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 demonstrated that decompositions of these tasks into workflows can succeed [21, 1, 40, 32]. These decompositions typically take the form of workflows, which are algorithmic sequences of tasks that manage interdependencies [2]. Workflows often utilize a first sequence of tasks to identify an area of focus (e.g., a paragraph topic [21], an error [1], or a concept [49, 50]) 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 [33], brainstorming [37, 48], and accessibility [24, 23, 25].

If decomposition is key to success in on–demand work, the question arises: what can, and can't, be decomposed? Moreover, how thinly can 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 [41, 42]. 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 [17, 26, 43]. In this sense, finer decompositions are seen as more robust — both to interruptions and errors [9] — even if they incur a fixed time cost. At the extreme, recent work has attempted demonstrated microtasks that take seconds [44, 5] or even fractions of a second [22]. However, workers perform better when similar tasks are strung together [26], or chained and arranged to maximize the attention threshold of workers [4]. Despite this, we as a community have leaned *into* the peril of low–context work, "embracing error" in crowdsourcing [22].

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 [1, 19], interpret tasks in different ways [18], and exhibit lower motivation [20] 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" [45], typically by automatically or manually generating higher–level representations for the workers to reflect on [10, 45, 19].

As the additional context necessary to complete a task diminishes, the invisible labor of *finding* tasks [30] has arisen as a major issue. Chilton et al. illustrate the task search challenges on AMT. Workers seek out good requesters [30] and then "streak" to perform many tasks of that same type [11]. Researchers have reacted by designing task recommendation systems [e.g. 12] and minimizing the amount of time that people need to spend doing anything other than the work for which they are paid [6].

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 where there was significant task homogeneity and low fixed costs of machinery and training [3].

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 [3]. 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 [39, 28]. 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 [3]. Instead, managers conducted "stop watch time studies" [31], 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 [14].

The fourth and final stage was narrow expertise training. Even after work is decomposed and measured, there may not be enough 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 [15]. 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 [46]. Where Graves directly implicates measurement as preventing scientific management from being fully utilized, modern crowd work is measuring and modeling every click, scroll, and keyboard event [36, 35]. 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 [16]. In contrast, today system–designers can share, modify, and instantiate environments like sites of labor in a few lines of code [27, 29]. 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 [13, 4, 9, 8, 22].

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 [34, 7]. Third, we used measurement to mathematically optimize workers' behaviors so that we could make the systems more efficient [47].

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 [38]. We might expect to see the rise of systems that scaffold workers into extremely narrow areas of expertise.

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 [26]. 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.