

Examining Crowd Work and Gig Work Through The Historical Lens of Piecework

Ali Alkhatib, Michael S. Bernstein, Margaret Levi

Computer Science Department and CASBS

Stanford University

{ali.alkhatib, msb}@cs.stanford.edu, mlevi@stanford.edu

ABSTRACT

The internet is empowering the rise of crowd work, gig work, and other forms of on-demand labor. A large and growing body of scholarship has attempted to predict the socio-technical outcomes of this shift, especially addressing three questions: 1) What are the complexity limits of on-demand work?, 2) How far can work be decomposed into smaller microtasks?, and 3) What will work and the place of work look like for workers? In this paper, we look to the historical scholarship on piecework — a similar trend of work decomposition, distribution, and payment that was popular at the turn of the 20th century — to understand how these questions might play out with modern on-demand work. We identify the mechanisms that enabled and limited piecework historically, and identify whether on-demand work faces the same pitfalls or might differentiate itself. This approach introduces theoretical grounding that can help address some of the most persistent questions in crowd work, and suggests design interventions that learn from history rather than repeat it.

ACM Classification Keywords

H.5.3. Information Interfaces and Presentation (e.g. HCI): Group and Organization Interfaces

Author Keywords

Crowd work; gig work; on-demand work; piecework

INTRODUCTION

The past decade has seen a flourishing of computationally-mediated labor. A framing of work into modular, pre-defined components enables computational hiring and management of workers at scale [66, 14, 82]. In this regime, distributed workers engage in work whenever their schedules allow, often with little to no awareness of the broader context of the work, and often with fleeting identities and associations [102, 92].

For years, such labor was limited to information work such as data annotation and surveys [80, 160, 167, 48, 116]. However,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI 2017, May 06 - 11, 2017, Denver, CO, USA

© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-4655-9/17/05...\$15.00

DOI: <http://dx.doi.org/10.1145/3025453.3025974>

physically embodied work such as driving and cleaning have now spawned multiple online labor markets as well [92, 147, 65, 141]. In this paper we will use the term *on-demand labor*, to capture this pair of related phenomena: first, *crowd work* [82], on platforms such as Amazon Mechanical Turk (AMT) and other sites of (predominantly) information work; and second, *gig work* [45], often as platforms for one-off jobs, like driving, courier services, and administrative support.

The realization that complex goals can be accomplished by directing crowds of workers has spurred firms to explore sites of labor such as AMT to find the limits of this distributed, on-demand workforce. Researchers have also taken to the space in earnest, developing systems that enable new forms of production (e.g. [11, 15, 115]) and pursuing social scientific inquiry into the workers on these platforms [125, 135]. This research has identified the sociality of gig work [51], as well as the frustration and disenfranchisement that these systems effect [71, 102, 104]. Others have focused on the responses to this frustration, reflecting on the resistance that workers express against digitally-mediated labor markets [92, 130].

This body of research has broadly worked toward the answer to one central question: *What does the future hold for on-demand work and those who do it?* Researchers have offered insights on this question along three major threads: First, what are the complexity limits of on-demand work — specifically, how complex are the goals that crowd work can accomplish, and what kinds of industries may eventually utilize it [139, 77, 166, 164, 108, 56]? Second, how far can work be decomposed into smaller microtasks [24, 98, 90, 26, 109]? And third, what will work and the place of work look like for workers [71, 70, 51, 104]?

This research has largely sought to answer these questions by examining extant on-demand work phenomena. So far, it has not offered an ontology to describe or understand the developments in worker processes that researchers have developed, or the emergent phenomena in social environments; nor has any research gone so far as to anticipate future developments.

Piecework as a lens to understand on-demand work

In this paper, we offer a framing for on-demand work as a contemporary instantiation of *piecework*, a work and payment structure which breaks tasks down into discrete jobs, wherein payment is made for *output*, rather than for *time*. We are not the first to relate on-demand work to piecework: in 2013, for

	<i>Observations in piecework</i>	<i>Mechanism</i>	<i>Implications for On-demand Work</i>
Complexity	Growth from simple tasks such as sewing to more complex composite outcomes on the assembly line floor.	Complexity was limited to tasks that could be easily measured and evaluated for payment by the piece.	Measurement and verification will remain persistent challenges that will limit complexity unless solved.
Decomposition	Work began sliced such that non-experts could perform each piece, but over time was sliced such that non-overlapping expertise was required for each step.	Scientific Management and Taylorism informed and drove decomposition by measuring and facilitating the optimization of smaller tasks.	After scientific management matured, piecework began specialized training to create experts in narrow tasks. A similar shift seems feasible with on-demand work.
Workers	Firms antagonized and exploited workers, leading workers to support one another independently, ultimately resulting in strong advocacy groups counterbalancing firms.	The features of piecework (independence and transience) were both the fulcrum managers used to exploit workers as well as the focal point around which workers bonded.	While worker frustrations are similar, the decentralized nature of on-demand work will limit collective action until there exist platforms to coordinate and exert pressure.

Table 1. Piecework and on-demand work have both wrestled with questions of how complex work can get, how finely-sliced tasks can become, and what the workplace will look like for workers. We connect piecework’s history (left) to the mechanisms that determined its outcomes to these three questions (center) in order to derive predictions for modern on-demand work (right).

example, Kittur et al. referenced crowd work as piecework briefly as a loose analogy [82]. Our goal in this paper is to inspect the relationship much more closely. But more than this, the framing of on-demand labor as a reinstantiation of piecework gives us years of historical material to help us make sense of this new form of work, and allows us to study on-demand work through a theoretical lens that is informed by years of rigorous, empirical research.

More concretely, by positioning on-demand labor as an instantiation or even a continuation of piecework, we can make sense of past events as part of a much larger series of interrelated phenomena (Table 1). We can reflect on differences in the features that impacted piecework historically and on-demand work today. And, to some extent, we can use these differences to offer some predictions of what on-demand work researchers and workers themselves might expect to see on the horizon. For example, we will draw on piecework’s scholarship on task decomposition, which was historically limited by shortcomings in measurement and instrumentation, and leverage that insight to suggest how modern technology affects this mechanism in on-demand work — namely, enabling precise tracking and measurement via algorithms and software.

We organize this paper as follows: first, we review the definition and history of piecework to make clear the analogy to on-demand work; and second, we examine the three major research questions above using the lens of piecework. For each question, we will contrast the perspective the piecework scholarship offers with on-demand labor’s body of research, identify similarities and differences, and then offer predictions for on-demand work.

A REVIEW OF PIECEWORK

The HCI community has used the term “piecework” to describe myriad instantiations of on-demand labor, but researchers have generally made this allusion in passing. Since we trace a much stronger parallel between (historical) piecework and (contemporary) on-demand work, a more comprehensive background on piecework will be useful. Specifically, first, we’ll define “piecework” as researchers in its field understand it; and second, we’ll trace the rise and fall of piecework at a high level, identifying key figures and ideas during this time. This section is not intended to be comprehensive: instead, it sets

up the scaffolding necessary for our later investigations of on-demand work’s three questions: complexity limits, task decomposition, and worker relationships.

What is piecework?: A primer and timeline

Aligning on-demand work with piecework requires an understanding of what piecework is. While it has had several definitions over the years, we can trace a constellation of characteristics that recur throughout the literature. We’ll follow this research, collecting descriptions, examples, and definitions, to develop a sense of piecework.

Piecework’s history traces back further than most would likely expect. Grier describes the process astronomers adopted of hiring teenage men to calculate equations in order to better-predict the trajectories of various celestial bodies in the night sky [52]. In the first half of the 19th century, George Airy was perhaps the first to rigorously put piecework-style decomposition to work; by breaking complex calculations into constituent parts, and training young men to solve simple algebraic problems, Airy could distribute work to many more people than could otherwise complete the full calculations.

Piecework began in the intellectual domain of astronomical calculations and projections, but it found its foothold in manual labor. Piecework took hold in farm work [117], in textiles [9, 120], on railroads [19], and elsewhere in manufacturing [131] by the mid-19th century. By 1847 we find a concise definition of piecework in Raynbird’s essay on piecework, particularly driven toward encapsulating the manual labor of farmwork. He does this by contrasting two paradigms: “the chief difference lies between the day-labourer, who receives a certain some of money . . . for his day’s work, and the task-labourer, whose earnings depend on the quantity of work done” [117]. Chadwick offers a number of illustrative examples: “payment is made for each hectare which is pronounced to be well ploughed . . . for each living foal got from a mare; . . . for each living calf got” [25]. This framing gives us an intuitive sense of piecework; “payment for results,” as he calls it, is not only common in practice, but well-studied in labor economics [43, 153, 154, 61].

It’s worth acknowledging that “this distinction [between piece-rates and time-rates] was not completely clear-cut” [59]. Employers adopted piece-rates in some aspects and time-rates in

others. The Rowan premium system, for example, essentially paid workers a base rate for time plus additional pay depending on output [126]. As Rowan's premium system guaranteed an hourly rate regardless of the worker's productive output *as well as* additional compensation tied to performance, workers were in some senses "task-labourers", but in other senses "day-labourers". This was just one of several alternatives to strict time- and piece-rate remuneration paradigms.

In the late years of the 19th century, Taylor — a mechanical engineer with an interest in work efficiency — began studying and formalizing the decomposition, tracking, and management of tasks [142]. In 1911 he published *The principles of scientific management*, concretizing an idea that had nebulously been forming, and which he had been working out himself, for years [143]. Scientific Management (and Fordism) thrust piecework into higher gear, especially as mass manufacturing and a depleted wartime workforce forced industry to find new ways to eke out more production capacity.

It may be worth thinking about piecework through the lens of its *emergent* properties to help understand it. Raynbird argues for the merits of piecework, pointing out that "piece work holds out to the labourer an increase of wages as a reward for his skill and exertion . . . he knows that all depends on his own diligence and perseverance . . . [and] so long as he performs his work to the satisfaction of his master, he is not under that control to which the day-labourer is always subject". The argument that "task-labourers" enjoy freedom from control crops up in Raynbird's and later Rowan's works [117, 126].

We see this sense of independence in myriad times, locales, and industries. Satre offers a look into the lives and culture of "match-girls", teenage women who assembled matchsticks in the late 19th century in London. Of interest was their reputation "... for generosity, independence, and protectiveness, but also for brashness, irregularity, low morality, and little education" [131]. Hagan and Fisher document piecework from 1850 through 1930 in Australia, finding similar notions of independence and autonomy among piecework newspaper compositors: "If a piece-work compositor . . . decided that he did not want to work on a particular day or night, the management recognised his right to put a 'substitute' or 'grass' compositor in his place" [55]. This sense of independence and autonomy appears to be a common thread of piecework.

Since workers could now choose their own schedule and style, a discussion arose on how best to manage pieceworkers. This conversation came to regard workers antagonistically [127], a far cry from the earlier rhetoric on piecework, which promised that pieceworkers would gladly work diligently and for as long as possible, as incentive-based pay rewarded exactly, and thus aligned the goals of both managers and workers [32].

Piecework opened the door for people who previously couldn't participate in the labor market to do so, and to acquire job skills incrementally. During World War II, women received training in narrow subsets of more comprehensive jobs, enabling work in capacities similar to conventional (male) workers [59]. Women previously had virtually no opportunities to engage in engineering and metalworking apprenticeships

as men did; now, they could be trained quickly on narrowly scoped tasks, demonstrate proficiency, and become experts. "Rosie the Riveter", an icon of 20th century America who represented empowerment and opportunity for women [63], would have been a pieceworker [35].

Piecework's popularity in the United States and Europe fell almost as quickly as it had climbed. Between 1938 and 1942, the proportion of metal workers under piecework systems had climbed steeply from 11% to 60% [58]. By 1961, Carlson finds, the proportion dropped to 8% [23]. He notes that, from 1973 to 1980, the holdouts of piecework — where more than 50% worked under incentive wage plans — were principally in clothes-making (e.g. hosiery, footwear, and garments). Hart and Roberts offer a number of explanations for the sudden demise of piecework. The salient suggestions include: 1) the emergence of more effective, more nuanced incentive models — rewarding teams for complex achievements, for instance; 2) the shifting of piecework industries such as manufacturing and textiles to other countries; and 3) the quality of "multidimensional" work, which was too difficult to evaluate [59].

In summary, piecework: 1) paid workers for *quantity* of work done, rather than *time* done, but occasionally mixed the two payment models; 2) afforded workers a sense of freedom and independence; and 3) structured tasks in such a way as to facilitate more narrowly scoped training and education.

Viewing on-demand work as a modern instantiation of piecework is relatively straightforward by this definition. First, platforms such as Amazon Mechanical Turk (AMT), Uber, Upwork, and TaskRabbit pay by the task, though some mix systems in similar ways to the Rowan system's combination of piece rate and time rate pay. Second, workers are attracted to these platforms by the freedom they offer to pick the time and place of work [102, 18]. Third, system developers as on Mechanical Turk typically assume no professional skills in transcription or other areas, and attempt to build that expertise into the workflow [110, 11]. Given this alignment, many of the same historical properties of piecework will apply to on-demand work as well.

Case studies in piecework

Throughout the paper, we will return to four case studies to frame our analyses: Airy's use of human computers; domestic and farm workers; the "match-girls" strike; and industrial and assembly-line workers. In introducing these cases at a high level, we'll trace the history of piecework while also framing the later analysis of the leading research threads we named earlier: complexity, decomposition, and relationships.

Airy's computers

In the 19th century, the calculation of celestial bodies had become a competitive field, and Airy needed to compute tables that would allow sailors to locate themselves by starlight from sea. This work ostensibly called for educated people who comprehensively understood mathematics. Airy realized that he could break the tasks down and delegate the constituent parts to human computers, or people who could compute basic functions. These human computers "... possessed the basic skills of mathematics, including 'Arithmetic, the use of

Logarithms, and Elementary Algebra' ” [52]. As a result, many of Airy's computers had relatively rudimentary educations compared to those that typically worked in the calculation of solar tables. Airy distributed tasks by mail, allowing work to be completed by a somewhat geographically distributed workforce, and paid for each piece of work completed.

The human computers captured several aspects of task decomposition that would become common. First, the work was designed such that it could be done independently and without collaboration. Second, the work was designed so that intermediate results could be quickly verified: Airy would have two workers each do the calculation, and another person compare their answers. Third, Airy identified ways to decompose the large task into narrowly-trainable subtasks.

Some of Airy's policies were more controversial, for example firing computers once they reached age 23. This practice ensured two outcomes that disfavored workers. First, it drastically reduced professional advancement, as workers' careers ended quickly, and without conventional backgrounds in mathematics they later struggled to find work for which their experience was meaningful. And second, it limited workers' ability to organize by ensuring that workers were in little communication with each other, and that they had almost no opportunity to recognize their circumstances and to coordinate.

Domestic and farmhand labor

The application of piecework to farm work in the late 19th century and later to manufacturing of small goods, such as garments and matches, at the turn of the 20th century proved to be a formative period for piecework as we would come to know it. Piecework regimes in farms and in homes engaged workers in assembling clothing. Textile manufacturers found that they could deliver fabric to people at their homes, asking them to sew together clothing. The manufacturers would later return to retrieve the finished garments, paying these workers for each piece of clothing completed. Farm work applied the idea of piecework by paying workers for tasks like picking bushels of fruit or bringing to birth animals [25].

Workers could, in principle, assemble as much or as little clothing as they wanted; the reality was more grim, as Riis documented in *How the Other Half Lives* in 1901 [120]. He found that workers endured bleak living conditions and worked long hours attempting to scrape together a living.

The match-girls' strike

Match-makers were some of the first workers in mass manufacturing to successfully rally for political causes. At the end of the 19th century, manufacturers had begun to employ teenage women to assemble matchsticks in factories. These women rallied first in the form of a march on parliament in 1871 to protest a proposed tax, and later (more famously) in what was later called “the match-girls strike of 1888” [131]. This later strike was sparked by a worker's arbitrary docking of pay, but much deeper resentment had been simmering for years. Match-girls were already frustrated with the arbitrariness of management, poor working conditions, and having to work with hazardous materials such as white phosphorus, the

improper handling of which caused serious, painful, disfiguring medical conditions in the bones and ultimately death.

Regardless of what prompted it, the lasting impact of the match-girls strike of 1888 was profound. This was one of the earliest and most famous successful worker strikes, and perhaps the beginning of “militant trade unionism” [131]. As Webb and Webb described, “the match-girls' victory turned a new leaf in Trade Union annals”: in the 30 years after the match-girls strike, the Trade Union Movement enrollment grew from 20% of eligible workers to over 60% [152].

Match-girls were some of the earliest to have formed a trade union, according to Booth's account in 1903. Satre noted that match-girls “... pooled their resources to purchase their plumes and clothes ... and expressed their solidarity through small [and major] strikes” [17]. But they were also, as Satre confesses, known for “brashness, irregularity, low morality, and little education” [131]. These were workers who treasured their independence, but also fiercely protected one another. “Brashness” may have detracted from their public image, but almost undoubtedly contributed to their sense of solidarity, making their propensity to act against such unfair treatment and poor conditions understandable and maybe predictable.

Industrial workers

Piecework might be most familiar in the context of industrial and factory work, which largely defined manufacturing through the 20th century. Before the factory assembly line arose, however, railway companies adopted piecework regimes at the turn of the 20th century. What followed was a flourishing of management practices, as railway companies worked to find effective ways to motivate and evaluate this skilled workforce of engineers. Graves takes up a case study of the Santa Fe Railway, finding that they employed “efficiency experts” to develop a “standard time” to determine pay for each task at the company informed by “thousands of individual operations”; Graves goes on to list some of the roles required to facilitate piecework in the early 20th century — among them, “piecework clerks, inspectors, and ‘experts’ ” [49]. This oversight, while controversial (especially among workers [73]), paved the way for piecework to grow substantially.

The 1930s represented a boom for piecework on an unprecedented scale, especially among engineering and metalworking industries. Hart and Roberts characterize the 1930s — and more broadly the first half of the 20th century — as the “hey-day” of piecework. They attribute this to the shortage of male workers, who would have gone through a conventional apprenticeship process affording them more comprehensive knowledge of the total scope of work.

Piecework found its way into the war effort during World War II. With the vast majority of men drafted into service, factories found themselves turning to a mostly female workforce that had neither the formal training nor years of experience that men would have had from apprenticeships. Rather than attempting to train this new labor force in every aspect of industrial work, these women were trained for individual tasks and correspondingly assigned to that or a similar task.

RESEARCH QUESTIONS

Research in crowdsourcing has spent the better part of a decade exploring how to grow its limits. This has largely involved iteratively identifying barriers to high-quality, complex work, then overcoming them through novel designs of systems, workflows, and processes (e.g. [11, 118, 81]). The question has become *whether* there are limits to on-demand work, and if so, what factors determine them. To this question, a number of contributions to the field have pressed for answers.

The exploration of on-demand labor's potential and limits has principally navigated three dimensions: First, what are the complexity limits of on-demand work? Second, how far can work be decomposed into smaller microtasks? And third, what will work and the place of work look like for workers? We'll explore these aspects of on-demand labor by connecting to corresponding piecework literature and comparing its lessons to the current state of on-demand labor.

Complexity Limits of On-Demand Work

A key question to the future of on-demand work is *what* precisely will become part of this economy. Paid crowdsourcing began with simple microtasks on platforms such as Amazon Mechanical Turk, but microtasks are only helpful if they build up to a larger whole. So, our first question: how complex can the work outcomes from on-demand work be?

The perspective of on-demand work

Kittur et al. were among the first to ask whether crowdsourcing could be used for more than parallelizing tasks [81]. Their work showed that it could, with proof-of-concept crowdsourcing of encyclopedia articles and news summaries — tasks which could be verified or repeated with reasonable expectations of similar results. Seeking to raise the complexity ceiling, researchers have since created yet more applications and techniques, including conversational assistants [86], medical data interpreters [86], and idea generation [164, 162, 163].

To achieve complex work, this body of research has often applied ideas from Computer Science to design new workflows. System designers leverage techniques such as MapReduce [81] and sequence alignment algorithms [85], arranging humans as computational black boxes. This approach has proven a compelling one because it leverages the inherent advantages of scale, automation, and programmability that software affords.

It is now clear that this computational workflow approach works with some classes of complex tasks, but the broader wicked problems largely remain unsolved. As a first example, idea generation shows promise [164, 162, 163], but there is as yet no general crowdsourced solution for the broader goal of invention and innovation [46]. Second, focused writing tasks are now feasible [78, 11, 108, 144, 2], but there is no general solution to create a cross-domain, high-quality crowd-powered author. Third, data analysis tasks such as clustering [30], categorization [7], and outlining [97] are possible, but there is no general solution for sense-making. It is not yet clear what insights would be required to enable crowdsourced solutions for these broader wicked problems.

Restricting attention to non-expert, microtask workers proved limiting. So, Retelny et al. introduced the idea of crowdsourcing with online paid *experts* from platforms such as Upwork. Expert crowdsourcing enables access to a much broader set of workers, for example designers and programmers. The same ideas can then be applied to expert “macro-tasks” [29, 54], enabling the crowdsourcing of goals such as user-centered design [118], programming [89, 42, 27], and mentorship [139]. However, there remains the open question of how complex the work outcomes from expert crowds can be.

The perspective of piecework

Piecework's body of research most squarely addresses complexity in two of the cases we looked at earlier: Airy's human computers and among industrial workers.

Airy's work on astronomical charts opened the door to greater task complexity by encoding the intelligence into the process rather than the people. Airy's computers had relatively limited education in mathematics, but by combining simple mathematical operations, Airy was able to create a complex composite outcome [52]. Likewise, in Ford's factories, no individual could build the entire car, but the process could emergently produce one.

But when piecework initially entered the American economy, it was not used for complex work. Without having designed complex work processes, piecework managers were restricted to available workers' skills such as sewing: it was infeasible to provide new pieceworkers with the comprehensive education that apprenticeships imparted [59]. So, initially piecework arose for farmwork, and as Raynbird and others discuss, the practice remained relatively obscure until it blossomed in the textile industry [117]. Complexity levels remained low at the turn of the 20th century as piecework saturated densely populated urban areas such as London and New York City [120].

Measurement also limited the complexity of piecework: only tasks that could be measured and priced could be completed via piecework. Earlier we discussed Graves's and later Brown's analysis of railway workers. They identified task homogeneity and measurement as key requirements for piecework to be successful. However, complex, creative work — which is inherently heterogeneous and difficult to routinize — was unsuitable [49]. Brown's description of “efficiency experts” would corroborate this: efficiency experts can effectively gauge how long known tasks should take, but would find themselves overwhelmed if they attempted to assess creative tasks like scientific research, which can take an arbitrary number of iterations before proceeding to a subsequent step.

Moreover, piecework was limited to tasks that could be quickly and accurately evaluated. Hart argue that evaluation limited piecework's complexity: at some point, evaluating multidimensional work for quality (rather than for quantity) becomes infeasible. In his words, “if the quality of the output is more difficult to measure than the quantity [...] then a piecework system is likely to encourage an over-emphasis on quantity ... and an under-emphasis on quality” [60]. Complex work, which is often subjective to evaluate, falls victim to this pitfall.

Comparing the phenomena

The research on piecework tells us that we should expect it to thrive in industries where the nature of the work is limited in complexity [19], and become less common as work becomes more complex. Has computation shifted piecework's previous limits of expertise, measurement, and evaluation?

In some ways, yes: technology increases non-experts' levels of expertise by giving access to information that would otherwise be unavailable. For example, taxi drivers in London endure rigorous training to pass a test known as "The Knowledge": a demonstration of the driver's comprehensive familiarity with the city's roads. This test is so challenging that veteran drivers develop significantly larger the regions of the brain associated with spatial functions such as navigation [100, 99, 137, 138, 159, 158]. In contrast, with on-demand platforms such as Uber, services such as Google Maps and Waze make it possible for people entirely unfamiliar with a city to operate professionally [136, 62]. Other examples include search engines enabling information retrieval, and word processors enabling spelling and grammar checking. By augmenting the human intellect [40], computing has shifted the complexity of work that is possible with minimal or no training.

Algorithms have automated some tasks that previously fell to management. Computational systems now act as "piecework clerks" [49] to inspect and modify work [71, 104]. However, these algorithms are less competent than humans at evaluating subjective work, as well as in their ability to exercise discretion, causing new problems for workers and managers.

Implications for on-demand work

Algorithms are undoubtedly capable of shepherding more complex work than the linear processes available to Airy and Ford. However, as work becomes more complex, it becomes increasingly difficult to codify a process to achieve it [41, 38]. So, while algorithms will increase the complexity ceiling beyond what was possible previously with piecework, there is a fundamental limit to how complex such work can become.

Technology's ability to support human cognition will enable stronger assumptions about workers' abilities, increasing the complexity of on-demand work outcomes. Just as the shift to expert crowdsourcing increased complexity, so too will workers with better tools increase the set of tasks possible. Beyond this, further improvements would most likely come from replicating the success of narrowly-slicing education for expert work as Hart and Roberts and Grier described in their piecework examples of human computation [52] and drastically reformulating macro-tasks given the constraints of piecework [59]. An argument might be made that MOOCs and other online education resources provide crowd workers with the resources that they need, but it remains to be seen whether that work will be appropriately valued, let alone properly interpreted by task solicitors [4]. If we can overcome this obstacle, we might be able to empower more of these workers to do complex work such as engineering, rather than doom them to "uneducated" match-girl reputations [131]. However, many such experts are already available on platforms such as Upwork, so training may not directly increase the complex-

ity accessible to on-demand work unless it makes common expertise more broadly available.

Evaluation remains as difficult for crowd work as it did for the efficiency experts. Reputation systems for crowdsourcing platforms remain notoriously inflated [64]. Ultimately, many aspects of assessment remain subjective: whether a logo made for a client is fantastic or terrible may depend on taste.

So, in the case of complexity, the history of piecework does not yet offer compelling evidence that on-demand work will achieve far more complex outcomes than piecework did. Improvements in workflows, measurement, and evaluation have already been made, and it's not immediately clear that the remaining challenges are readily solvable. However, on-demand work will be far more broadly distributed than piecework historically was — reaching many more tasks and areas of expertise by virtue of the internet.

Decomposing Work

At its core, on-demand work has been enabled by decomposition of large goals into many small tasks. As such, one of the central questions in the literature is how finely-sliced these microtasks can become, and which kinds of tasks are amenable to decomposition. In this section, we place these questions in the context of piecework's Taylorist evolution.

The perspective of on-demand work

Many contributions to the design and engineering of crowd work consist of creative methods for decomposing goals. Even when tasks such as writing and editing cannot be reliably performed by individual workers, researchers have demonstrated that the decomposition of these tasks into workflows can succeed [81, 11, 144, 108]. These decompositions typically take the form of workflows, instantiated as algorithmically managed sequences of tasks that resolve interdependencies [14]. Workflows often utilize a first sequence of tasks to identify an area of focus (e.g., a paragraph topic [81], an error [11], or a concept [163, 165]) and a second sequence of tasks to execute work on that area. This decomposition style has been successfully applied across many areas, including food labeling [110], brainstorming [134, 162], and accessibility [86, 85, 87].

If decomposition is key to success in on-demand work, the question arises: what can, and can't, be decomposed? More pointedly, how thinly *should* work be sliced and subdivided into smaller and smaller tasks? The general trend has been that smaller is better, and the microtask paradigm has emerged as the overwhelming favorite [145, 146]. This work illustrates a broader sentiment in both the study and practice of crowd work, that microtasks should be designed resiliently against the variability of workers, preventing a single errant submission from impacting the agenda of the work as a whole fully exploiting the abstracted nature of each piece of work [69, 88, 148]. In this sense, finer decompositions are seen as more robust — both to interruptions and errors [29] — even if they incur a fixed time cost. At the extreme, recent work has demonstrated microtasks that take seconds [149, 21] or even fractions of a second [83]. However, workers perform better when similar tasks are strung together [88], or chained and arranged to maximize the attention threshold of workers [20].

Despite this, we as a community have leaned *into* the peril of low-context work, “embracing error” in crowdsourcing [83].

The general lesson has been that the more micro the task, and the more fine the decomposition, the greater the risk that workers lose context necessary to perform the work well. For example, workers edit adjacent paragraphs in inconsistent ways [11, 78], interpret tasks in different ways [74], and exhibit lower motivation [79] without sufficient context. Research has sought to ameliorate this issue by designing workflows to help workers “act with global understanding when each contributor only has access to local views” [150], typically by automatically or manually generating higher-level representations for the workers to reflect on [30, 150, 78].

As the additional context necessary to complete a task diminishes, the invisible labor of *finding* tasks [102] has arisen as a major issue. Chilton et al. illustrate the task search challenges on AMT [31]. Workers seek out good requesters [102] and then “streak” to perform many tasks of that same type [31]. Researchers have reacted by designing task recommendation systems (e.g. [33]) and minimizing the amount of time that people need to spend doing anything other than the work for which they are paid [22].

The perspective of piecework

Four major stages characterize decomposition in the history of piecework. The first stage was decomposition of an expert task such that it could be done by non-experts. This was arguably the main innovation of Airy’s human computers. Rather than hire expert computers, Airy identified ways to break down astrological calculations into steps that could be completed with only a basic knowledge of mathematics. Likewise, Brown argued that piecework arose in industries with homogeneous tasks and low fixed costs of machinery and training [19].

After decomposing tasks for amateurs, the second major stage was to apply the same methods to domain experts. Unlike Airy’s human computers, railway engineers had significant expertise [19]. As Brown noted, however, it was still possible to discretize and measure their work. Thus, experts such as railway engineers became pieceworkers as well.

Third, decomposition led to quantification and scientific management. What can be modularized can be measured, and what can be measured can be optimized. With Taylor’s formalization of scientific management in *Taylorism* (and Henry Ford’s eponymously named *Fordism*), piecework in the early and mid-20th century surged, especially in industrial work. Scientific management promised that the careful measurement of workers would yield higher efficiency and output [143, 95]. While Brown points out that piecework dramatically advanced the instrumented measurement of workers, in Taylor’s time highly instrumented, automatic measurement of workers was all but impossible [19]. Instead, managers conducted “stop watch time studies” [107], using completion times to inform per-task compensation, similarly to the efficiency experts hired in the Santa Fe Railway, but substantially more precise. The distillation of work into smaller units ultimately bottomed out with tasks as small as could be usefully measured [49].

The fourth and final stage was narrow expertise training. Even after work is decomposed and measured, there were not enough qualified workers available to do it. So, as World War II raged and there was a dearth of skilled workers, managers trained women just enough to be able to complete their tasks [59]. Over time, these women could gain proficiency and gain broader expertise.

Comparing the phenomena

Where measurement and instrumentation were limiting factors for historical piecework, computation has changed the situation so that a dream of scientific management and Taylorism — to measure every motion at every point throughout the workday and beyond — is not only doable, but trivial [151]. Where Graves directly implicates measurement as preventing scientific management from being fully utilized [49], modern crowd work is measuring and modeling every click, scroll, and keyboard event [129, 128]. The result is that on-demand work can articulate and track far more carefully than piecework historically could.

A second shift is the relative ease with which the metaphorical “assembly line” can be experimented with and measured. Historical manufacturing equipment could not quickly be assembled, edited, and redeployed [67]. In contrast, today system-designers can share, modify, and instantiate environments like sites of labor in a few lines of code [93, 96]. This opportunity has spurred an entire body of work investigating the effects of ordering, pacing, interruptions, and other factors accelerating scientific management that would have been all but impossible as few as 20 years ago [34, 20, 29, 28, 83].

Implications for on-demand work

If decomposition in piecework progressed in four stages, we have seen three of them in on-demand work so far. First, as with piecework, on-demand work began by decomposing tasks so that anyone could complete them, as with data labeling on Amazon Mechanical Turk. Second, we began to modularize and measure external expertise (e.g., software engineering, design) so that it could be brought into crowdsourcing systems [118, 27]. Third, we used measurement to mathematically optimize workers’ behaviors so that we could make the systems more efficient [155].

The fourth stage, then, appears likely to occur: narrow training of workers for these decomposed expert tasks. There is demand for skilled workers in many crowdsourcing tasks, and systems to help train workers [139]. We might expect to see the rise of systems that scaffold workers into extremely narrow areas of expertise, for instance using online courses as proof of expertise in a specific domain necessary for a microtask.

Finally, improved measurement and lowered costs of production have made it feasible to apply piecework methods to many domains where it may not have historically been possible. The limit is no longer measurement precision, but human cognition. Task switching and other cognitive costs make it difficult to work on tasks so far decontextualized from their original intention [88]. There will of course be tasks that can be decomposed without much context, and these will form the most fine-grained of microtasks. However, other tasks cannot be

freed from context — for example, logo design requires a deep understanding of the client and their goals.

Workers' Relationships to their Work

HCI and CSCW have historically framed themselves around supporting work. While all artifacts have politics, the recent shift into computational labor systems has directly impacted the lives and livelihoods of workers in new ways. So, it's imperative to ask: What will the future look like for the workers who use these systems?

The perspective of on-demand work

Who are the crowd workers and what draws them to crowd work? Early literature emphasized motivations like fun and spare change, but this narrative soon shifted to emphasize that many workers use platforms such as Amazon Mechanical Turk as a primary source of income [75, 68, 8]. Despite this, Mechanical Turk is a disappointingly low-wage worksite for most people in the United States [68, 102, 53]. Thus, those who choose to opt out of the traditional labor force and spend significant time on Mechanical Turk are especially motivated by the opportunity for autonomy and transience between tasks [75]. While some describe Turkers as powerless victims or even unaware of what's going on, this framing is increasingly being rejected by workers and designers as “cast[ing] Turkers as dopes in the system.” [70].

Workers' relationships with requesters are fraught. The unbridled power that requesters have over workers, and the resultant frustration that this generates, has motivated research into the tense relationships between workers and requesters [50, 130]. Workers are often blamed for any low-quality work, regardless of whether they are responsible [102, 104]. Some research is extremely open about this position, blaming unpredictable work on “malicious” workers [47] or those with “a lack of expertise, dedication [or] interest” [133]. Workers resent this position, and for good reason. Irani and Silberman highlighted the information asymmetry between workers and requesters on AMT, which led to the creation of Turkopticon, a site which allows Turkers to rate and review requesters [71]. Dynamo then took this critique on information asymmetry and power imbalances further, designing a platform to facilitate collective action among Turkers to changes to their circumstances [130].

Researchers have also begun to appreciate the sociality of crowd workers. Because the platforms do not typically include social spaces, workers instead congregate off-platform in forums and mailing lists. There, Turkers exchange advice on high-paying work, talk about their earnings, build social connections, and discuss requesters [102]. Many crowd workers know each other through offline and online connections, coordinating behind-the-scenes despite the platforms encouraging independent work [51, 161]. However, the frustration and mistrust that workers experience with requesters does occasionally boil over on the forums.

The perspective of piecework

Early observers believed that workers were strongly motivated by the autonomy of working in the piecework model. Clark observed textile mill pieceworkers and reported, “When he works by the day the Italian operative wishes to leave before

the whistle blows, but if he works by the piece he will work as many hours as it is possible for him to stand” [32]. However, the emergent trend contrasted with this early rhetoric, as when workers began instituting “The Fix”, deliberately slow work to game efficiency experts [127]. Piece workers, Roy found, would form acrimonious relationships with their managers.

Soon, workers began resisting piecework regimes. The match-girls engaged in their famous strike of 1888, particularly pushing to abolish the fines that were taken out of their wages. Soon others followed suit, including women in the garment industry in Philadelphia who established collective bargaining rights [39] and national coal miners who effected an individual minimum wage in 1912 [122].

Many worker organizations began weighing in against piecework and the myriad oversights it made in valuing workers' time [73, 119]. As mounting attention increasingly revealed problems in piecework's treatment of workers, workers themselves began to speak out about their frustration with this new regime. Organizations representing railway workers, mechanical engineers, and others began to mount advocacy in defense of workers [73, 119]. Pieceworkers' relationships with their employers eventually developed a pattern of using laborer advocacy groups [94, 5, 103, 72]. Following the template of the match-girls, collective action grew to become a central component of negotiating with managers [57, 113].

Relative to the modern on-demand workers, there is a noticeable dearth of information on the interpersonal relationships among pieceworkers beyond the match-girls at the end of the 19th century. Nevertheless, we can offer some observations: primary sources indicate that labor organizations wished for workers to identify as a collective group, “not only as railroad employees but also as members of the larger life of the community” [73]. Doing this, Ostrom and others later argued, would facilitate collective action and perhaps collective governance [114, 57, 113]. Riis also contributed to this sense of shared struggle and endurance by documenting pieceworkers in their home-workplaces, literally bringing to light the grim circumstances in which pieceworkers lived and worked [120].

Comparing the phenomena

There was generally less written about work quality concerns for historical pieceworkers than there is in modern on-demand work. Why the difference? One possibility is that, by writing web scripts and applying them to many tasks, a small number of spammers have an outsized influence on the perception of bad actors. Another possibility is that historical pieceworkers faced much more risk in shirking: it was much harder for pieceworkers to move to a new location and find a new job. Today, Mechanical Turk workers can work for a dozen or more different groups in the span of a day. A third possibility: online anonymity breeds distrust [44], and where pieceworkers could be directly observed by foremen and known to them, online workers are known by little more than an inscrutable alphanumeric string, like A2XJMS2J2FMVXX.

The relationship between workers and employers has also shifted: while historically the management of workers had to be done through a foreman, foremen of the 20th century have

largely been replaced by algorithms of the 21st century [92]. Consequently, the agents managing work are now cold, logical, and unforgiving. While a person might recognize that the “attention check” questions proposed by Le et al. and others ensure that malicious and inattentive workers are stopped [91, 111], some implementations of these approaches only seem to antagonize workers [104]. As Anderson and Schmittlein wrote in 1984, “when performance is difficult to evaluate, imperfect input measures and a manager’s subjective judgment are preferable to defective (simple, observable) output measures” [6]. This frustration has only grown as requesters have had to rely on automatic management mechanisms. Only a few use the equivalent of human foremen [54, 84].

Relative to the history of collective action for pieceworkers, on-demand workers have struggled to make their voices heard [130, 70, 71]. With workers constantly drifting through these platforms, and with many part-time members, it’s extremely difficult to corral the group to make a collective decision [130]. Even when they can, enforcement remains a challenge: while pieceworkers could physically block access to a site of production and convince other workers to join them, online labor markets provide no facilities for workers to change the experience of other workers. This is a key limitation — without it, workers cannot enforce a strike.

Implications for on-demand work

The decentralization and anonymization of on-demand work, especially online crowd work, will continue to make many of its social relationships a struggle. While some workers get to know each other well on forums [102, 51], many never engage in these social spaces. Without intervention, worker relationships and collectivism are likely to be inhibited by this decentralized design. One option is to build worker centralizing points into the platform, for example asking workers to vote on each others’ reputation or allowing groups of workers to collectively reject a task from the platform [156].

The history of piecework further suggests that relationships between workers and employers might be improved if employers engaged in more human management styles. Instead of delegating as many management tasks as possible to an algorithm, it might be possible to build dashboards and other information tools that empower modern crowd work foremen [84]. If the literature on piecework is to be believed, more considerate *human* management may resolve many of the tensions.

Reciprocally, crowd work may be able to inform piecework research in this domain. There exists far less literature about piece workers’ relationships than there does today about on-demand workers’ relationships. Two reasons stand out: first, modern platforms are visible to researchers in ways that the sites of piece work labor were not. Second, Anthropology stands on a firmer theoretical and methodological basis than it did at the turn of the 20th century. Malinowski, Boas, Mead, and other luminaries throughout the first half of the 20th century effectively defined Cultural Anthropology as we know it today; *participant-observation*, the *etic* and the *emic* understanding of culture, and *reflexivity* didn’t take even a resemblance of their contemporary forms until these works [101, 16,

106]. On-demand labor today may give us an opportunity to revisit open questions in piecework with a more refined lens.

DISCUSSION

In our analysis of on-demand work via the piecework lens, three issues arise: 1) the hazards of predicting the future, 2) utopian and dystopian visions, and 3) a research agenda. We will attempt to grapple with these questions here explicitly.

The Hazards of Predicting the Future

The past isn’t a perfect predictor for the future; as Scholz cautions, “it would be wrong to conclude that in the realm of digital labor there is nothing new under the sun” [132]. Our analysis is limited by the differences, foreseen and unforeseen, between historical piecework and modern on-demand work. For example, unlike physical work environments, people can (and often do) make one-off contributions to online communities [105]. While we have attempted to identify some likely parallels and divergences between piecework and on-demand work, we can’t claim to have accounted for everything.

But this does not mean that attempting to draw meaningfully from historical scholarship would be folly; enough of piecework can and does inform on-demand work that HCI and CSCW researchers might seek out historical framings for other phenomena of study as well. While we can only speculate one of (perhaps many) possible futures, history does allow us to articulate and bound which futures appear more likely.

Rosenberg and others have contributed substantially to the practice in part by clearly limiting the extent of their claims — only offering, for instance, “to narrow our estimates and thus to concentrate resources in directions that are more likely to have useful payoffs” [123]. Using this approach, our method of relating history to modern socio-technical systems may be a useful tool for researchers attempting to make sense of ostensibly new phenomena. In other words, offering “that past history is an indispensable source of information to anyone interested in characterizing technologies” [124].

Utopian and Dystopian Visions

An easy narrative is to characterize the future of on-demand labor at one of two extremes. On one hand, crowd work researchers imagine the application of crowdsourcing as a potentially bright future that enables the achievement of near-impossible goals and career opportunities [140, 82, 15, 139]. On the other hand, researchers warn that on-demand labor will create exploitative sites of dispossession [132], discrimination [37], and invisible, deeply frustrated workers [71, 13].

A uniquely challenging facet of this domain is the public attention that it has garnered. Activists have described speculative work as having “essentially been turned into modern-day slaves” [10]. Meanwhile, advocates have described it as “a project of sharing aimed at providing ordinary people with more economic opportunities and improving their lives” [36].

Piecework teaches us that, without appropriate norms and policies, the dystopian outcome has happened and may happen again. The piecework nature of on-demand work induces us “to neglect tasks that are less easy to measure” [3], rewarding

us not for creativity but predictability; payment for this work may ultimately be determined by algorithms that fundamentally don't understand people; the layers between us and our managers might eventually become "defective (simple, observable)" algorithms [6], just like those which already frustrate on-demand workers [92, 130, 71]. However, social policy has advanced since the early 1900s, so as on-demand work grows, a repeat of *How the Other Half Lives* [120] seems less likely.

On the other hand, while piecework's nascent years were grim, what followed was a century of some of the most potent labor advocacy organizations in modern history [59, 103]. Even today, the *geist* of the labor union revolution inspires collective action and worker empowerment around the world. Recently, in India, workers across the nation engaged in the largest labor strike in human history [1]. If labor advocacy groups can find ways to effect change in on-demand work as some have called for [76], then the future of on-demand labor may follow the same trajectory of worker empowerment that piecework saw.

The history of piecework suggests that the utopian and dystopian outcomes will *both* occur, in different parts of the world and to different groups of people. When piecework largely disappeared in the United States, outsourcing appeared — creating major labor issues around the world. It's entirely possible that we will create a new brand of flexible online career in developed countries, while simultaneously fueling an unskilled decentralized labor force in developing nations. As designers and researchers, this prompts the question: which outcome are we attempting to promote or avoid for whom?

A Research Agenda

Piecework also helps bring into focus the areas of research that might bear the most fruit. We return to the three questions that motivated this paper: 1) What are the complexity limits of on-demand work? 2) How far can work be decomposed into smaller microtasks? 3) What will work and the place of work look like for workers?

While we have arguably outpaced piecework with regard to the limits on the complexity of work, the most complex and open-ended wicked problems [121] remain the domain of older human collectives such as governments and organizations. In addition, we can learn from the piecework literature as it relates to the stymieing effect that mismanagement has on workers; research into complexity limits should focus on finding new ways to manage workers, in particular using humans (perhaps other crowd workers) to act as modern "foremen".

Piecework researchers looking into decomposition pointed out long ago that piecework is saddled by a lower limit on decomposition: as Bewley mentions, "piecework does not compensate workers for time spent switching tasks" [12]. We've since studied this phenomenon in crowd work both observationally [31] and experimentally [88]. We should consider whether this remains a worthwhile area to explore; unless the work we put forth directly affects the costs of task-switching — for instance, the cost of suboptimal task search, or the cognitive burden of changing tasks — we may only make incremental advances in micro-task decomposition. When the cognitive

cost of understanding a task and its inputs outstrips the effort required to complete it, decomposition seems a poor choice.

Finally, we turn to the relationships of crowd workers. The crowd work literature here can convincingly speak back to the piecework scholarship perhaps more than in the other sections. The tools that are available to us today — not just technical, but *methodological* — make it possible to discover, study, and partner with crowd workers in ways that were unimaginable to piecework researchers. Bigam engages in crowd work [13] not just because it's possible, but because our community appreciates the importance of approaches such as participant-observation and ethnography as a whole [112].

We should also take a moment to explore the opportunity to discuss the ethics of on-demand labor, as Williamson does [157]. The literature on the history of piecework does not frame the question as *whether piecework is inherently ethical or unethical*, instead asking *what conditions render it exploitative*. The literature we have brought to bear suggests that exploitation occurs when conditions harm workers directly or indirectly, such as in sweatshops and agricultural work with pesticides, or where employers systematically underpay or overwork laborers by contemporary standards.

The question then is whether given socio-technical infrastructures systematically harm, underpay, or overwork workers. For example, Amazon Mechanical Turk does not directly require any rate of payment for work, but its design encourages employers to engage in exploitative behavior: piece-rate pay, for example, does not value workers' task search time; additionally, task design interfaces undeniably frame workers as unreliable by recommending replication with multiple workers rather than trusting and paying individual workers more.

Piecework also has lessons for a number of other research questions in crowdsourcing. For example, future work could more deeply explore the evolution of scientific management as it relates to crowdsourcing optimization; quality control approaches between the two eras; and further analysis of incentive structures.

CONCLUSION

On-demand work is not new, but a contemporary instantiation of piecework. In this paper, we reconsider three major research questions in on-demand work using piecework as a lens: 1) What are the complexity limits of on-demand work? 2) How far can work be decomposed into smaller microtasks? 3) What will work and the place of work look like for workers? We draw on piecework scholarship to inform analyses of what has changed, what hasn't, and may change soon. Reciprocally, we believe that modern on-demand work can teach us about the broader phenomenon of piecework as well. If history really does repeat itself, the best we can do is be prepared.

ACKNOWLEDGEMENTS

This work was supported by the Stanford Cyber Initiative and National Science Foundation award IIS-1351131.