associated incentive increases could also increase crowdworkers' motivation and commitment to the CW platform in the long run. How these concepts and theory are transferred from regular employment to CW is illustrated in Figure 2 and argued in detail below.

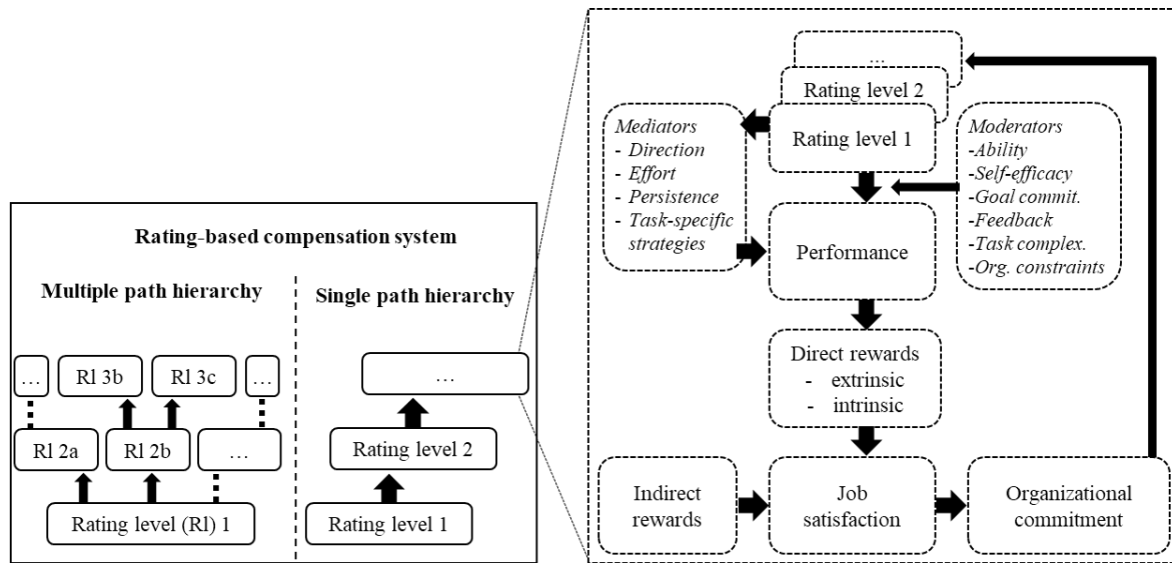


Figure 2. Rating-based compensation systems as an equivalent of internal labor markets and deferred compensation on crowdworking platforms (left part) and rating levels as challenging goals (right part, inspired by Watzka (2017))

In general, internal labor markets are an “[...] administrative unit, [...], within the pricing and allocation of labor is governed by a set of administrative rules and procedures” (Doeringer & Piore, 1985: 8–9). In contrast to the external labor market, the internal labor force has exclusive rights to jobs filled internally, continuity of employment, and even at entry ports, they are protected from direct competition by workers in the external labor market (Doeringer & Piore, 1985). The main reasons that lead to internal labor markets in an organization are the need for skill specificity, on-the-job-training and customary law (Doeringer & Piore, 1985). Skill specificity leads to internal labor markets, because it increases the proportion of training costs borne by the organization and it increases the absolute level of such costs (Doeringer & Piore, 1985). On-the-job trainings makes the skills which it produces highly specific to the context in which they were acquired (Doeringer & Piore, 1985). Lastly, it is also customary at the work place to have an unwritten set of rules based largely upon past practice (Doeringer & Piore,

1985). Depending on the effort and cost required, it is more beneficial for organizations to retain employees with such a set of skills specific to the organization.

All three reasons mentioned above for considering an internal labor market also apply to CW platforms with few exceptions. Even though specialized CW platforms require crowdworkers with a high skill specificity, they are also dependent on already qualified crowdworkers (Boons et al., 2015; Schulten & Schaefer, 2015). This means, crowdworkers with suitable general knowledge of certain task types who will acquire platform-specific knowledge over time without further training. For this reason, on-the-job training and familiarization with platform customs are an essential part of the work on CW platforms, making experienced and skilled crowdworkers very valuable. This is especially the case since there are no employment contracts involved and these crowdworkers could leave at any time. However, the consequences for a platform trading in demanding tasks are more serious in the long run, since the loss of valuable crowdworkers also means the loss of information about the platform's and client's customs.

Internal labor markets seem to be suitable not only on CW platforms, a RBCS already provides similar functionality. That means that each higher rating level is the next higher position in the platform hierarchy that is filled internally with registered crowdworkers and comes with assigned compensation. Even multiple development paths with individual rating levels for each task type are observed. Not exclusive to platforms with a RBCS, but as long as crowdworkers don't violate the platform's guidelines, they can usually rely on the continuity of that relationship. Another important functionality is the protection against direct competition from crowdworkers outside the platform to a certain extent, as crowdworkers usually start with a lower rating level on a platform and thus do not represent immediate competition for already higher rated crowdworkers.

What distinguishes RBCSs from conventional internal labor markets is that although a higher rating is associated with tasks with higher quality requirements, these tasks are at least similar to tasks for lower ratings. It is therefore unlikely that expert crowdworkers who move up to the next rating level will suddenly find themselves insufficiently qualified for the tasks, known as the Peter principle (Lazear, 2004). Even if this were the case, these crowdworkers would face a degradation to a lower rating level, which is unusual in internal labor markets for regular employments (Doeringer & Piore, 1985). Internal labor markets also rather use payment based on an employee's position within an organization because performance is harder to measure (Lazear & Gibbs, 2014) than on CW platforms, which explains the use of performance-based

pay. Despite some differences, the similarities between RBCSs and internal labor markets seem to prevail.

Furthermore, forms of deferred compensation and acknowledging crowdworkers membership duration on a CW platform are also part of RBCSs and their similarity to internal labor markets. According to Lazear (1990), 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”. Due to time-consuming and challenging thresholds for each rating level, higher rating levels and associated incentive raises also become more likely with a longer membership duration. As a result, crowdworkers who remain with the CW platform for a longer period of time receive as a bonus task prizes that may exceed their productivity or the average compensation on the market. This also means that crowdworkers who have recently registered on the platform and have not yet validated their qualifications may be paid below their productivity. A circumstance which possibly attracts rather crowdworkers who intend to stay on the CW platform and be active there.

Unlike deferred compensation in regular employment, crowdworkers who have been working on a CW platform for a long time do not have to worry about being laid off because their payments become too expensive the longer they work for the platform. From a more neoclassical perspective, an organization usually has an interest in laying off employees who are paid above their productivity when the savings from the time an employee was usually paid below his or her productivity early in his or her career in the organization have been used up (Lazear, 1990). This would not necessarily be the case for CW platforms with a RBCS due to direct performance monitoring and an interest in continuing the relationship with difficult to replace skilled crowdworkers (Boons et al., 2015; Schulten & Schaefer, 2015). Therefore, crowdworkers who cannot maintain a required level of performance are more likely to face a downgrade to a lower rating level and be paid accordingly than to have the relationship terminated. In this way, CW platforms can offer higher compensation to motivated, engaged crowdworkers and a competitive pricing to their clients.

Based on similarities to these concepts, a RBCS is not only a promising tool to motivate valuable crowdworkers to the respective CW platform, but also to commit platforms to crowdworkers. Since specialized CW platforms rely on their expert crowdworkers to satisfactorily complete mediated tasks, they also commit themselves through higher payments and other granted privileges. A point often overlooked, even in regular employment relationships, is the importance of an organization’s commitment to its employees (Atchison,

1991; Lawler III, 1986; Osterman, 1988). A variety of authors argue that employees perform better, demonstrate more citizenship behavior, and express a higher level of commitment to an employer when they worked in an exchange that is equally beneficial for the employer and its employees (Atchison, 1991; Lawler III, 1986; Osterman, 1988). These benefits, however, cannot be realized unless employers are willing to offer some degree of security or career investments in its employees (Williamson, Wachter, & Harris, 1975). Therefore, CW platforms, especially for challenging tasks, are well advised to send reliable signals through an appropriate compensation and development opportunities for reciprocating crowdworkers.

Rating levels of a RBCS are likely to serve as such credible signals and as desirable goals for crowdworkers, as it enables them to actively improve their working conditions. Building on this rating-as-goal premise, Figure 2 illustrates that Locke and Latham's (2002) goal-setting theory provides the theoretical basis to explain how the internal platform hierarchy, each rating level, and their incentive increases can increase crowdworkers' motivation and commitment in the long run.

In general, according to Locke and Latham's (2002) goal-setting theory, well-specified and challenging goals can substantially promote an employee's performance and his or her commitment to the organization. This can be explained by a particular circular process. First, there is a well-researched effect of specified and challenging goals on an employee's performance (Locke & Latham, 2002). This relation is particularly moderated by the employee's abilities, perceived self-efficacy of her or his own skills, the commitment to the goal, received feedback, task complexity and organizational constraints such as technical resources (Locke & Latham, 2002). Second, goals can also direct attention to relevant activities, activate and maintain effort, and foster task-specific strategies to further positively influence performance (Locke & Latham, 2002). Employees who were able to achieve the goal are usually rewarded extrinsically, for example by a higher income, a bonus, or even a promotion, and intrinsically by, for instance, the feeling of success (Locke & Latham, 2002). This is not only associated with increases in job satisfaction and organizational commitment, but also with an increased willingness of the employee to commit themselves to the next achievable goal (Locke & Latham, 2002). In turn, this commitment can initiate the next round of this process just described for a different goal.

Applied to the CW context, it seems plausible that a CW platform with a RBCS can also initiate such a cycle that gradually increases crowdworkers' performance and commitment to the platform. It is likely that crowdworkers who ascent the internal platform hierarchy experience

that they are able to progress with each higher rating level and overcome challenges such as the rating-specific thresholds by their own. In addition, achieving a higher rating level is of importance for crowdworkers as it basically determines a crowdworkers' working conditions on the respective platform such as income, status and other offered privileges. Taken together, this describes a process that can promote a crowdworkers' satisfaction and commitment to the platform, not only for rational but also for emotional reasons, as crowdworkers' may feel valued by the platform. Under these assumptions, a RBCS is a promising tool that is both beneficial for CW platforms and crowdworkers.

4 Discussion

4.1 Implications

The main contribution of this paper is to add long-term RBCSs, a newly formed term, to the CW literature and to show similarities to established concepts of regular employment, namely internal labor markets and forms of deferred compensation. Due to the lack of specification of RBCSs in the literature, this paper specifies and discusses key components of RBCSs based on 32 reviewed German-speaking CW platforms with and without such a system. Despite the observed variations in performance and behavior thresholds, visualizations, and assigned incentives in RBCSs, commonly used components were observed for specific platform types. Common to all observed CW platforms with a RBCS is that their work environment is highly competitive and crowdworkers are selected and assigned to task pools based on their reputation and qualifications. Crowdworker qualities relevant to the CW platform are represented by their individual rating level. These rating levels put crowdworkers in a hierarchical or specific order. The higher the rating level, the higher the compensation and privileges of the individual crowdworker. For this reason, similarities to internal labor markets and deferred compensation in regular employment have been discussed. In fact, RBCSs have several similarities and can be considered as a possible equivalent on CW platforms, with a few exceptions, such as pay by performance rather than by position and the absence of employment contracts. That moving up within a rating hierarchy could have positive effects on crowdworkers' performance and behavior to the platform is provided by the goal-setting theory as a theoretical basis. Overall, RBCSs seem to be a promising tool to commit both specialized CW platforms and experienced crowdworkers.

Understanding long-term compensation for crowdworkers such as RBCSs is crucial, given the ongoing relationships between CW platforms and their registered crowdworkers (Giard et al.,

2019). In particular, highly specialized CW platforms mediating design or text creation tasks, for example, rely on experienced crowdworkers. These experts are not only valuable because of their skills, they are also familiar with the practices of the platform and its clients, which is difficult to replace (Schulten & Schaefer, 2015). Since crowdworkers' reputations and qualifications are not reliably transferable between CW platforms, platforms need to collect relevant information about their registered crowdworkers themselves (Gadiraju et al., 2017). Therefore, they need to ensure continuous participation and commitment of their valuable crowdworkers for an ongoing business (Boons et al., 2015; Schulten & Schaefer, 2015). Crowdworkers, on the other hand, may feel valued and receive more appropriate compensation for achieved reputation and validated qualifications. This could pave the way for a more mutually beneficial relationship between CW platforms and experienced crowdworkers.

The long-term view of RBCSs and employment relationships between CW platforms and crowdworkers does not contradict the findings of previous CW research. Rather, it complements the short-term view with a long-term perspective of work on CW platforms and offers a promising long-term compensation system. Moreover, this work transfers extant work on employment relationships from personnel economics to CW. Although CW differs in its form of work organization through an online platform, the interactions between a platform and its crowd nevertheless have similarities to employment relationships in regular organizations.

This has practical implications for researchers and CW platforms. Researchers should pay more attention to the similarities of CW platforms to regular work organizations and thus to the extant body of work on employment relationships. More in-depth research should be conducted to determine what other concepts and theories might be applicable to CW. Presumably, this would contradict the claim of a new and disruptive form of work (Howe, 2006) and rather emphasize that CW is primarily a rearrangement of already known processes and information flows, like other new forms of work organization (Puranam, Alexy, & Reitzig, 2014). Given the similarities to regular employment, CW platforms may also be well advised to take more responsibility for their registered crowdworkers. Mainly because of the lack of employment contracts, CW platforms need other means to increase their attractiveness. This is because once the working conditions appear unacceptable for a crowdworker, there are almost no barriers to leaving the platform. Accordingly, CW platforms should anticipate this movement.

4.2 Limitations and future research

This work is not without limitations. First, although the CW platforms were selected according to defined criteria, there may be some sort of observer bias in the selection of platforms. This

limitation should be mitigated by the breadth of the dataset of CW platforms reviewed and, additionally, by having an otherwise uninvolved assistant verify that platforms meet the criteria and whether other platforms may be missing. Second, the selection and collection of relevant platform characteristics may not completely avoid any loss of information. This issue was addressed by drawing on previous work that focused on characteristics of CW platforms and by continuously expanding these characteristics during the web search. In order to gain more detailed insights into the design of the platforms, both a client and a crowdworker account were created on each platform mentioned in this paper, as far as the platform allowed. Third, the study was partially limited by the fact that it focused in particular on platforms with at least a German-speaking web interface. However, the platforms studied are globally accessible, and not all were founded in Germany. This limitation was chosen to represent a CW market that is broadly accessible to German crowdworkers without potential language barriers and to increase comparability as these platforms attempt to target similar workforces. Furthermore, the focus on the German CW market was of particular interest as this research was funded by the Ministry of Culture and Science of the German State of North Rhine-Westphalia as part of the interdisciplinary research program "Digital Future". Fourth, no empirical evidence for the effectiveness of RBCSs is provided. Possible effects of such a compensation system, for instance on crowdworkers' performance or commitment to the platform, are entirely based on related work and by drawing on existing work on employment relationships from personnel economics. Although this work is intended to be a conceptual introduction to the under-researched topic of RBCSs in CW research, empirical evidence would still be a meaningful addition to this work.

Based on these limitations, there are also some implications for future research. Research on RBCSs should aim to provide empirical evidence on the effectiveness of RBCSs on different platform types and their effects, e.g., on crowdworkers' motivation and commitment to the CW platform. Considering that CW research relies strongly on exploratory or cross-sectional datasets (Hensen et al., 2021; Zhao & Zhu, 2014), longitudinal data would be desirable in this context, especially to explore performance and commitment trends over time. Interesting research questions that could also be addressed include: Are RBCSs more effective at committing and motivating crowdworkers than non-reputational fixed task prize systems? How do differently rated crowdworkers differ in terms of performance and behavior? How many rating levels should be implemented in a RBCS? What mechanisms can explain the commitment and performance of CW platforms with a RBCS? What incentives can be derived from these mechanisms? How do groups of committed crowdworkers differ in terms of

performance and behavior? Answering some of these questions can pave the way for a better understanding of a more desirable digital work environment and more mutually beneficial employment relationships in CW.

5 References

- Atchison, T. J. 1991. The employment relationship: un-tied or re-tied? *Academy of Management Perspectives*, 5(4): 52–62.
- 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.
- Berg, J. M. 2016. Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers. *Comparative labor law and policy journal*, 37(3): 506–543.
- 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. 2010. Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. *Information, Communication & Society*, 13(8): 1122–1145.
- Brabham, D. C. 2012. The myth of amateur crowds. *Information, Communication & Society*, 15(3): 394–410.
- Brawley, A. M., & Pury, C. L. S. 2016. Work experiences on MTurk: Job satisfaction, turnover, and information sharing. *Computers in Human Behavior*, 54: 531–546.
- Buettner, R. 2015. A systematic literature review of crowdsourcing research from a human resource management perspective, *2015 48th Hawaii International Conference on System Sciences*: 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.

- Doeringer, P. B., & Piore, M. J. 1985. *Internal Labor Markets and Manpower Analysis* (1st ed.). Armonk, New York, USA: Me Sharpe.
- Durward, D., Blohm, I., & Leimeister, J. M. 2016. Crowd Work. *Business & Information Systems Engineering*, 58(4): 281–286.
- Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. 2012. Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2): 189–200.
- Feng, Y., et al. 2018. Gamification artifacts and crowdsourcing participation: Examining the mediating role of intrinsic motivations. *Computers in Human Behavior*, 81: 124–136.
- Fieseler, C., Bucher, E., & Hoffmann, C. P. 2019. Unfairness by Design? The Perceived Fairness of Digital Labor on Crowdworking Platforms. *Journal of Business Ethics*, 156(4): 987–1005.
- 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.
- Füller, J., Hutter, K., & Fries, M. 2012. Crowdsourcing for Goodness Sake: Impact of Incentive Preference on Contribution Behavior for Social Innovation, *Interdisciplinary Approaches to Product Design, Innovation, & Branding in International Marketing*: 137–159: Emerald Group Publishing Limited.
- 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.
- Gerrits, L., & Verweij, S. 2013. Critical Realism as a Meta-Framework for Understanding the Relationships between Complexity and Qualitative Comparative Analysis. *Journal of Critical Realism*, 12(2): 166–182.
- 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.
- Hemsen, P., et al. 2021. Crowdsourcing for paid work as a new form of employment relationship: A content analysis. *Working Paper*.
- Ho, C.-J., et al. 2015. Incentivizing High Quality Crowdwork. *Proceedings of the 24th International Conference on World Wide Web*: 419–429.
- Howe, J. 2006. The Rise of Crowdsourcing. *Wired Magazine*, 14(6): 1–5.
- 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.
- Ikeda, K., & Bernstein, M. S. 2016. Pay It Backward: Per-Task Payments on Crowdsourcing Platforms Reduce Productivity. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*: 4111–4121.
- 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.
- Kuek, S. C., et al. 2015. *The global opportunity in online outsourcing*. World Bank Group: Washington, DC.
- 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.
- Lee, C., et al. 2015. Explore the feasibility of adopting crowdsourcing for innovative problem solving. *Industrial Management & Data Systems*, 115(5): 803–832.
- Leimeister, J. M., et al. 2016. *Systematisierung und Analyse von Crowd-Sourcing-Anbietern und Crowd-Work-Projekten*. Düsseldorf: Hans-Böckler-Stiftung.
- 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.
- Ming, Y., Yiling, C., & Yu-An, S. 2013. The Effects of Performance-Contingent Financial Incentives in Online Labor Markets. *27th AAAI Conference on Artificial Intelligence (AAAI'13)*.
- Ming, Y., Yiling, C., & Yu-An, S. 2014. Monetary Interventions in Crowdsourcing Task Switching. *Proceedings of the Second AAAI Conference on Human Computation and Crowdsourcing (HCOMP 2014)*: 234–241.
- Osterman, P. 1988. *Employment futures: Reorganization, dislocation, and public policy*: Oxford University Press, USA.
- 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.
- 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.
- Ragin, C. C. 2000. *Fuzzy-Set Social Science* (1st ed.). Chicago: University of Chicago Press.
- Ragin, C. C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond* (1st ed.). Chicago: University of Chicago Press.
- Ragin, C. C., & Davey, S. 2016. Fuzzy-Set/Qualitative Comparative Analysis 3.0. Irvine, California: Department of Sociology, University of California. *Irvine, California: Department of Sociology, University of California*. <http://www.socsci.uci.edu/cragin/fsQCA/software.shtml>.
- Riedl, C., & Seidel, V. P. 2018. Learning from Mixed Signals in Online Innovation Communities. *Organization Science*, 29(6): 1010–1032.
- Riedl, C., & Wooley, A. W. 2017. Teams vs. Crowds: A Field Test of the Relative Contribution of Incentives, Member Ability, and Emergent Collaboration to Crowd-Based Problem Solving Performance. *Academy of Management Discoveries*, 3(4): 382–403.
- Rose, M. 1998. Performance-related pay in schools: An assessment of the green papers. *NUt*:

London.

- Ryan, R. M., & Deci, E. L. 2000. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary educational psychology*, 25(1): 54–67.
- Schneider, C. Q., & Wagemann, C. 2012. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis* (1st ed.). Cambridge: Cambridge University Press.
- 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.
- Watzka, K. 2017. *Zielvereinbarungen in Unternehmen: Grundlagen, Umsetzung, Rechtsfragen* (2nd ed.). Wiesbaden: Springer Gabler.
- Williamson, O. E., Wachter, M. L., & Harris, J. E. 1975. Understanding the Employment Relation: The Analysis of Idiosyncratic Exchange. *The Bell Journal of Economics*, 6(1): 250–278.
- Yang, J. 2011. *Incentive and Culture: Shaping Information and Social Dynamics in Online Information Sharing Systems*. Dissertation. University of Michigan. Michigan.
- Yang, Y. 2012. *Open innovation contests in online markets: Idea generation and idea evaluation with collective intelligence*. Dissertation. Temple University Graduate Board. Philadelphia, USA.
- Zhao, Y., & Zhu, Q. 2014. Evaluation on crowdsourcing research: Current status and future direction. *Information Systems Frontiers*, 16(3): 417–434.
- Zogaj, S., Bretschneider, U., & Leimeister, J. M. 2014. Managing crowdsourced software testing: A case study based insight on the challenges of a crowdsourcing intermediary. *Journal of Business Economics*, 84(3): 375–405.