Bargaining with the Black-Box

Designing and Deploying Worker-Centric Tools to Audit Algorithmic Management

DAN CALACCI, MIT Media Lab, USA
ALEX PENTLAND, MIT Media Lab, USA

The increasing prevalence of large-scale labor aggregation platforms, worker analytics, and algorithmic decision-making by management raises the question of whether workers can use similar technologies to advocate for their own goals. Yet, there are inherent challenges in building worker-centric tools that collect, aggregate, and share data in responsible and ethical ways. In this paper, we present the design and deployment of the Shipt Calculator, a tool developed in collaboration with non-profit worker groups that allows app-based delivery workers to track and share aggregate data about their pay, increasing wage transparency. We first discuss the design challenges inherent to building worker-centric technologies, particularly for informally organized workers, and ground our discussion in the history of worker inquiry and co-research. We then describe some principles from this history and our own lessons in designing the Calculator that can be applied by future researchers and advocates seeking to build technical tools for organizing campaigns. Finally, we share the results of using the Calculator to audit an app's shift to a black-box pay model using data contributed by 140 workers in the Summer of 2020, finding that although the average pay per-order increased under the new payment model, almost half of workers experienced an unannounced pay cut during the shift, and many workers worked shifts that paid under their state's minimum wage. Finally, we discuss how tools like the Calculator demonstrate the important role that aggregate worker data, and a new Digital Workerism, can serve in creating and maintaining a more balanced platform economy.

Additional Key Words and Phrases: Platform Work; Organizing; Labor; Data Advocacy

ACM Reference Format:
Dan Calacci and Alex Pentland. 2022. Bargaining with the Black-Box: Designing and Deploying Worker-Centric Tools to Audit Algorithmic Management. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 428 (November 2022), 24 pages. https://doi.org/10.1145/3570601

1 INTRODUCTION

On October 15, 2020, a group of workers took to the streets of Birmingham, Alabama, to go on strike. Their company had recently overhauled management, claiming that this change would result in fairer pay for all workers they hire and contract. But, after the change, some workers reported that this new management structure suddenly resulted in a slimmer paycheck [5]. These workers were part of the two-hundred-thousand person workforce contracted by Shipt, an app-based delivery company acquired by Target in 2017, and their demands meant more than just asking for a raise or fairer working conditions. They meant demanding wage transparency from the algorithm serving as their newly redesigned manager.

The problem that these workers faced—securing wage and algorithm transparency—is an active area of research and advocacy by labor and design scholars [27, 28, 39, 52]. Recent reporting suggests that the algorithms used by several app-based delivery platforms have enabled digitally-mediated wage theft, with pay algorithms e.g. effectively removing tips given to workers by customers [56].

Authors' addresses: Dan Calacci, dcalacci@media.mit.edu, MIT Media Lab, Cambridge, MA, USA, 02139; Alex Pentland, sandy@media.mit.edu, MIT Media Lab, Cambridge, MA, USA, 02139.

2573-0142/2022/11-ART428
https://doi.org/10.1145/3570601

While transparency or open access to the payment algorithms themselves would make auditing wage fairness and theft a cut-and-dry issue, many of the algorithms used by platforms are considered trade secret or intellectual property, limiting the ability to audit, reverse engineer, or share their implementations [16].

These limitations have led scholars, workers, and advocates to explore how worker-driven self-reporting techniques might produce insight into platform worker conditions [34, 39, 56]. Some of the most commonly cited statistics published on delivery-app worker wages come directly from organized workers responding to "self-report questionnaires" or online surveys from sources such as the popular The Rideshare Guy [51]. While such surveys are an important form of modern *worker inquiry*, a concept with a storied history in labor organizing and theory [18, 26], they face severe limitations. For modern workers whose jobs involve algorithmically-mediated management, surveys alone are not sufficient for digital organizing and issue campaigns. New digital tools that facilitate worker data sharing and aggregation, algorithmic audits, and research into workers' conditions are needed.

This paper presents the design and deployment of the Shipt Calculator, a worker-led SMS chatbot tool designed to audit an algorithmic pay change instituted by Shipt, an app-based delivery company, in the Summer and Fall of 2020. The tool was designed in collaboration with worker-organizers and advocates, and was oriented around a specific organizing campaign. Data contributed by workers reveals that although Shipt described its new pay model as fair and improved, over 40% of workers that participated in our study received an unannounced pay cut.

To our knowledge, this paper is the first contribution to the CSCW community that audits an algorithmic change to platform worker pay performed in collaboration with worker-organizers. In addition to describing the audit and its results, this paper offers two other main contributions. First, we present an extended background and discussion that frames the Calculator's development as a step towards "Digital Workerism", an emerging field of digital organizing that leverages data and worker collectives to effect change. Second, we share the design process behind the tool, documenting how researchers collaborated with worker-organizers and how worker feedback was incorporated into its design.

This paper is organized into three main sections. First, we discuss the problems of information asymmetry, algorithmic and wage transparency, and worker organizing that advocates and researchers face in the context of platform-based work. We then present our case study of the Shipt Calculator and the results of using the Calculator to audit how Shipt's transition to a black-box pay model impacted worker pay in the Summer of 2020, finding that 41% of workers studied received an unannounced pay cut. Finally, we discuss design lessons from building and using the calculator in an organizing campaign, the limitations of the technology design, and steps for future work in creating tools that build additional capacity for digital worker inquiry in algorithmically mediated labor and in worker organizing more generally.

2 RELATED WORK & CONTRIBUTIONS

This paper rests at the intersection of two long-standing threads of research in the CSCW community: (1) tool design for activism, and (2) systems to help increase wages and the well-being of workers.

2.1 Designing Activist Tools

Other tools built in the CSCW community have also been designed to aid in specific kinds of activist actions and campaigns. *Out of Site* is a chrome extension that helps streamline the complex actions involved in participating in a consumer boycott through altering the webpages that users visit [41]. Other projects focus on the problem of calling people to engage in collective action, such as the *Botivist* project , which uses an automated system to "test different strategies for calling volunteers to action" [50]. These projects focus on reducing friction in taking collective actions (boycotting, collective discussion), and allow organizers to create their own campaigns. While the Shipt Calculator also focuses on a specific collective action—an algorithmic audit—the collective action is dependent on data sharing and aggregation between participants, not just reducing friction in collective action.

Because the Shipt Calculator is a data-based tool, it is also in conversation with work like [55] and [54], which explore the idea of "data leverage" and the power of "data strikes", respectively. Data Leverage refer to "data-related actions that harm [an] organization's technologies or help its competitors' technologies" [55]. In contrast to "harm-based" data leverage, where e.g. users of a service delete their data to harm an algorithm's effectiveness, the authors also present what they call "Conscious Data Contributions" (CDC). CDC are efforts by consumers to share data with an organization they support to increase market competition and put pressure on a organization or firm they would like to impact. In the case of the Shipt Calculator, participants instead share data with a loose activist group—worker-organizers, advocates, and researchers—with the goal of both understanding their working conditions and revealing the impact of a company's algorithm on its workers.

2.2 Systems to Increase Worker Well-Being

This work is also in conversation with other projects that aim to increase wages, well-being, and material conditions for workers. There is a thread of work in the CSCW and CHI communities focused on improving working conditions for workers by researching tools that are designed to "keep…workers both happy and productive" [37]. These tools include predicting and optimizing information worker "state transitions" [37], chat-based interventions to prevent "cyberloafing" [53], and peer systems for upskilling and improving work quality [19].

Although such studies engage directly with workers, many are often built from the perspective of management or "the boss". They uncritically approach improving well-being as a pathway to greater worker productivity (the "happy-productive worker hypothesis"), with increased productivity often being the implicit end-goal of a study's intervention. This perspective can risk contributing to modern-day scientific management and algorithmic control of workers [39], issues that this project grapples with directly.

This is in contrast with other studies that place worker perspectives and needs first. Projects like Turkopticon [36] and We Are Dynamo [48] have investigated designing collective action tools for distributed platform workers that focus on worker experience. Projects such as [20] offer systems that may ease the psychological burden that content moderators face at work, without the 'productive' focus. Other work focuses explicitly on strategies that workers can take to increase wages learned from other workers [49], or crowdsourced wage analysis on platforms like Mechanical Turk [33].

This paper continues this line of work attempting to shift the 'default' standpoint of CHI and CSCW research from the perspective of management to that of the worker. Unlike many of the studies mentioned above, this paper documents a *worker-led* research project whose goals were defined by workers themselves, and expands the scope of worker-led research to include delivery workers.

2.3 Outside the Academy

Much work that focuses on building tools for platform workers exists outside the academy, either build by workers or advocates directly, often trying to fill a similar gap to this work. The Driver's Seat Cooperative [3] is a data cooperative that enables gig workers to share and aggregate data on their working conditions, including pay. Deliveroo Unwrapped, a project by the Independent Workers' Union of Great Britain, allows delivery workers to upload invoices to uncover their hourly pay and other details, similar to the Shipt Calculator [2].

There is also a short history of workers using the platforms *themselves* as a tool to coordinate collective action. For example, Seattle-based workers' rights group Working Washington helped coordinate the "Blitz Up" campaign, which coordinated Postmates workers to refuse jobs that did not include bonus (also called "promo" in other apps) pay [45].

The burgeoning field of *Worker Data Science* includes a handful of non-profit and union-driven projects that take aim at more traditional work contexts. The Time Project [1] is a data aggregation initiative for workers in film and TV in the UK that hopes to shed light on the true working hours of film and TV workers. WeClock is a general-purpose data collection tool that unions and organized workers can use to collect reliable data on working conditions through self-surveys and sensor data from phones [7].

The work presented here is unique in calling for researcher-worker collaboration and solidarity, and in its algorithmic audit of a gig work platform. Unlike other projects that have estimated platform worker pay, we present evidence that an algorithmic change impacted workers' pay in a way that would otherwise be invisible.

2.4 Contributions

To our knowledge, this is the first contribution to the CSCW community that documents a worker-led audit of a platform-based delivery app pay algorithm. This paper highlights three main contributions: (1) Explicitly framing the study and tool within the historical standpoint of Digital Workerism and digital organizing within and outside of the CSCW community; (2) The authors' process of facilitating worker inquiry through co-developing a tool with workers and organizers; (3) An analysis of data shared through the tool that audits Shipt's shift to a black-box pay algorithm and its impact on worker pay.

3 BACKGROUND

3.1 Information Asymmetry, Organizing, and Digital Workerism

While many platform-mediated workers (e.g. crowd work, rideshare, and food delivery work that is app-based) are defined in US labor law as independent contractors, scholars and advocates have made clear that the reality is much more complex [11, 32, 44, 47]. Workers are managed and directed through algorithmic and automated management techniques that, as Rosenblat and Stark detail, often produce "the equivalent effects of what most reasonable observers would define as a managed labor force" [47]. They argue that some of the control exerted by these algorithmic management techniques emerges through a general pattern of *information asymmetry*, characterized by withholding information from workers that "would otherwise help them make informed choices about their decisions" [47]. While Rosenblat and Stark focus on patterns within Uber such as information about jobs, surge pricing determinations, and other details, the general problem of *data transparency* at work is much older.

Early proponents of scientific management, or Taylorism, pioneered by Frederick Winslow Taylor in the 20th century, idealized then-novel technologies such as specialized photography [42] and time studies as a way to objectively determine worker wages [24, 29]. As these methods advanced across the management frontier in the US, they faced intense union and worker resistance, eventually leading to the development of union and worker-led strategies to exert some control over the scientific management process, many of which revolved around data transparency and worker self-research. In a historical survey of union manuals from 1947 to 1967, Khovanskaya et al. identify three main ways that unions advocated for workers' rights in the face of scientific management: securing access to the data that employers used to set wages and working conditions, contesting wage-setting by conducting thorough reviews (and counter-reports) of employer-led studies, and approaching "participation" in workplace decision-making (including decisions around time studies) selectively [39]. While they recognize that some of these strategies are unfeasible today, specifically mentioning that labor advocates "will likely need to gather and aggregate data about workers themselves" because the algorithms (and data) used to set dynamic wages are considered the intellectual property of management, how workers can practically execute on such a strategy remains unclear. This gap of data-sharing and analysis, which we further detail below, is one of the core problems in modern digital organizing.

First, access to data about worker experience can be difficult to access or collect. For many platform workers, the data needed to effectively contest dynamic wages and working conditions is collected exclusively by the platforms themselves and flows unilaterally from the worker to the firm. But data such as work history, pay and tip details, GPS traces, and interaction data, common information collected by app-based platforms, could all be considered crucial information for workers seeking to understand their collective conditions. One avenue to gain access to this data is through Data Subject Access Requests (DSARs), a legal mechanism for consumers to request machine-readable exports of data they produce while interacting with platforms [14, 43]. Although generally, scholars have found that DSARs are not adequately accommodated by firms [14], there have been some notable successes in requesting worker data from platform companies. Perhaps the leader in this space with respect to labor is Worker Info Exchange (WIE), a non-profit group that supports platform workers in the EU [8]. WIE regularly facilitates the submission of successful DSARs to platforms on behalf of workers, and was party to a May 2021 EU legal battle that resulted in Ola and Uber being ordered to reveal algorithmic decision-making details and share data requested by WIE on behalf of workers [8, 35].

While DSARs can be an effective way to access data that management and employers already have and collect, they fall short as a singular strategy for data-led organizing efforts. As other design scholars have noted, data is a situated object, and practices around its collection and retention often reflect the position of its steward [25]. This means that the data available through DSAR requests, even in aggregate, invariably reflect the perspective of management and so can be limiting as an organizing tool. For example, in its response to the data access requests in the May 2021 legal case, Uber noted that drivers may not be able to receive full GPS data from their work history out of concerns for end-user privacy (notably, such GPS data is information that drivers can feasibly collect themselves).

The alternative to accessing information about workers from employers through legal means or negotiation is to collect that information independently. Such an approach has the potential to create a more holistic understanding of working conditions: questions asked by e.g. distributing worker surveys can provide a far richer picture than DSARs alone. It also benefits from the flexibility afforded by hand-rolling a collection strategy and research design. Some campaigns may only need a simple survey, while efforts such as auditing an algorithm's fairness may require much more infrastructure. The complexity of modern organizing needs, however, hints at the shortcomings of this approach.

To start, understanding how a managerial change impacted worker well-being or studying the supposed fairness of an algorithm are complicated research problems, and devising a plan may require expertise in research design to execute properly. Beyond research design, *how* to collect the relevant data for a given study is also a difficult problem for even well-organized workers. This is in part due to the increasing technological complexity of low-wage worker management: to quantify the impact of a change in a delivery app payment algorithm, a survey, even if performed longitudinally, is simply not granular enough. Instead, data about workers' GPS locations, their driving habits, records of their tips, and records of their overall pay may be needed to effectively make a convincing case. The challenge of collecting all of this data completely and securely, independently from the platform, is daunting.

Relatedly, whether the data used is from a DSAR or from independent efforts, analyzing and creating rigorous or convincing research outputs from a design requires an increasing amount of technical knowledge. Like in the mid-20th century, when unions struggled to hire or find technically trained staff to design and carry out their own time studies to challenge management [39], it is difficult for modern groups of organized workers to find people or resources who can carry out data collection and analysis of worker conditions rigorously. Even seemingly simple cases can require a surprising amount of technical expertise. For example, in order to challenge a driver deactivation by Uber, Worker Info Exchange analyzed a location dataset (obtained through a DSAR) from a driver's phone and cross-referenced device activity to those location records, a task that requires knowledge of GPS data, geospatial analysis techniques, and data visualization experience.

Successful leverage of collective data for campaigns and publicity point to the importance of collective rights and voice for the future of algorithmically-mediated labor rights. As legal scholar Valerio De Stefano writes, while individuals may be granted rights to access data or contest algorithmic decision-making under various data protection laws, it "could not be sufficient … Individuals should not be left alone to cope with the intricacies of this technology when they want to comprehend and contest the consequences of its applications on them" [21]. Indeed, when considering questions like the one addressed in this paper—how a new payment algorithm instituted by management impacts worker pay—meaningful inferences simply cannot be made in the individual context. Instead, a collective leveraging of data rights is needed.

3.2 Technological Interventions in the Spirit of Digital Workerism

One solution to these challenges is for researchers to bridge this technology gap by embarking on research projects and agendas that fulfill crucial questions important to worker experience. Computationally competent researchers can co-design inquiry through adopting methods from participatory research and strategies from the history of worker inquiry [12, 58], dating from the mid-1960s. Participatory research and action research [15, 30, 57] provide tools that guide researchers seeking to center and create research questions on the basis of subjects' experience, while worker inquiry provides a lens of organizing, coalition and power-building, and historical context that may help guide research agendas [17, 18, 60].

First, a note on the history of worker inquiry and workerism. Started in the 1960s in Italy as a response to radically changing working conditions, worker inquiry was one part of the *Operaist* (translated to "workerist"), project that focused on exploring workers' experience of work and workplaces [17, 60]. While developed under the same roof, the perspectives of its authors are fragmented, with competing political perspectives and epistemologies [60]. Strategies for inquiry ranged from approaches based on the modern sociology and social science of the time to analyses of class relations and general studies of production processes. The goals of inquiry were under a constant self-critique, with perspectives that ranged from valuing basic knowledge of worker experience [46] to ideas of research as a political tool for coalition-building [18].

While the inquiry of this era focused sometimes myopically on analyses of class relations or general studies of production processes, the notion of a worker-driven research agenda, as Carmichael [18] points out, has particular salience in the contemporary age of rapid workplace change driven by technology, where unilateral decisions made by technology designers can impact millions of workers. The same is true of the general workerist strategy of transforming "hot inquiry"—a study of a specific workplace question or struggle—into a longer-term practice of *co-research*, which might contribute toward a "lasting disposition towards inquiry or research among workers" [18]. Such a disposition has the potential to build a more empowered and capable worker movement.

However, in the digital era where inquiry responsive to algorithmic management may require more complex systems of tools for data collection, analysis, and stewardship, the prospect of developing a "longer-term practice" takes on a new meaning. Building "capacity" no longer means only developing a disposition or a politics based on a "right to research" [13], but also on the real development of concrete tools that make research more attainable, actionable, and accessible to workers. This strategy of leveraging digital tools for worker-driven inquiry has been described as a kind of "Digital Workerism" or "Digital Worker Inquiry": a research practice enabled by modern technology that seeks to simultaneously answer research questions and build a capacity for future digital organizing [9, 26, 59].

As Karen Gregory notes, this practice raises concerns about the power relations between workers and researchers [31], a conversation echoed in the participatory design literature that if left unaddressed may have serious downstream consequences. Rather than simply applying participatory methods in digital social science research to the workplace context, Digital Workerism must concern itself with the standpoint of the research itself. It involves two main pillars. Firstly, Digital Workerist research must arrive from a worker-driven perspective [59]. The "hot inquiry" questions addressed by a Digital Workerist approach are those that address core worker experience, regardless of their relation to algorithms or digital technology at work: "is the driver ratings system biased against black workers?"; "Is everyone getting an appropriate amount of break time?", and should be questions that are devised and asked by workers themselves. Second, a Digital Workerist approach should aim to produce a lasting *capacity* to repeat this inquiry or move beyond the individual questions themselves. In building digital tools, such a capacity can arrive both as a longer-term organizing or research agenda, but also in the form of sustainable, open-source software and tools that are usable by organizers and workers without direct researcher involvement. Such an approach has the potential to both answer crucial questions about modern working conditions as well as develop workers' ability to ask and answer such questions independently in the future.

4 DESIGNING THE SHIPT CALCULATOR

In this section, we use the frame of Digital Workerism to situate the design of the Shipt Calculator as a tool both to solve a specific "hot inquiry" problem—wage transparency and algorithmic pay auditing for a specific group of platform-based delivery workers—and as a system that may help "build capacity" for similar future campaigns through open-source design and an easily-deployable infastructure. We start with an overview of the problem that the Calculator was designed to solve, discuss its design process, including how relevant stakeholders such as workers and worker advocates were involved, and describe its final design.

4.1 The Problem

The Calculator started as a partial solution to a "hot inquiry" problem that a group of platform-based delivery workers faced in the Summer to Fall of 2020. Shipt, a major app-based grocery delivery firm purchased by Target in 2017, employed a transparent and simple algorithm to pay its workers before the Spring of 2020. Workers were paid \$5 per order completed, plus a commission of 7.5% of an order total. For example, if a worker accepted an order for \$50 worth of goods from Target, they would be paid \$3.75 commission, plus a \$5 base pay, for a total of \$8.75, before customer tip. In the Spring of 2020, Shipt started testing an opaque new algorithm (dubbed 'V2') for worker pay that many workers reported as effectively reducing their pay. In press releases, Shipt claimed that their new algorithm, based on a number of variables such as order amount, mileage driven, and

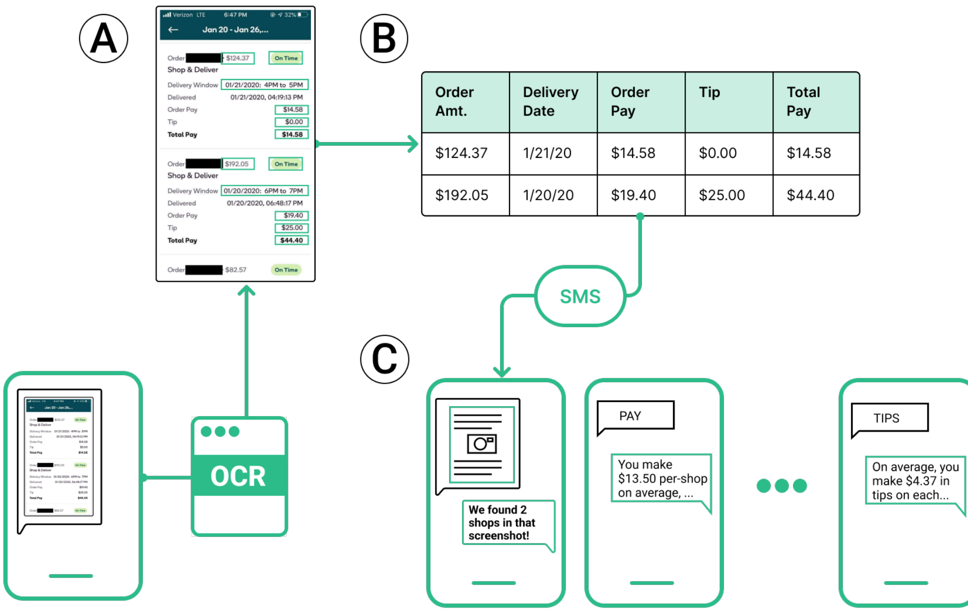

Fig. 1. Illustration of how the bot worked. Workers would send a screenshot (A) of their pay history, and the Calculator would use OCR to parse out relevant pieces of data and store it in a structured way (B). Workers then received a validation text informing them that the bot had parsed their screenshots. Workers could then use various commands to learn more about their pay (C).

time-to-shop in stores, would be more fair to workers and provide pay more commensurate with an order's effort than their previous approach. The complexity, variability, and opaque nature of this new algorithm, however, made it impossible for these claims to be verifiable by workers.

4.2 Early Approach by Organizers

In the Summer of 2020, researchers were introduced to worker-organizers and a non-profit worker organization, Coworker.org [4], who had begun an initiative to measure how the new 'V2' algorithm impacted workers. The Shipt Shopper app allows workers to view their work histories, but does not offer a way to export this information. In order to collect pay data from workers, organizers had started soliciting screenshots of worker job histories from the app, an example of which is shown in Figure 1A. Workers would text, email, or otherwise share dozens of screenshots of their pay history with organizers. Each screenshot shows multiple orders that a worker completed. Organizers would then manually transcribe each order into a spreadsheet, recording details such as order base pay, order tip, and the date of the order. Because the earlier pay structure was transparent, organizers marked an order as paid using the 'V2' system if its base pay differed from the $5 base pay plus 7.5% commission scheme.

While this approach worked for initial organizing efforts, it had two main shortcomings. First, manually transcribing details for each order was time consuming. Many shoppers had hundreds of orders they had completed that needed to be transcribed, and manually entering the details for each order took many hours of organizer time. Second, while workers could donate their pay data to the effort, it was logistically infeasible to give workers any individual feedback about their pay, such as whether their pay had changed or the percentage of their pay that came from tips.

4.3 Technical Approach and Research Questions

The first steps we took with organizers was to collaborate on both a general technical approach to scale the system and a set of broad research questions that could guide further development and design decisions.

4.3.1 Deciding a Technical Approach and Design Goals. To demonstrate a technical approach that could help organizers, we first prototyped an OCR script that parsed worker screenshots into a spreadsheet using the Tesseract OCR engine in Python [10, 40]. After presenting the prototype, we proposed two main technical approaches: building a website where workers could upload their screenshots, or building an SMS texting bot. Organizers overwhelmingly preferred the texting bot because of accessibility and simplicity. Many platform workers do not own a desktop computer, so any solution had to be easily accessible from a mobile device. In addition, organizers were interested in maintaining the tool over time, and developing an SMS bot meant less developer overhead on designing a website, displaying information reliably on different devices, and onboarding for workers interested in participating and donating data. We then developed a proof-of-concept prototype with Twilio [6] that could parse individual screenshots sent over MMS.

4.3.2 Defining Research Questions. After deciding on a technical approach and making our proof-of-concept prototype, we held several open discussions over Zoom with organizers to decide on research questions to focus on. Researchers contributed knowledge on what was understandable and computable using the data we could collect, while organizers offered contextual understanding of what information workers needed most. Between meetings, organizers would hold one-on-one conversations with workers about their concerns regarding the pay change and would voice those questions in our meetings. We decided on five main research questions through consensus with organizers with the goal of answering them using data collected through the tool. The questions and some motivation for each (informed by organizers) are below:

- **RQ1**: How was the V2 algorithm rolled out to workers? Was it consistent?
- **RQ2:** Is there a detectable difference in pay between the V1 and V2 algorithm? If so, what is the degree of difference?
- **RQ3**: How many workers have been impacted by the algorithm change?
- **RQ4**: Does the new pay algorithm impact some workers more than others?
- **RQ5:** What is the average hourly pay of workers, before expenses? How does the V2 algorithm impact this hourly pay?

RQ1 is important because an understanding of how V2 rolled out may help organizers and workers understand how future pay and feature changes might spread across the workforce. Answering **RQ2** would bring clarity to many workers' perceptions that the V2 algorithm had resulted in lower pay overall, and information about any difference (particularly if V2 paid out lower than V1) could be used by organizers in campaigns to mobilize workers. **RQ3** would provide a sense of scale that could be used to make estimates about the broader Shipt worker population. **RQ5** is a "holy grail" of gig work research and organizing inquiry—hourly pay is notoriously difficult to calculate or collect for gig workers because of the sporadic nature of the work and algorithm design. Understanding how V2 might impact hourly wages can also provide a benchmark that helps the public understand the conditions of gig workers generally.

4.4 Design Goals

From these early conversations, we converged on a small set of design goals for the tool, which we also decided on via consensus:

1. The system should be easy to use by non-technically savvy workers.

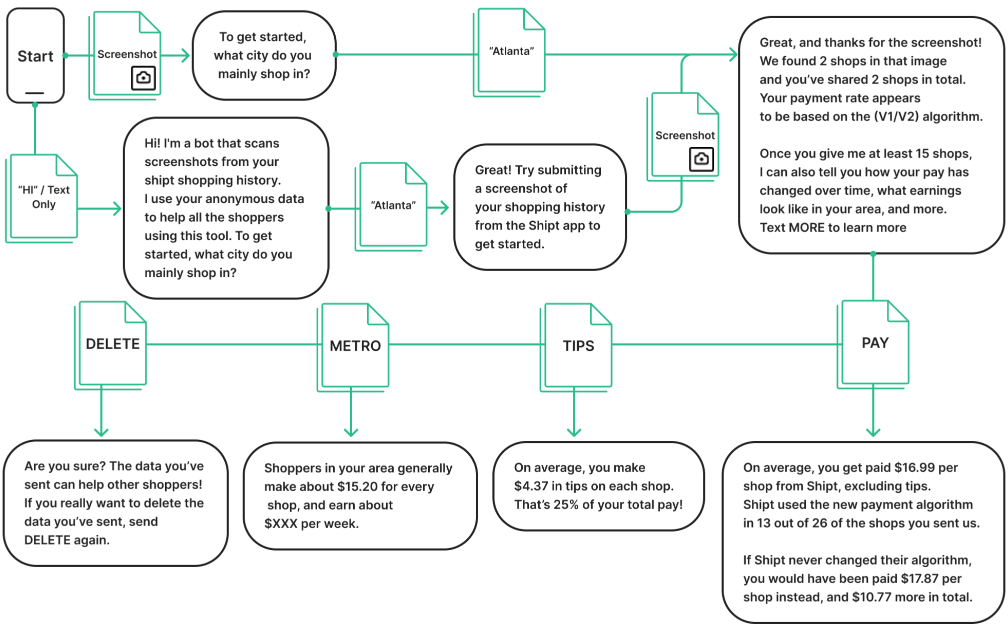

Fig. 2. The texting flow co-designed with organizers. Organizers were given direct access to the working flow diagram through Google Slides to enable co-editing of the bots' script. Blocks represent pieces of text sent to workers, while lines between blocks represent commands or responses a worker might send to the bot. Not shown is a response from the bot detailing what each command does, the result of the 'CONTACT' command, which shared contact info of organizers and researchers, and the 'INFO' command, which summarised the goals of the project.

2. Workers should receive some useful information in return for contributing data as "compensation" for contributing and to incentivize participation.
3. The system should allow workers to delete and download their own data securely.
4. The system should operate entirely over SMS/MMS for simplicity.

To make sure the tool met all these goals, we collaborated on the user flow that workers would move through while interacting with the bot, including the order of commands and specific wording and text sent to workers. This flow then informed the technical development and "features" that needed to be added to the system. To do this, we created a shared deck in Google Slides that organizers and researchers could edit and comment on. Researchers created the first version of the flow, which was then commented on and edited by organizers. The final texting flow is shown in Figure 2, with final organizer edits in green to demonstrate the co-design process.

We did not perform any rigorous user studies as part of this project, instead opting to let organizers share the bot with workers and collect informal feedback. To ensure our design accomplished goal (1), we ensured that the bot would clearly describe its purpose and what workers had to do in order to contribute their pay information. We added three commands to the bot for design goal (2): "PAY", "METRO", and "TIPS". Each of these commands would provide some information to a worker based on the data they contributed, highlighting any changes due to the 'V2' algorithm that could be identified. Table 1 shows some example text returned by the bot in response to these commands.

Table 1. Shipt Calculator commands available to workers.

Command	Description	Example Response
METRO	Compare pay to workers in same metro area	Shoppers in your area generally make about 15.20 for every shop, and earn about 423 per week.
PAY	Pay and algorithm information	On average, you get paid 16.99 per shop from Shipt, excluding tips. Shipt used the new payment algorithm in 13 out of the 26 shops you sent us. If Shipt never changed their algorithm, you would have been paid 17.87 per shop instead, and 10.77 more in total.
TIPS	Average tip and percentage of total pay in tips	On average, you make 4.37 in tips on each shop. That's 25 percent of your total pay!

After contributing at least ten shops, the "TIPS" and "PAY" keywords would calculate and send the workers' average pay and tip statistics, including what percentage of their shops we inferred as being paid by the new black-box 'V2' algorithm. If a worker shared that they worked in a metro area that at least ten other workers also registered under, the "METRO" keyword would share average pay per-order and per-week for all shoppers in that metro area. Metro areas shared by workers were normalized using the Google geocoding service.

For design goal (3), we included an "EXPORT" keyword for workers which would generate a temporary link that allowed a worker to download their parsed pay data in a CSV format.

Importantly, organizers were also given access to an 'export' command with an associated password that changed daily, created by researchers. This command would export aggregate information from all workers that organizers could use for their own campaigning and efforts, allowing both researchers and organizers to access and analyze the data. Full data access was restricted to researchers.

4.5 Scaling and Iterating

To test the reliability of the OCR and texting bot, we used a collection of example screenshots from the initial organizing efforts to benchmark the system. This ensured that edge cases such as partially-cropped jobs, jobs with special "promo" pay, and details such as delivery time and window were all parsed correctly.

After collecting several hundred screenshots, it became clear that using Google Sheets as a backend for data storage would not scale to the thousands of screenshots and orders needed to audit workers' pay changes. After discussing this problem with organizers, we moved to storing worker data in Firebase, a real-time unstructured database hosted by Google. After moving backends, the data shared by workers was no longer directly visible to organizers, so we developed another keyword, with an associated, developer-generated passphrase, that would allow exports of all worker data for organizers. Importantly, the link given to organizers to export data only lasted for a few hours, reducing the chance of it being shared and used outside of the core organizing team.

We made several other iterative changes to the bot. While testing, we would receive feedback from some workers with varying device details about their pay being parsed incorrectly. Researchers investigated each one of these reports and adjusted the OCR algorithm to accommodate these

Table 2. Sample data collected by the bot by scanning shopper screenshots. 'After rollout' indicates if the delivery date was after the official V2 rollout. 'Alg. Version' is inferred, and is equal to 'V1' if the Order Pay matched what a worker would have been paid under the V1 algorithm. Shopper IDs randomly assigned here for privacy.

Shopper ID	Order Total	Order Pay	Tip	Delivery Date	Alg. Version
3910	$170.03	$17.75	$0.00	2020-09-12	V1
2272	$68.09	$10.11	$10.00	2020-08-23	V1
2011	$115.39	$13.65	$21.76	2020-08-23	V1
1755	$187.55	$14.00	$21.05	2020-09-28	V2
2682	$317.30	$25.52	$70.98	2020-08-29	V2

cases. Early exports that workers could download were simple exports of the database back-end translated into a CSV file. After hearing from some workers that they had trouble understanding the exported data, we simplified the formatting of exports to be more easily understood.

4.6 Incentivizing Participation and Recruitment

Early testers were solicited through online forums by organizers and through word-of-mouth recruitment. When organizers and researchers agreed the tool was sufficiently tested, a single phone number was created for the bot that could be shared widely with workers and organizers in mid-July 2020. We relied on organizers for recruitment, as it was important for organizers that the tool be framed as an organizing effort as well as a research tool. Organizers posted about the bot on social media channels and in online worker forums.

5 THE STUDY AND RESULTS

5.1 Data and Methods

By October 2020, the Shipt Calculator had received a total of 5684 orders from 201 workers. To create the final dataset we use for analysis, we perform some basic filtering and data checks. First, the OCR tool and script we developed to parse screenshots [10] made several mistakes in parsing order pay amounts. For example, significant digits were occasionally dropped in parsing, turning e.g. a $16 order into a $6 one. To partially account for this, we enforced data consistency by including only orders where the total pay was consistent with the recorded tip and base order pay. We also exclude orders with a payout over $55 before tips, for a similar reason, and any orders with an order total (the amount a customer purchased through Shipt) over $500. Finally, we include only workers who submitted at least 10 orders to the tool. These filters result in a final dataset of 5271 orders from 140 workers.

Although screenshots shared by workers include many details about each order—a full detail of fields that the final version of the Calculator parsed can be found in Table 2—the Shipt Shopper app does not indicate to workers whether an order was paid out under the old payment scheme ('V1') or the new, 'V2' black-box algorithm. To infer which worker orders were paid out under this new scheme, we computed the counterfactual order pay—that is, what the pay would have been under the old V1 scheme—and compared it to the actual pay a worker received. We'd like to clarify briefly that because we are not simply comparing aggregate pay between orders paid via V1 vs orders paid via V2, any differences we detect are not due to any pattern in deployment that Shipt may have taken or some hidden covariate in our worker population. Instead, it is simply the difference between what a worker was paid under V2, and what they would have been paid under V1.

Table 3. Summary statistics of 'real' orders paid through the different algorithm versions, excluding counterfactual analyses.

Alg. Version	# Shops	Mean Pay	Pay Std.	Median Pay	Mean Tip
V1	2152	$12.36	$4.80	$10.98	$10.17
V2	3119	$13.58	$5.15	$12.23	$10.48

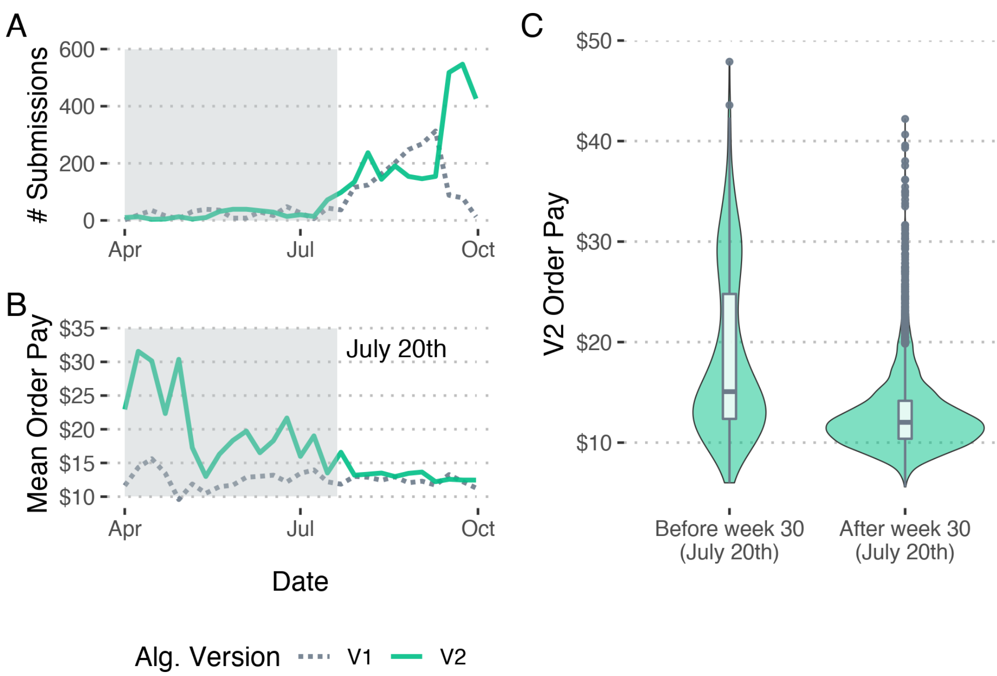

Fig. 3. Examining order pay variability. Figure A shows the number of total submissions over time by algorithm type. Figure B shows the mean order pay across all orders for each algorithm type. Order pay stabilized significantly after week 30 (July 20th), the week the tool was first launched. In both Figures A and B, the gray area indicates weeks before week 30 which were not used in analysis, and the black vertical line indicates when Shipt officially moved to the V2 algorithm. Figure C shows the difference in variability in order pay before and after week 30.

This counterfactual pay was computed using the algorithm that Shipt had previously published: $5 base pay per order, plus a 7.5% commission on the total order amount. Any order with an observed payout different from this 'counterfactual' V1 pay was labeled as a V2 order. Using this method, we identify a total of 2152 orders that were paid using the V1 algorithm, and 3119 orders paid using V2. Table 2 shows a sampling of 5 rows from our final dataset with labels identifying the pay algorithm version, while 3 shows some summary statistics of orders paid out using the different algorithms. While the mean and median pay are similar, pay under V2 appears to vary more than under V1.

We only use orders that occurred after week 30. Per-order pay became much more stable after this week, in part due to the increase in submissions beginning around week 28 (mid July), when the tool was first launched. Figure 3A and 3B show the number of submissions and average base order pay over time by algorithm, respectively, demonstrating that order pay became much more stable after the 30th.

Workers also reported their "metro" area to the tool on sign-up. While some included state names (e.g. "birmingham al"), others reported only their city ("mesa"). We assigned states to each worker by using the Google Geocoding API to match reported metro areas to states. Three workers reported metros that were un-matchable. While we used these workers' pay data, their information is excluded from analyses such as those in 5.2.4 that require an inferred work location.

5.2 Results

Once the basic infrastructure had been tweaked and deployed, workers were recruited to participate in the program through word-of-mouth (snowball) recruitment and posts by organizers on platform worker forums. Once 100 workers had contributed data to the platform (a heuristic threshold we set), we began planning and executing our analysis. The specific research questions co-designed with organizers are listed below:

- **RQ1**: How was the V2 algorithm rolled out to workers? Was it consistent?
- **RQ2:** Is there a detectable difference in pay between the V1 and V2 algorithms? If so, what is the degree of difference?
- **RQ3**: How many workers have been impacted by the algorithm change?
- **RQ4**: Does the new pay algorithm impact some workers more than others?
- **RQ5:** What is the average hourly pay of workers, before expenses? How does the V2 algorithm impact this hourly pay?

5.2.1 The Rollout of the Black Box (RQ1). Figure 4C shows a sampling of worker-submissions over time. While most workers experienced a singular transition point where their pay migrated from the V1 scheme to the new V2 algorithm, when this change occurred was unpredictable. Some workers had been paid using a new scheme as early as February, while others were still being paid using the V1 algorithm as late as mid-October, days before the official rollout went into effect on October 15. Figure 4B shows the average percentage of orders in each week that were paid through V1 or V2, averaged over 2-week periods. By June, over 50% of orders on average were paid through the V2 algorithm. We could not find a consistent pattern in the order of the V2 rollout across shoppers.

5.2.2 Measuring the Impact of the 'V2' Algorithm (RQ2 & RQ3). To compare V2 and V1 pay, we simply take all orders identified as paid under V2, and then calculate what that worker would have been paid under the V1 scheme—for more information on this 'counterfactual' dataset, see Section 5.1. Figure 5A shows the change in average pay per-order that each worker experienced under the new V2 algorithm. Overall, the average workers' pay increased under V2: the new algorithm paid $0.57 more on average for each worker, a 8% increase in order pay from what workers would have been paid under V1.

This is explained through disaggregating the data. When disaggregated, this pay change is not distributed equally across all workers. Figure 5A shows a binned distribution of each workers' average pay change from V1 to V2. While many workers saw an increase in their average order pay, 42% of workers effectively received a pay cut. Half of all workers who made less under V2 received a pay cut of at least 10% on average.

5.2.3 The Change Impacted Some Workers Unevenly (RQ4). Do workers that received a pay cut overall consistently get paid less under V2? Figure 5B shows the average dollar difference in order pay for workers for each week from August through October. Workers paid less on average under V2 were consistently paid less each week, rather than having some weeks where V2 paid more and

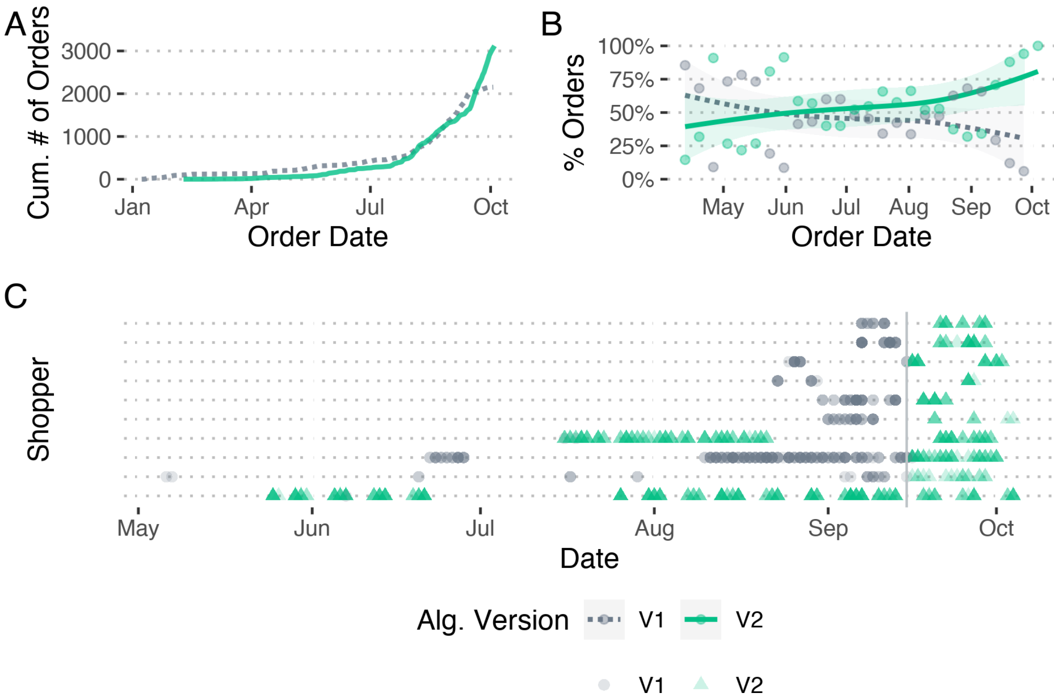

Fig. 4. Algorithm prevalence over time. Figure A shows the cumulative number of orders by participants over time. By mid-July, workers had received more total V2 orders than V1, and the transition to V2 was almost completely finished by October. Figure B shows the percentage of orders that are of either algorithm, averaged over a 2-week window. By June, workers were receiving pay under the V2 algorithm more often than V1 on average. Figure C shows submissions over time for a sampling of 10 workers. While some workers had been recording pay under the new V2 algorithm since February, others' pay algorithms did not change until Sept. 15 (vertical line), when the official rollout of V2 began. Some workers also saw intermittent V1 or V2 jobs in their workstreams, instead of a clean break from one payment scheme to another.

some weeks where it paid less. This suggests that there is something about the V2 algorithm that punished certain shoppers over others. Unfortunately, our research design limits the inferences we can make on this point. We find no pattern between overall pay change and location (state) the only other piece of information on workers we collect besides pay.

5.2.4 Shifts and Hourly Pay (RQ5). Using the delivery times recorded on worker pay stubs, we estimate the average time taken to complete an order, and use that information to estimate workers' average hourly pay. For each worker, we group their orders by day, and define a 'shift' for a day as the stretch of time from their first delivery to their last delivery of the day. This is a rough approximation of worker behavior, because workers can take breaks between jobs or work other apps, so these estimates are necessarily conservative, offering an upper-bound on the number of hours a worker may work in a single day. Calculating shifts this way results in an average shift length of 8.2 hours with a standard deviation of 2.5 hours: the workers contributing data to this study are, on average, working shifts of about eight hours. Figure 6A shows a visualization of worker shifts, ordered by start time, across our dataset. A large majority of shifts start between 8 A.M and 12 P.M, and last until at least 4 P.M, similar to a normal workday.

Fig. 5. Impact of new V2 algorithm on worker pay. Figure A bins workers by the average pay change they experienced per-order when moving from V1 to V2. While many workers (almost one third) experienced a pay increase on average, 41 percent of workers have experienced some form of pay cut in the transition to the V2 algorithm. Figure B shows the average pay difference across workers over time, revealing that the pay change through the V2 algorithm appears to be consistent across workers. Workers who effectively had their pay cut under V2 are consistently making less under the new regime, while workers who are making more on average are consistently making more.

By calculating the number of orders a worker completes during each shift, we estimate the average amount of time a worker spends on each order, and their corresponding hourly pay using their average order pay under each algorithm. Figure 6B shows the distribution of average order length across workers, revealing that on average, it takes a worker 1.1 hours to complete an order, with a wide distribution ranging from as short as half an hour to as long as two hours. This average time means that for many workers, the base pay they receive for a given order is their most recent hourly base wage. Figure 6C shows a box plot of worker hourly pay across our two groups of interest: workers who received a pay cut and those who did not. This verifies that workers receiving a pay cut under V2 saw a hit to their effective hourly wages.

Figure 7 shows how this hourly pay compares to the minimum wage laws in each shopper's state as inferred from their metro location (see section 5.1). Dashed lines indicate the minimum wage for that state in 2020 as reported by the Department of Labor, while each boxplot represents the distribution of calculated hourly wages by algorithm. We only include states with at least ten reported "shifts" (days of work) from at least 5 different shoppers.

Shipt's base pay often results in an effective wage lower than state minimum wage laws in some states. Pay distributions in Washington, California, New York, Minnesota, and Ohio all show some shifts where workers were paid under minimum wage. At times, the pay from Shipt is far below

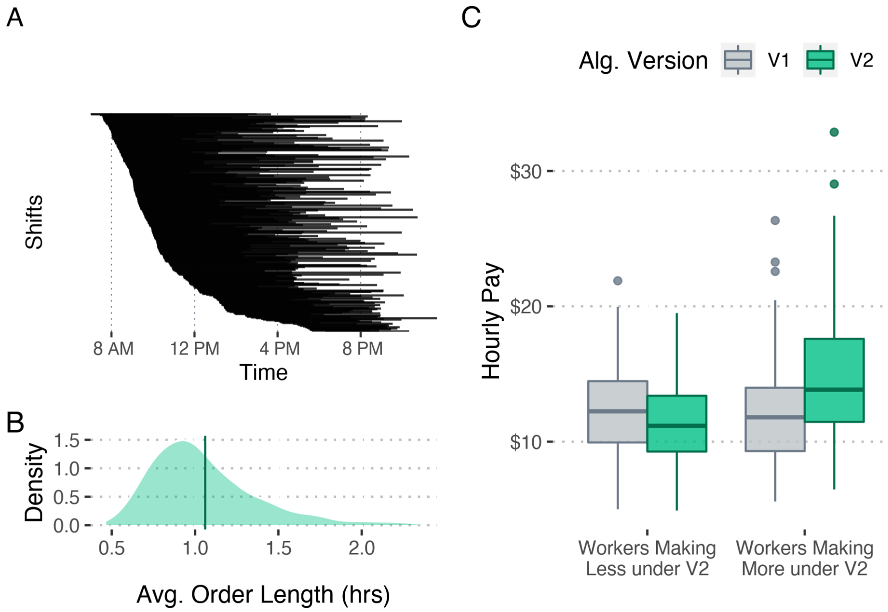

Fig. 6. Worker shift patterns and hourly pay. Shifts are calculated as the number of hours between the first delivery time and the last delivery time of a given day. Figure A shows example shifts, represented by solid lines, taken by workers in our dataset, sorted by start time. Figure B shows the distribution of average time taken to complete an order for each shopper, calculated by averaging the number of orders completed in a shift divided by its length in hours. The line represents the mean order duration across all workers. Based on these metrics, Figure C shows the distribution of hourly pay workers make before tips. The solid and dashed lines shows the median hourly pay under the V1 algorithm and the V2 algorithm, respectively.

even already-low state minimum wage statutes. For example, in Ohio, where the minimum wage was $8.70 in 2020, the median hourly pay under V2 was $7.57 . In total, across all states, 35% shifts we calculated resulted in an hourly pay under minimum wage.

Also interesting is the wide variability in pay distributions both between and within states. For example, hourly pay in Florida ranges from under $10 per hour to over $20 per hour under V1, an enormous range. Variability in pay is one important condition in defining economic precarity. While Shipt's pay in Florida appears to reliably stay above the state minimum wage, this wide variability likely makes it difficult for workers to plain their financial futures.

Interestingly, states with similar minimum wage laws, like California and New York or Hawaii and Minnesota, do not demonstrate similarity in pay distribution. While reverse engineering the V2 algorithm is beyond the scope of this paper, it does hint that Shipt's new payment model is not dependent on minimum wage laws. Similarly, states with very different costs of living, such as New York and Minnesota, do not have largely different pay distributions. Surprisingly, the lower end of the hourly wage in New York is much lower than in Minnesota, even though cost of living in New York is much higher.

5.3 Tool Usage

Although we did not run any user studies, we do have some general usage data and metadata from workers' submissions to the texting bot. We did not record logs from the texting service detailing users' requests for statistics like their pay or tip analyses. In this section, we leverage this

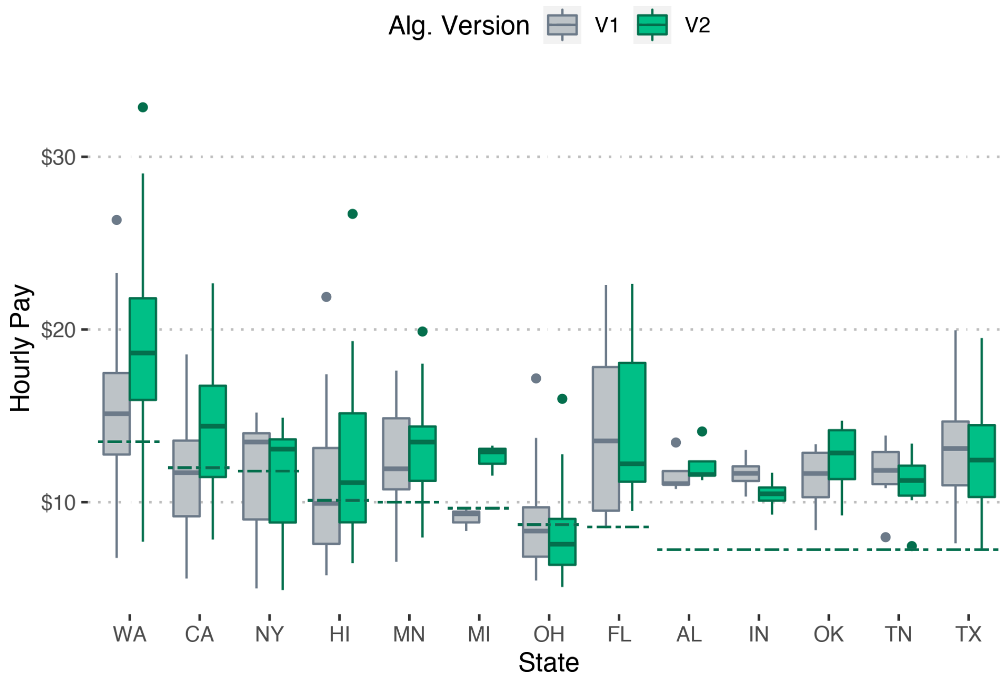

Fig. 7. Worker effective hourly wage (excluding tips) by state and algorithm version. Dashed lines represent the state minimum wage in 2020. States are ordered by their minimum wage. States with fewer than 10 shifts or fewer than 5 shoppers are not shown.

information to gleam some general insights into how the tool was used and derive some insights that might be useful for future projects like the Shipt Calculator.

Many workers who submitted screenshots to the tool submitted in one large 'dump'—47% of workers submitted all their orders to the bot in a single day, rather than across several days. Figure 8B shows the distribution of the number of unique days users interacted with the tool. This makes some sense, as while sending images to a texting service is relatively low-maintenance, it is easier to do in large batches (which the tool supported). This suggests that when building a tool like the Calculator, designers should expect a large chunk of workers to interact with it a single time. Workers who submitted all their orders in a single day did not submit significantly different numbers of orders than other workers, however.

We do not have saved data on the requests for analysis (e.g. the "TIPS", "PAY", or "METRO" commands) that workers submitted, but any future work following this tool's design should strongly consider saving this data and analyzing it to estimate capacity for long-term usage. While over 50% of workers interacted with the tool over at least two days, the fact that so many workers only submitted screenshots on one day means it is unclear whether workers found the feedback beneficial enough to use it long-term.

On average, workers submitted orders as far back as 45 days before the day they interacted with the tool, with some workers sharing pay from as far back as 8 months. One to two months may be the natural amount of time that these workers are comfortable sharing their pay data, or it may simply feel like a natural span from which workers thought their pay patterns might be uncovered. Figure 8A plots the number of unique days each worker worked against the span of their work

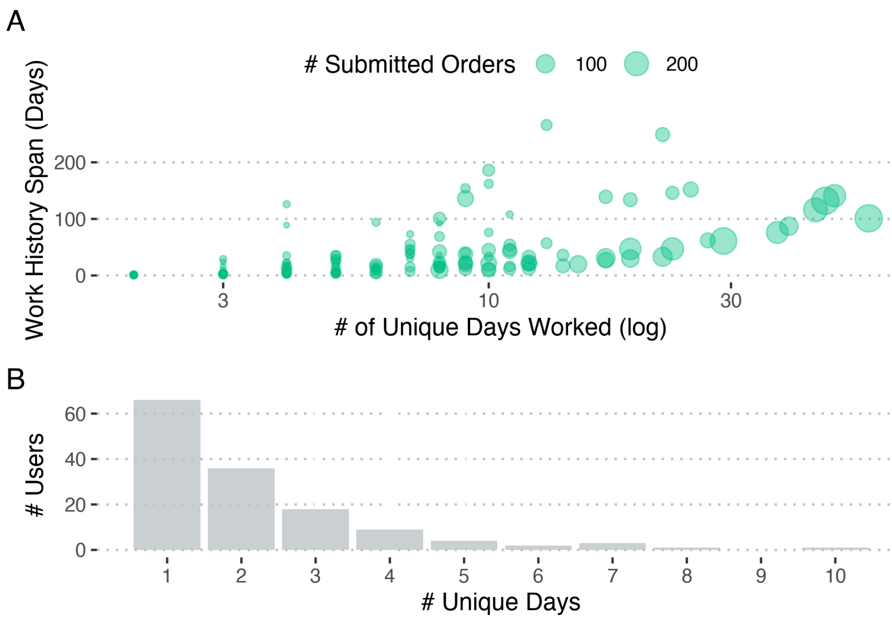

Fig. 8. Tool usage data. Figure A shows the number of days between a worker's first and last order they submitted plotted against the total number of unique days they delivered at least one order. For example, if a worker worked half of the days in a 100-day range, they would have worked 50 unique days. Size of each point indicates how many total orders they submitted to the tool. Figure B shows the total number of unique days each user submitted orders using the tool. Most workers submitted orders over the course of 4 days or fewer.

history, showing that in general, workers who worked more days did contribute a history that spanned more time, although this was not always the case.

6 DISCUSSION

In this paper, we present the Shipt Calculator, a tool co-designed with workers and organizers to aggregate and analyze gig worker pay. We share some design lessons from developing the tool, and present an analysis of how Shipt's shift to a black-box pay algorithm impacted worker pay.

As platform-based gig work rises, so do the potential harms of algorithmically driven management and pay structures that create opaque working conditions. As others have argued [31], the gig economy is a bellwether at the forefront of managerial and employer-driven data science and control. However, this also makes it a fruitful testing ground for experiments in collective data access, the future of worker-driven technologies, and the potential for academic collaborations with workers more generally, extending the *Workerist* spirit into the 21st century and further into the academy.

Here, we demonstrate that collective action through data can be used to audit an algorithmic change in gig worker pay. The nuanced results of our analysis—that some workers received a pay cut while others saw a raise, that the effective wage of workers is so varied—points not only to the poor working conditions of gig work and the one-sided nature of control in the field, but also at the importance of transparency in the gig work marketplace. Workers who participated in this project were able to see the reality of their pay only through a months-long collaboration between a computationally competent researcher and a non-profit organizing group. Workers who are participating as "independent contractors" in an open marketplace should not have to go to such lengths to understand how their pay may change between apps, or indeed, between weeks.

6.1 The Importance of Data Access

Our analysis also points to the limitations of dynamic pricing for workers. 29% of workers in our analysis worked at least one shift below their state's minimum wage. While Shipt workers are tipped, this is not always a guarantee across the platform work landscape. Doordash, for example, notoriously skimmed tips to meet minimum wage guarantees for an unknown amount of time before workers caught on, prompting a class-action lawsuit [38]. In addition, 26% of jobs collected by our tool were entirely un-tipped, leaving many workers with only their base order pay for many jobs. Collective data access and review could be a crucial future affordance for regulators and policymakers seeking to understand the state of platform work more generally.

Data access for platform workers is also a larger project than just "bargaining with the black box" for higher wages or investigating work experience. Collecting and aggregating screenshots of worker pay only scratches the surface of the vast amount of data that platforms collect from workers in the course of their daily work. Platforms such as Shipt deal in what some scholars describe as "dual value production" [23]: while platforms ostensibly produce value through the services they provide (like delivery), they also generate profit through speculation on the value of the data they hold. Workers accessing and leveraging the data they produce at work for their own goals is then also an implicit exercise in interrogating the kinds of value they add to the firm. Projects like Driver's Seat, which seek to create data products co-owned by workers that exchange their donated data for future dividends, are experiments in this realm, but more experiments and collaborations are needed to explore how workers can share in the value their data creates while at work.

6.2 Data Stewardship and Governance

Data stewardship and governance is another important part of the Digital Workerist landscape. These new paradigms also call for new ways of managing and stewarding worker data. The data collected by our tool is sensitive: for some workers who participated, Shipt was their only income source, potentially allowing the orchestrator of the tool to estimate their full income and work history if phone numbers or other identifiers were stored and made available. While much of the data donated by workers in academic projects like the Shipt Calculator will be governed under Institutional Review Board agreements with strict consent rules, when these projects leave the realm of research and are, as Digital Workerism hopes, adopted by workers and organizers themselves, this will no longer be the case.

Responsible participation in the Digital Workerist movement will mean developing tools and frameworks for organizers and workers to govern how datasets like the one analyzed here are used and stored. Legal mechanisms, like the cooperative structure of the Driver's Seat Cooperative, are one solution, although a cooperative structure implies a requirement of dividends to worker-owners, necessitating a profit model from worker data. Data trusts are another solution, with several proposals in recent years addressing how bottom-up structures may be used to create responsible data stewardship models [22].

6.3 Capacity-Building and Future Use

One of the goals for this project was to "build capacity" for worker-organizers and advocates to collect their own data and run their own analyses using similar methods. At the time of writing, the Calculator is under active development to expand it's application to other platforms, but has not been adopted for use by any worker groups without researcher involvement. Deploying the tool and analyzing the resulting data is not yet an automated process; having a computationally savvy team member is a large obstacle for worker groups who would like to use the Calculator for their own campaigns. In this way, researcher-driven analysis and reporting is a major weakness of this work. Future work should consider how to create low or no-code data analysis tools that organizers can use to generate their own reports from their own campaigns.

We also do not study long-term usage of the tool, which we view as room for future work. Interesting future avenues include expanding the tool to enable worker groups to create their own campaigns and data collection strategies while studying how campaigns and workers use the tool over time. Such a study might also investigate more thoroughly how different incentive mechanisms impact worker participation. While we offer some customized statistics to incentivize workers to contribute data, we do not study or test their effectiveness in soliciting participation or motivating collective action.

6.4 Limitations

There are many limitations to our study that we should acknowledge here. Future work should explore more deeply the differences in design approaches between participatory design and worker co-research, and derive and share some principles for co-designing technological tools with sensitive worker populations. Second, the wage study we present is also limited. Shipt is a large platform, and the number of workers surveyed in this work is miniscule compared to their broader workforce. While we are able to accurately report pay differences between the V1 and V2 algorithms for the workers we studied, the hourly wage estimates reported here are rough approximations due to the limitations of worker invoices and the scale of our data. Future work could report broader results from more platforms and locales, and devise a more rigorous approach to estimating hourly wage that includes expenses and differences in working schedules. Finally, while we speculate about the value of data governance for workers, the project presented here uses a traditional data collection strategy; the data collected by the Calculator is not "owned" by workers. Future work should investigate ways of designing technologies that allow workers more full control over their data and how it is used.

7 CONCLUSION

This paper presents the design and deployment of the Shipt Calculator, a tool designed in collaboration with workers and organizers that aggregates platform worker pay and provides some wage transparency to workers. We then discuss using the Calculator to audit an algorithmic pay change made by Shipt in the Summer to Fall of 2020, finding t hat the change resulted in an unannounced pay cut for 41% of participating workers, and find that many workers work shifts with pay under their state's minimum wage. We present the Calculator as a kind of case study for a modern form of *Digital Workerism*, a movement introduced by other scholars that we argue should include building concrete technological tools co-designed with workers with the goal of creating a capacity for future worker inquiry and research. While the Calculator presents useful findings for a specific group of workers, it is one part of a broader effort to empower workers through digital technology and their data.

ACKNOWLEDGMENTS

Thank you to Willy and Vanessa for tirelessly organizing and being open to worker-led inquiry. We would also like to thank Amos Toh and Lena Simet for comments and direction on this work, Christina Colclough, Jonnie Penn, and Nathan Freitas for their support and helpful comments, and the reviewers for their helpful remarks that led to a greatly improved version of this paper.

REFERENCES

[1] 2020. About | Time Project. https://thetimeproject.co.uk/about

[2] 2020. Deliveroo Unwrapped · Riders Roovolt. https://data.ridersroovolt.com/

[3] 2020. Driver's Seat Cooperative. https://driversseat.co/

[4] 2020. Home - Coworker. https://home.coworker.org/

[5] 2020. Target's Delivery Workers Are Staging a Walkout. https://www.vice.com/en/article/z3b9kw/shipt-workers-are-staging-a-walkout

[6] 2020. Twilio: Communication APIs for SMS, Voice, Video & Authentication. https://www.twilio.com/

[7] 2020. WeClock. https://weclock.it/

[8] 2020. Worker Info Exchange. https://www.workerinfoexchange.org

[9] 2021. Digital Worker Inquiry: Data, Solidarity, Leverage. https://digitalworkerinquiry.com/

[10] 2021. *Tesseract OCR*. tesseract-ocr. https://github.com/tesseract-ocr/tesseract

[11] Jeremias Adams-Prassl. 2019. What if your boss was an algorithm? Economic Incentives, Legal Challenges, and the Rise of Artificial Intelligence at Work. *Comp. Labor Law & Policy Journal* 41 (2019), 123.

[12] Romano Alquati. 2019. Co-research and worker's inquiry. *South Atlantic Quarterly* 118, 2 (2019), 470–478.

[13] Arjun Appadurai. 2006. The right to research. *Globalisation, societies and education* 4, 2 (2006), 167–177.

[14] Jef Ausloos and Pierre Dewitte. 2018. Shattering one-way mirrors – data subject access rights in practice. *International Data Privacy Law* 8, 1 (Feb. 2018), 4–28. https://doi.org/10.1093/idpl/ipy001

[15] Manijeh Badiee, Sherry C. Wang, and John W. Creswell. 2012. Designing community-based mixed methods research. *American Psychological Association* (2012).

[16] Matthew T. Bodie. 2021. The Law of Employee Data: Privacy, Property, Governance. *Indiana Law Journal* 97 (2021).

[17] David Broder. 2020. The Autumn and Fall of Italian Workerism. *Catalyst* 3 (2020). Issue 4. https://catalyst-journal.com/2020/03/the-autumn-and-fall-of-italian-workerism/

[18] Patrick Carmichael. 2020. Postdigital Possibilities: Operaismo, Co-research, and Educational Inquiry. *Postdigital Science and Education* 2, 2 (Apr. 2020), 380–396. https://doi.org/10.1007/s42438-019-00089-0

[19] Chun-Wei Chiang, Anna Kasunic, and Saiph Savage. 2018. Crowd coach: Peer coaching for crowd workers' skill growth. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–17.

[20] Brandon Dang, Martin J. Riedl, and Matthew Lease. 2018. But who protects the moderators? the case of crowdsourced image moderation. *arXiv preprint arXiv:1804.10999* (2018).

[21] Valerio De Stefano. 2019. " Negotiating the Algorithm": Automation, Artificial Intelligence, and Labor Protection. *Comp. Lab. L. & Pol'y J.* 41 (2019), 15.

[22] Sylvie Delacroix and Neil D. Lawrence. 2019. Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance. *International Data Privacy Law* 9, 4 (Nov. 2019), 236–252. https://doi.org/10.1093/idpl/ipz014

[23] Niels van Doorn and Adam Badger. 2020. Platform Capitalism's Hidden Abode: Producing Data Assets in the Gig Economy. *Antipode* 52, 5 (2020), 1475–1495. https://doi.org/10.1111/anti.12641

[24] Horace B. (Horace Bookwalter) Drury. 1918. *Scientific management; a history and criticism.* AMS press. http://archive.org/details/cu31924002406647

[25] Catherine D'Ignazio and Rahul Bhargava. 2020. *Data visualization literacy: A feminist starting point.* Amsterdam University Press, 207–222. https://doi.org/10.2307/j.ctvzgb8c7.19

[26] Sai Englert, Jamie Woodcock, and Callum Cant. 2020. Digital Workerism: Technology, Platforms, and the Circulation of Workers' Struggles. *tripleC Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society* 18, 1 (Jan. 2020), 132–145. https://doi.org/10.31269/triplec.v18i1.1133

[27] Cynthia Estlund. 2010. Just the facts: the case for workplace transparency. *Stan. L. Rev.* 63 (2010), 351.

[28] Cynthia Estlund. 2014. Extending the case for workplace transparency to information about pay. *UC Irvine L. Rev.* 4 (2014), 781.

[29] Frank Bunker Gilbreth. 1914. *Primer of Scientific Management.* Van Nostrand.

[30] Davydd J. Greenwood, William Foote Whyte, and Ira Harkavy. 1993. Participatory action research as a process and as a goal. *Human relations* 46, 2 (1993), 175–192.

[31] Karen Gregory. 2021. 'Worker Data Science' Can Teach Us How to Fix the Gig Economy. *Wired* (2021). https://www.wired.com/story/labor-organizing-unions-worker-algorithms/

[32] Kathleen Griesbach, Adam Reich, Luke Elliott-Negri, and Ruth Milkman. 2019. Algorithmic control in platform food delivery work. *Socius* 5 (2019), 2378023119870041.

[33] Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage, Chris Callison-Burch, and Jeffrey P. Bigham. 2018. A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174023

[34] Ellie Harmon and M. Six Silberman. 2019. Rating working conditions on digital labor platforms. *Computer Supported Cooperative Work (CSCW)* 28, 5 (2019), 911–960.

[35] Worker Info Exchange. 2020. Gig workers score historic digital rights victory against Uber & Ola. https://www.workerinfoexchange.org/post/gig-workers-score-historic-digital-rights-victory-against-uber-ola-2

[36] Lilly C. Irani and M. Six Silberman. 2013. Turkopticon: Interrupting worker invisibility in amazon mechanical turk. 611–620.

[37] Harmanpreet Kaur, Alex C. Williams, Daniel McDuff, Mary Czerwinski, Jaime Teevan, and Shamsi T. Iqbal. 2020. Optimizing for happiness and productivity: Modeling opportune moments for transitions and breaks at work. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.

[38] Catie Keck. 2020. *DoorDash Tip-Skimming Scheme Prompts Class Action Lawsuit Seeking All Those Tips That Didn't Go to Drivers*. https://gizmodo.com/doordash-tip-skimming-scheme-prompts-clash-action-lawsu-1836820630

[39] Vera Khovanskaya, Lynn Dombrowski, Jeffrey Rzeszotarski, and Phoebe Sengers. 2019. The Tools of Management: Adapting Historical Union Tactics to Platform-Mediated Labor. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–22.

[40] Matthias A. Lee. 2021. *Python Tesseract*. https://github.com/madmaze/pytesseract

[41] Hanlin Li, Bodhi Alarcon, Sara Milkes Espinosa, and Brent Hecht. 2018. Out of site: Empowering a new approach to online boycotts. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–28.

[42] Richard Lindstrom. 2000. "They All Believe They Are Undiscovered Mary Pickfords": Workers, Photography, and Scientific Management. *Technology and Culture* 41, 4 (2000), 725–751. http://www.jstor.org/stable/25147595

[43] René Mahieu and Jef Ausloos. 2020. Recognising and Enabling the Collective Dimension of the GDPR and the Right of Access. (Jul. 2020). https://doi.org/10.31228/osf.io/b5dwm

[44] Ruth Milkman, Luke Elliott-Negri, Kathleen Griesbach, and Adam Reich. 2020. Gender, Class, and the Gig Economy: The Case of Platform-Based Food Delivery. *Critical Sociology* (2020).

[45] Sara Ashley O'Brien. 2019. On-demand workers are protesting – using the apps they work for - CNN. https://edition.cnn.com/2019/06/14/tech/postmates-blitz-up-worker-protest/index.html

[46] Raniero Panzieri. 1994. *Spontaneità e organizzazione: gli anni del Quaderni rossi, 1959-1964*.

[47] Alex Rosenblat and Luke Stark. 2016. Algorithmic labor and information asymmetries: A case study of Uber's drivers. *International Journal of Communication* 10 (2016), 27.

[48] Niloufar Salehi, Lilly C Irani, Michael S Bernstein, Ali Alkhatib, Eva Ogbe, and Kristy Milland. 2015. We are dynamo: Overcoming stalling and friction in collective action for crowd workers. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 1621–1630.

[49] Saiph Savage, Chun Wei Chiang, Susumu Saito, Carlos Toxtli, and Jeffrey Bigham. 2020. Becoming the super turker: Increasing wages via a strategy from high earning workers. In *Proceedings of The Web Conference 2020* (2020). 1241–1252.

[50] Saiph Savage, Andres Monroy-Hernandez, and Tobias Höllerer. 2016. Botivist: Calling volunteers to action using online bots. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 813–822.

[51] The Rideshare Guy. 2020. Here's What Drivers Are Saying About the Coronavirus. https://therideshareguy.com/uber-drivers-can-survive-the-coronavirus/

[52] Molly Tran and Rosemary K. Sokas. 2017. The gig economy and contingent work: An occupational health assessment. *Journal of occupational and environmental medicine* 59, 4 (2017).

[53] Vincent W.-S. Tseng, Matthew L. Lee, Laurent Denoue, and Daniel Avrahami. 2019. Overcoming Distractions during Transitions from Break to Work using a Conversational Website-Blocking System. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300697

[54] Nicholas Vincent, Brent Hecht, and Shilad Sen. 2019. "Data strikes": evaluating the effectiveness of a new form of collective action against technology companies. In *The World Wide Web Conference*. 1931–1943.

[55] Nicholas Vincent, Hanlin Li, Nicole Tilly, Stevie Chancellor, and Brent Hecht. 2021. Data leverage: A framework for empowering the public in its relationship with technology companies. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 215–227.

[56] Working Washington. 2019. Delivering Inequality, What Instacart Really Pays, and How the Company Shifts Costs to Workers.

[57] William Foote Whyte, Davydd J. Greenwood, and Peter Lazes. 1991. Participatory action research: Through practice to science in social research. *Participatory action research* (1991), 19–55.

[58] Jamie Woodcock. 2014. The workers' inquiry from Trotskyism to Operaismo: A political methodology for investigating the workplace. *Ephemera: theory & politics in organizations* 14, 3 (2014), 493–513.
[59] Jamie Woodcock. 2021. Towards a Digital Workerism: Workers' Inquiry, Methods, and Technologies. *NanoEthics* 15, 1 (2021), 87–98.
[60] Steve Wright. 2007. Back to the future: Italian workerists reflect upon the operaista project. *Ephemera: Theory & Politics in Organization* 7 (2007), 270–281.

Received January 2022; revised April 2022; accepted May 2022