Gamification as a Means of Improving Performance in Human Operator Processes

by

Aaron Alexander Small
B.A. Chemistry, Brown University, 2010

Submitted to the MIT Sloan School of Management and the Institute for Data, Systems, and Society in Partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration
and
Master of Science in Engineering Systems

In conjunction with the Leaders for Global Operations Program
at the
Massachusetts Institute of Technology

June 2017
© 2017 Aaron Small. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author
MIT Sloan School of Management, MIT Engineering Systems Division
May 12, 2017

Certified by__________________________
Steve Spear, Thesis Supervisor
Professor of Management, MIT Sloan School of Management

Signature of Author
MIT Sloan School of Management, MIT Engineering Systems Division
May 12, 2017

Certified by__________________________
Bruce Cameron, Thesis Supervisor
Director, Systems Architecture Lab

Accepted by__________________________
Maura Herson, Director of MIT Sloan MBA Program
MIT Sloan School of Management

Accepted by__________________________
John N. Tsitsiklis, IDSS Graduate Officer
Clarence J. Lebel Professor of Electrical Engineering
Gamification as a Means of Improving Performance in Human Operator Processes

by

Aaron Small

Submitted to the MIT Sloan School of Management and the Institute for Data, Systems, and Society on May 12th, 2017 in Partial Fulfillment of the Requirements for the Degrees of Master of Business Administration and Master of Science in Engineering Systems

Abstract

The Amazon fulfillment center network is the backbone of Amazon's e-commerce business. To achieve higher efficiency and lower cost, Amazon invests heavily in robotic technology. In some buildings, robots automatically store and retrieve shelving units, delivering them to operators who can interact with product at fixed stations. This has greatly increased throughput in buildings with the technology, while adding new constraints. During periods of peak demand, throughput is limited by the number of stations available and the average operator rate at those stations.

This thesis examines how this constraint can be relieved by increasing average operator rate. Time-in-motion studies, video analysis, historical data analytics, and A/B testing suggest that modifications to the station design and operator process do not yield consistent sustainable improvements in performance. Learning curve analysis suggests that operator motivation and engagement are key factors in driving increased performance. Operators perform at a rate of roughly 239 units per hour stowed, with a standard deviation of 48 units per hour. However, operators demonstrate an average maximum sustainable rate of 283 units per hour with a standard deviation of 64 units per hour.

Review of available research on motivation and engagement suggests that gamification methods could be cheaply and easily employed to increase operator motivation and engagement, and have realized 30% improvements in similar manufacturing settings. A cost analysis shows that a similar implementation at Amazon is likely to yield a high return on investment, with a base-case net present project value of over $100 million.

The thesis concludes by describing a custom gamification system for Amazon that could efficiently alleviate the throughput bottleneck for one type of operator station. This approach is not only widely applicable across different process at Amazon, but also similar human operator processes in the manufacturing and warehouse settings.

Thesis Advisor: Bruce Cameron, Thesis Supervisor
Title: Director, Systems Architecture Lab

Thesis Advisor: Steven J. Spear
Title: Senior Lecturer, MIT Sloan School of Management
Acknowledgements

This thesis is a product of countless hours of heading in the wrong direction before beginning the journey of heading in one that might be right.

At MIT I first need to thank my advisors, Bruce Cameron and Steve Spear. Your comments and advice always helped me to look at the project with fresh eyes. I also need to thank the LGO program, both for teaching me everything I know about operations, and for helping us get the most out of the process. And my classmates, for being the only ones who could truly commiserate when the process wasn’t going my way.

A huge group from Amazon deserves thanks for helping me. First, Jeff Messenger, my manager, who let me explore the dead-ends, interesting asides, and fun questions on the way to my thesis. The final product and the experience are better for it. I also owe thanks to the leadership team at MKE1, including Brian Urkiel, Mike Stone, Rob Hahn, and Andrew Bignell, who helped me navigate life as an Amazon intern, entertained me while I dreamed big, and joined me in having some fun at the same time. The team at Amazon Robotics not only listened politely to my annoying questions about the stow system, but always answered them thoroughly. For that I am greatly indebted to John Cohn, Ibi Horvath, Eli Gallaudet, Vishnu Ayyagari, Nathan Parks, and Bill Lynch. Finally, Brian Donato, who looked over my project from Seattle and challenged me to keep pushing for something deeper.

Finally, thanks to my family for helping me feel at home in far-away Wisconsin, for supporting me unconditionally, and for inspiring me to be curious and courageous.
Table of Contents

List of Figures .................................................................................................................. 7

List of Tables .................................................................................................................... 8

1 Introduction .................................................................................................................... 9
   1.1 Amazon Fulfillment ................................................................................................. 9
   1.2 Problem Statement ................................................................................................. 9
   1.3 Project Hypothesis ................................................................................................. 10
   1.4 Project Approach .................................................................................................. 10
   1.5 Thesis Overview ................................................................................................... 10

2 Operations at Amazon .................................................................................................. 11
   2.1 Warehouse Operations ........................................................................................ 11
   2.2 Fulfillment Centers .............................................................................................. 12
   2.3 Amazon Robotics Sortable Fulfillment Centers .................................................... 13
   2.4 Stow at AR Sortable Fulfillment Centers ............................................................. 13

3 System Architecture Solutions Yield Limited Results in Highly Optimized Processes ......................................................................................................................... 14
   3.1 Selecting System Architecture Improvements ..................................................... 16
   3.2 Selected Tests and Results .................................................................................... 19
   3.3 Discussion of Results ........................................................................................... 19

4 Motivation and Engagement Show Serious Potential ................................................. 20
   4.1 Learning Curve Analysis ....................................................................................... 20
   4.2 Interviews and Surveys ......................................................................................... 23
   4.3 Conclusions from Analyses .................................................................................. 24
   4.4 Gamification as a Target ....................................................................................... 24

5 Gamification Literature Review ................................................................................... 26
   5.1 Motivation ............................................................................................................... 26
   5.2 Gamification ........................................................................................................... 27
   5.3 Gamification in the Manufacturing/Material Handling Environment .................... 29

6 Potential Impact of Gamification .................................................................................. 29
   6.1 Cost Calculation Showing Potential Impact ......................................................... 30

7 Gamification Proposal .................................................................................................. 31
   7.1 Design Framework ................................................................................................. 31
   7.2 Design Elements and Rationale ............................................................................ 32

8 Implementation and Next Steps ................................................................................... 37
   8.1 How Could This be Implemented? ......................................................................... 37
   8.2 What Comes Next? ................................................................................................. 39
### 9 Generalizability and Further Study

<table>
<thead>
<tr>
<th>9.1</th>
<th>Generalizability</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.2</td>
<td>Further Areas of Study</td>
</tr>
</tbody>
</table>

### 10 Conclusions

<table>
<thead>
<tr>
<th>Bibliography</th>
</tr>
</thead>
</table>
List of Figures
Figure 1: Amazon Fulfillment Network Schematic ......................................................... 11
Figure 2: Comparison of Fulfillment Center Types ......................................................... 12
Figure 3: Robot Drive and Shelving Pod ....................................................................... 13
Figure 4: Simplified Stow Process Diagram .................................................................. 14
Figure 5: Visual Representation of Common Stow Performance Inputs ....................... 15
Figure 6: Pareto of Top Five Non-Value-Added Activities ............................................ 18
Figure 7: Top Stow Error Messages by Resolution Time ................................................ 18
Figure 8: Aggregate Learning Curve for All Stowers .................................................... 21
Figure 9: Rate Distributions at Various Points in Stow Learning Curve ....................... 21
Figure 10: 4-day Moving Average Rates Versus Cumulative Hours Worked for Top Per... 22
Figure 11: Distribution of Average and Maximum Stow Rates ...................................... 22
Figure 12: Distribution of Average/Max Stow Rate ...................................................... 23
Figure 13: Operators with 4-day Moving Average Rates Far in Excess of Their Average ....... 23
Figure 14: Stow Operators Survey Results - “What Things Would Most Motivate You to Be a Better Stower?” ............................................................. 24
Figure 15: Target Increase Versus Current Average Rates ............................................. 30
Figure 17: Performance Graph Mockup at User Station ................................................. 33
Figure 18: Performance Graph Mockup in Web Portal .................................................. 33
Figure 19: Streak and Lifetime Achievement Mockup in Web Portal ......................... 34
Figure 20: Badges Mockup in Web Portal ..................................................................... 34
Figure 21: Individual Competition Mockup at User Station ......................................... 35
Figure 22: Team Competition Mockup at User Station ............................................... 35
Figure 23: Avatar Mockup in Web Portal ..................................................................... 36
Figure 24: Compelling Customer Story Mockup at User Station .................................. 37
Figure 25: Gamification Implementation Roadmap ....................................................... 38
List of Tables
Table 1: Simplified Fulfillment Process .................................................................12
Table 2: Stow System Input Categories ...............................................................16
Table 3: Sample Opportunity Tracker ................................................................17
Table 4: Summary of System Modification Test Results ......................................19
Table 5: Example Employee Rate Data Format ..................................................20
Table 6: Types of Formal Incentives ..................................................................25
Table 7: Psychological Needs and Their Corresponding Game Design Elements ....27
Table 8: Summary of Gamification Impact in Warehouse Setting .......................29
Table 9: Estimated Cost Savings Under Different Scenarios ..............................30
Table 10: Sensitivity Analysis for Project Cost .....................................................31
Table 11: Psychological Needs and Their Corresponding Game Design Elements .....32
1 Introduction

1.1 Amazon Fulfillment

Amazon.com is the world’s largest retailer, generating over $100 billion in revenue in 2015 (Amazon.com 2016). Founded in 1994 by Jeff Bezos to sell books online, Amazon has expanded over the last twenty-three years to offer the largest product variety of any retailer. They have entered diverse businesses outside of retail, including electronic devices, media and content, and web services. Currently, e-commerce retail and order fulfillment remains the core of their business. This core is supported by a massive global network of distribution centers (known internally as fulfillment centers), cross docks, receive centers, and other types of warehousing facilities totaling over 100 million square feet (Amazon.com 2016). Inventory is sourced from a variety of major and minor wholesale distributors. Through their “Fulfillment by Amazon” network, they also store and ship inventory owned by third parties, including many small suppliers. While Amazon has the largest e-commerce fulfillment network in the world, they still have significant room for growth. As they expand in the US and overseas, they will continue to look for ways to improve throughput and lower cost in their fulfillment network.

1.2 Problem Statement

Amazon’s competitive advantage on their core e-commerce business comes in large part from their ability to reduce the cost of order fulfillment. Order fulfillment represents the cost of delivering a product to a customer, excluding of the cost of goods sold. Components include shipping of the product to a warehouse, storage of the product, order picking, packaging, and shipping to the final destination. A majority of the cost of fulfillment is incurred inside fulfillment centers, which are large warehouses where Amazon stores inventory that is available for sale, and picks and ships items that have been ordered. One way Amazon has kept costs low is by investing in massive, technologically advanced fulfillment centers that speed and automate order fulfillment. The latest generation fulfillment centers use large robotic inventory floors that efficiently route shelving units to human operators at fixed stations. Robotic inventory has greatly increased productivity and lowered VCPU, but has created a new set of operational challenges.

In buildings without robotic inventory, Amazon can increase throughput during times of peak demand by hiring additional workers. Throughput in those buildings is effectively constrained by the number of people that can fit inside. However, the robotic system has a fixed number of operator stations where humans can interact with inventory. Once all stations are occupied, additional throughput can only be achieved by improving the operator rate at each station.

As the robotic inventory system and surrounding processes become ever more optimized, Amazon continues to search for innovative solutions that squeeze more performance out of each station.
This thesis focuses on finding a solution to this problem for one of the major processes where human operators interact with inventory in the robotic inventory system, known internally as ‘stow.’ The definition of stow and its place in the overall fulfillment process will be discussed in Section 2 of the thesis. The goal was to find a solution that would yield a large throughput improvement opportunity for the stow process at a relatively high return on investment for the company. An artificial minimum target of 10% improvement was set at the outset of the project, measured as an increase in average operator stow rate. Average rate is calculated as the number of units stowed in a building for a given time period, divided by the number of stow operator hours worked during the same period. It is expressed in units per hour or UPH.

Using current average operator stow rate, all-in hourly labor cost, and the typical number of units stowed per year in an Amazon Robotics building, a 10% rate increase would translate to a per-building annual savings of $1.75 million in labor costs. At the time of this thesis’ writing, Amazon had six such buildings, with plans to open dozens more over the next several years. If the process improvements are translatable to the analogous process of pick, the savings would be roughly doubled in each building.

1.3 Project Hypothesis
The hypothesis of the thesis is that increasing the intrinsic motivation of human operators is an untapped, high-potential area for throughput improvement that can be realized quickly, cheaply, and easily by adding game-based elements to the operator interface.

1.4 Project Approach
The project started with a kaizen approach. A series of hardware, software, and business process modifications were selected and tested in the warehouse environment. However, the results of these tests revealed only small improvements to throughput at significant cost.

A search for additional areas of opportunity included an analysis of employee learning curve, employee surveys, and one-on-one interviews. All strongly suggested that employee motivation remained a largely untapped lever for driving increased throughput.

Based on these findings, a comprehensive literature review was performed on employee motivation. This review revealed the power of gamification as a tool for increasing worker motivation. A gamification system was proposed based on the latest gamification research and tailored specifically to the Amazon environment. Finally, an implementation plan was created to provide a roadmap for building this system into the Amazon fulfillment network.

1.5 Thesis Overview
This thesis starts with background on Amazon’s fulfillment network, specifically focusing on Amazon Robotics fulfillment centers. It then provides details on stow, the specific process under study.
Next, it reviews the results of several tests related to hardware, software, and business process modifications and proposes why no significant improvements were achieved. It describes a series of analyses that were performed to find other areas of opportunity. In doing so, it makes the case that worker motivation is a fertile area for throughput improvement.

This is followed by a literature review of worker motivation and gamification. This section demonstrates that gamification is an extremely good target for improving intrinsic motivation. An attempt is then made to demonstrate that 10% or greater throughput improvement is potentially achievable using these methods, but that even much smaller improvements should yield positive returns on investment.

Next, a gamification system is described, tailored specifically to the stow process at Amazon. An implementation plan is included, showing how this could best be rolled out at Amazon.

The thesis concludes with a discussion of the generalizability of the gamification approach and recommends additional areas of study.

2 Operations at Amazon

2.1 Warehouse Operations

Amazon.com operates a global distribution network, the cornerstones of which are its fulfillment centers (FCs). Fulfillment centers range in size from about 100,000 square feet to over 1,000,000 square feet and serve as the storage point for all Amazon-owned inventory and a significant proportion of 3rd-party “Fulfillment by Amazon” inventory.

Fulfillment centers are supplied with product either directly from a vendor, or via a crossdocking facility that accepts vendor shipments and routes them, using Amazon-owned freight, to the FC network as a means for balancing inventory. A schematic of the network is shown in Figure 1.

![Figure 1: Amazon Fulfillment Network Schematic](image)
FCs are split into several functional areas that mirror the fulfillment process: receive, stow, pick, pack, and ship. A process schematic and description is shown in Table 1 below.

<table>
<thead>
<tr>
<th>Receive</th>
<th>Stow</th>
<th>Pick</th>
<th>Pack</th>
<th>Ship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incoming product from vendors and third party sellers is virtually checked into the building and processed to ensure quality.</td>
<td>Products are placed onto inventory shelves for storage. Once they are placed on shelves, they are available for sale on Amazon.com.</td>
<td>After a customer order is placed, product is retrieved from the appropriate shelf and placed it in a container to be sent to the next process step.</td>
<td>Customer orders are packed into cardboard boxes, with dunnage, and if required a gift receipt and message.</td>
<td>Packages are labeled with the correct address and postage. They are routed to a loading dock where they are placed into trucks for delivery to customers or to a downstream facility.</td>
</tr>
</tbody>
</table>

Table 1: Simplified Fulfillment Process

In between these steps are some key handoffs. With the exception of the connection between receive and stow, in which products are transported on pallets to stow stations, all other steps are linked by a complicated network of automatic conveyance that ensures incoming product, picked product, or packed orders are routed to the appropriate next step.

Since all products must be placed on an inventory shelf before they are available for sale on Amazon.com, during peak demand times the critical processes of stow and pick tend to become bottlenecks for the process.

2.2 Fulfillment Centers

There are four types of Amazon Fulfillment centers which are shown in matrix in Figure 2 below. The two dimensions are sortable versus non-sortable, and traditional versus robotic.

Figure 2: Comparison of Fulfillment Center Types
Sortable fulfillment centers process products that are less than 18" in their longest dimension. This makes it possible for all products to fit on standardized inventory shelves. Non-sortable fulfillment centers process larger or irregularly shaped products.

Traditional fulfillment centers use fixed shelving aisles where all inventory is stored. Operators walk up and down the aisles with carts and hand scanners, locating specific shelves to stow or pick products from. Robotic fulfillment centers store inventory on moveable shelves, called pods, that are transported by robotic drives to stations on the perimeter of the inventory field.

This thesis limits discussion specifically to Amazon Robotic sortable fulfillment centers, the most highly automated fulfillment centers in Amazon’s network.

2.3 Amazon Robotics Sortable Fulfillment Centers

The set of process steps at Amazon Robotics sortable fulfillment centers is similar to those at traditional sortable fulfillment centers. What distinguishes these FCs is how inventory is stored and interacted with by human operators. An Amazon Robotics inventory field consists of hundreds or thousands of steel and fabric-framed inventory shelving pods, roughly 4'x4'x8'. These pods have shelf space on all four sides that can accept products varying in depth from 6" to 18". The pods are packed tightly in a grid, with aisles left for ease of movement. The pods are moved by robotic drives that are able to position themselves underneath a pod, raise it off of the floor, and drive it to a new location as shown in Figure 3. Generally, the pods are driven to the perimeter of the inventory field where operators can either place inventory on or retrieve inventory from the shelves.

![Figure 3: Robot Drive and Shelving Pod](image)

2.4 Stow at AR Sortable Fulfillment Centers

Within Amazon Robotics sortable fulfillment centers, this thesis focuses specifically on the stow process. Stow takes inbound products that have been received into the building, and uses human operators to place them onto robotic inventory shelves. The products are then available for sale on Amazon.com website and ready to be picked when needed.
The stow process takes place at a stow station like the one shown in Figure 4. Pods with available space for products are driven to the station by robotic drives. The station has a metal rack for staging products that are ready to be stowed. These products are usually contained in yellow plastic containers known as totes. The station also includes a handheld scanner that the operator uses to scan totes, products, and inventory shelves (known as bins). A screen positioned above the station indicates where the operator is in the workflow and lets them know when there are errors.

There are four main steps to the stow process: scan a tote, scan an item, place item in bin, scan a bin. The three scans communicate to the system where an item is coming from, what the item is, and where the item is going to. This ensures inventory accuracy is maintained by Amazon’s inventory management system. A process visual is shown in Figure 4.

3 System Architecture Solutions Yield Limited Results

The natural place to look for improvement opportunities in complex and highly automated system is changes to the system architecture. The stow system can be conceptualized as six interacting components. Those are the user interface, the robotic system, upstream processes, station hardware, standard work, and labor management.

The user interface is everything contained in the logic and display of the screens that operators use during the stow process. These screens provide instruction and feedback to the operator and allow them to interact with the system by selecting setting and options. For example, if an operator experiences a problem with a product, they can summon help from a manager using a menu on the screen.
The robotic system includes the software and hardware that comprise the robotic inventory floor. Components include the robotic drives that transport inventory, the shelving pods that hold inventory, and the software that directs the movements of the robots. An example of changes that can be made to this system are modifications to the algorithm that decides when pods are empty enough to be brought to a stower, or changes to the hardware of the robots that makes them operate more quickly.

Upstream processes include choices that are made at preceding process steps that can affect stow. Examples include the types of containers that the products come in for stowing. Yellow plastic totes contain fewer products than large carts, but carts tend to contain more quality errors. These decisions have direct impact on the rate at which products can be stowed.

Station hardware includes the physical components of the station where the operator works. The rack where products are staged before stowing, the ergonomic mat where stowers stand, and the ladder used to stow products to upper shelves are all examples of this hardware. Changes could include modifications to the rack to reduce wasted steps, the addition of platforms to make stowing to higher shelves easier, and ergonomic improvements that reduce fatigue.

Standard work is the set of procedures that stowers are trained to follow to complete the stow task. Currently, stow is broken into the four steps discussed in the previous section. However standard work also includes many smaller tips, such as how to effectively scan for empty space, where to place and not place products, and how to arrange your work station. Changes could include new instructions that add or remove steps or guidelines from the process.

Finally, labor management includes all the processes that Amazon uses to hire, train, and motivate stowers. Training at the during a worker’s tenure, performance feedback, and motivation techniques all fall into this category. Examples of changes include raising the bar for hiring, changing the compensation structure, or changing the way managers communicate with stowers.

All these system components are summarized below in Figure 5 and Table 2.
3.1 Selecting System Architecture Improvements

Several system architecture changes were selected and tested in the warehouse to evaluate their effect on throughput. The method for selecting these changes relied on several analytic techniques, including solicitation of opportunities from managers and stowers, a video activity study, and an analysis of error messages received at the station. These methods are summarized in the sections below.

3.1.1 Qualitative Input from Managers and Stowers

Informal interviews were held with dozens of employees, both managers and stowers, to solicit ideas for improvement to the current process. All ideas were tabulated into a master list, and ideas were categorized by the impacted system component. Each was given a score on two dimensions. First was a scope score that indicated the difficulty of implementing such a change. The second was an impact score that indicated the
perceived impact of the change. Both scope and impact scores were determined qualitatively based on the judgement of the managers and stowers who recommended the change. Nearly 70 ideas were collected in total. A sample of the master opportunity tracker is shown in Table 3 below.

<table>
<thead>
<tr>
<th>System Component</th>
<th>Opportunity</th>
<th>Proposal</th>
<th>Scope</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Station Hardware</td>
<td>Ladder</td>
<td>Add sliding ladder to station, which is easier and faster for stowers to access, reducing time spent with ladder and increasing stows to higher bins.</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Station Hardware</td>
<td>Platform</td>
<td>Add a raised stowing platform to allow for easier stowing of top bins without a ladder.</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Robotic System</td>
<td>Template colors</td>
<td>Color different size templates in different colors to help stowers quickly identify which templates are arriving next.</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Robotic System</td>
<td>Template mix</td>
<td>Optimize template mix for product mix at FC.</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>User Interface</td>
<td>Time display</td>
<td>Add a toggable time display so associates know when to go to break, lunch, and end of shift.</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Labor Management</td>
<td>Time off task</td>
<td>Minimize time off task through proper tracking and building design. Consider things like breakroom and bathroom placement.</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Station Hardware</td>
<td>Tote height</td>
<td>Modify height of UNEX rack for more ergonomic stowing to power zone.</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Standard Work</td>
<td>Tote placement</td>
<td>Allow for tote placement on the floor and stacking on top of totes in the first slot</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Labor Management</td>
<td>Training</td>
<td>Implement training that accelerates the learning curve for stowers.</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3: Sample Opportunity Tracker

3.1.2 Video Activity Study

Every stow station has a security camera positioned above it. This allows for observation of stower behavior without influencing due to observation effects. A time-in-motion study was designed around this security footage in order to identify the proportion of time stowers were spending on various non-value-added tasks. To systematize the process, a set of activity buckets were defined for the stow. A coding system was used to categorize stower behavior into these buckets and time stamp it appropriately. 9 hours of stowing was observed at randomly selected stations and times across all shifts. The activities and time stamps were recorded and the resulting data was analyzed.

The Pareto chart below in Figure 6 shows the five most commonly observed non-value added activities. Note that value added time is not shown here, but comprises about 70% of overall time. 30% of time is lost to non-value added activities.

Top time-wasting activities included:

1. Rearranging containers, which refers to the time stowers spending sorting yellow totes at their station to position the right products closest to the robotic inventory pods.
2. Dropping/emptying containers, which refers to the process of indicating the system when a yellow tote has been completely emptied and moving that tote to the side.
3. Resolving exceptions, which refers to dealing with error messages that the user interface presents to the stower. Error messages can arise for a variety of reasons, as discussed in the next section.

4. Break away from station, which refers to stowers leaving their station. It is impossible to tell from the video feed what the reason for the break was.

5. Using the ladder, which refers to retrieving and using a ladder to stow products to higher shelves.

---

3.1.3 Error Message Analysis

Based on the fact that error messages appeared as a large non-value-added time component in the above analysis, additional analysis was used to determine the relative frequency of different error messages and their median resolution time. A script was written that scraped data logs from all stow stations within one warehouse and analyzed the data to determine the types of error messages that were being triggered and the time it took for stowers to resolve them. The final result was an ordered list of error messages by total time lost (the product of frequency and average resolution time). Figure 7 shows the top five error message types.
3.2 Selected Tests and Results

As a result of the previous analyses, some system changes were selected for further testing. The goal was to identify changes that were both quickly implementable and had the potential to have a large impact on stow rates. The following four changes were selected:

1. Platform Stow: Addition of a raised platform to allow operators to more easily reach all shelves
2. UI Changes: Modification to the operator user interface to consolidate features to a single screen and reduce interaction time
3. Automatic Containers: Addition of an automatic container detection system to reduce the time that operators spend manipulating product containers
4. Angled Stations: Modification of the station layout to make product containers more accessible to operators and reduce time spent walking

These four changes were made to a set of stations within one of Amazon’s fulfillment centers. Operator rate was recorded at each of those stations, and compared with control stations that did not have the modification. A single sample is considered to be one stower working for a shift, which is roughly equal to eight hours of work. At the end of the study, the mean difference between the two groups was taken and the difference was tested for statistical significance. The results of these system modifications were not promising. As shown in Table 4, results were either not statistically significant, or decidedly negative.

<table>
<thead>
<tr>
<th>Test</th>
<th>Study Date Range</th>
<th>Test Samples (stower*days)</th>
<th>Control Samples (stower*days)</th>
<th>Mean Difference</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform Stow</td>
<td>5/15/16 - 5/30/16</td>
<td>72</td>
<td>122</td>
<td>+12 units/hour</td>
<td>0.058</td>
</tr>
<tr>
<td>UI Changes</td>
<td>4/8/16 - 4/17/16</td>
<td>67</td>
<td>42</td>
<td>+10 units/hour</td>
<td>0.21</td>
</tr>
<tr>
<td>Automatic Containers</td>
<td>5/2/16 - 5/12/16</td>
<td>74</td>
<td>70</td>
<td>-29 units/hour</td>
<td>0.005</td>
</tr>
<tr>
<td>Angled Stations</td>
<td>5/18/16 - 5/7/16</td>
<td>38</td>
<td>55</td>
<td>+4 units/hour</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Table 4: Summary of System Modification Test Results

3.3 Discussion of Results

Why were there no obvious improvements from system changes? One possibility is that the right system changes have not yet been tried. This is certainly possible, and there may remain many undiscovered changes that yield improvements in throughput. However, despite the tests described above, and approximately a dozen other improvement initiatives that took place on the stow process in a similar time period, none produced large and consistent improvement in average stow rate.

Another possibility is that the robotic stow system currently in place at Amazon is effectively at a local performance maxima. While a drastic overhaul of the system could produce a step change in performance, small tweaks are more likely to move the system away from optimal performance that towards it.
4 Motivation and Engagement Show Serious Potential

All the discussion up to this point has largely ignored the most critical aspect of the system: the human operator. Much effort has been placed into optimizing the standard work they follow, the training they receive, and the ergonomics of the system they work in. However, relatively little thought has gone into the psychology and motivation required to perform their assigned task. Given that the system modifications proposed thus far did not yield measurable results, a closer look at the human side of the equation is warranted.

In order to get a sense of how motivation and engagement might affect performance, an analysis of stow rate data was done to understand what other factors might affect rate, and how stowers perform at various points in their learning curves.

4.1 Learning Curve Analysis

A large dataset containing rate data for every stow operator working at three Amazon Robotics sortable warehouses was used as the basis for analysis. These three sites were chosen because data was available from the date they opened until June, 2016 when the data was retrieved. Example data is shown in Table 5.

<table>
<thead>
<tr>
<th>Building Code</th>
<th>Unique User ID</th>
<th>Workday</th>
<th>Shift</th>
<th>Hours Worked</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPA1</td>
<td>exampleuid1</td>
<td>2016-06-25</td>
<td>Days</td>
<td>7.3</td>
<td>341.2</td>
</tr>
<tr>
<td>MKE1</td>
<td>exampleuid2</td>
<td>2016-06-27</td>
<td>Nights</td>
<td>6.8</td>
<td>272.8</td>
</tr>
<tr>
<td>BWI2</td>
<td>exampleuid3</td>
<td>2016-06-29</td>
<td>Nights</td>
<td>4.5</td>
<td>185.2</td>
</tr>
<tr>
<td>BWI2</td>
<td>exampleuid4</td>
<td>2016-07-01</td>
<td>Nights</td>
<td>5.3</td>
<td>199.8</td>
</tr>
</tbody>
</table>

Table 5: Example Employee Rate Data Format

This data allowed for a detailed analysis of stower performance, as hours worked and dates could be correlated to build a picture of individual stower performance over time. This data could also be aggregated to understand how stowers perform on average. For all subsequent analysis, data points where a stower worked for less than 4 hours or the daily rate was less than 75 units per hour were excluded to ensure that partial workdays and stations accidentally left logged in did not skew the results.

Figure 8 shows average rate across all stowers as a function of cumulative hours worked in stow. The data show that, on average, stowers improve consistently until about 700 cumulative hours worked before plateauing.
However, this aggregate data hides underlying variability. Figure 9 shows distributions of stower rates at various points in the learning curve. For example, stowers with 10 cumulative hours of stowing (less than two days of work), have an average rate of 179.1 units per hour, with a standard deviation of 58.1 UPH. Stowers with 500 cumulative hours of stowing (translating to about 4 months of experience) have an average rate of 248.2 units per hour, with a standard deviation of 65.5 units per hour. This standard deviation holds roughly constant throughout the learning curve.

Note that all these distributions have a rightward skew, as do distributions of lifetime average rate discussed below. Investigating the right tails of these distributions could give insight into how to increase performance.

Figure 10 shows 11 stowers who maintained a 4-day moving average rate greater than 350 units per hour after they had had 400 hours of stow experience. Notably, one operator was able to achieve a 4-day moving average rate of nearly 700 units per hour. These are not one time spikes, but represent consistently high performance. In a system where 10% improvements are difficult to find, certain workers are capable of consistently outperforming the average by 100% or more.
Some employees may not be capable of achieving such high rates, due to factors like limited mobility, height, or physical fitness. While this is true for some stowers, evidence from learning curves suggests this is not generally the case. Figure 11 below compares the distribution of lifetime stow rates for stowers (excluding their first five weeks, when they are still learning the process) with the distribution of their maximum four-day moving average stow rates. The mean difference between average stow rates and maximum sustained stow rates is 44.3 units per hour. Approximately 5% of stowers achieve sustained peak rates of over 400 units per hour.

Figure 12 shows the distribution of average stow rates as a fraction of maximum stow rates. On average, after their first five weeks, stowers stow at about 85% of their demonstrated maximum rate.
Figure 13 highlights this fact by showing five operators (chosen from many more examples), who achieved 4-day moving average rates far in excess of their overall average performance. Interestingly, these spikes rarely happen at the end of an operator’s tenure, indicating that they are not the results of breakthroughs in technique or experience.

4.2 Interviews and Surveys

To add qualitative detail to this analysis, I interviewed the 10 highest performing stowers at one fulfillment center to understand why they stow faster than average. While different and sometimes contradictory reasons were proposed, the similarity between them is that they consistently observed their own performance data. Here is an example quote:

"I always check to see if my name is on the board [of top stowers]. If I don’t see myself in the top five, I know I need to step it up the next day."

Survey results from a randomly selected set of 24 stowers suggest that, in general, stowers are motivated by the same things, with increased access to performance data being number one. Some less popular but commonly cited reasons include receiving more vouchers, known as ‘swag bucks’ or ‘vendor bucks’ that can be redeemed for
merchandise at the company store. These vouchers are given out at the discretion of managers for good performance. While the sample size of 24 is limited, it points towards potential areas of opportunity for increasing stower motivation.

Figure 14: Stow Operators Survey Results - “What Things Would Most Motivate You to Be a Better Stower?”

4.3 Conclusions from Analyses

From the above analyses, we can conclude several things:

1. Operator experience is not the only factor determining average rate throughout learning curve. Motivation and engagement also play a role, and they may wane over time.
2. The station design is not the rate-limiting component of the stow system, since many operators are capable of sustaining rates as high as two times the average rate.
3. Many operators have large performance reserves, and do not perform near their demonstrated maximum sustainable rate.
4. Operators that achieve rates far in excess of average are highly motivated by access to knowledge about their own performance.

These four conclusions suggest that finding ways to raise operator motivation and engagement could unlock large performance improvements. Additionally, the final conclusion suggests that a powerful way to do this is also extremely simple: provide operators with access to data about their own performance.

4.4 Gamification as a Target

A suitable target for a rate-improving intervention should be low cost, easy to implement and administer, and applicable to all stowers. There are many ways of influencing operator motivation, but few satisfy all of these constraints.

4.4.1 Problems with Formal Incentives

When looking for ways to motivate employee performance, an obvious starting point is formal incentive programs. Formal incentive programs include things like direct monetary compensation, pay-for-performance, promotions, job perks, and non-monetary compensation. These methods are summarized in Table 6 below.
These methods can be effective, and many are already used in one form or another at Amazon. However, formal incentive programs come with problems that make them unattractive as a first choice for tackling the problem.

Direct monetary compensation and pay-for-performance have the obvious downside of being expensive. Cost savings from improvements in operator rate come primarily from the corresponding reduction in labor cost necessary to complete a given amount of work. By directly increasing labor costs, you cut into the benefit of any performance gains. A ten percent increase in pay may only result in a five percent performance boost, making the whole endeavor a losing proposition. While quantifying this tradeoff is hard difficult (and is dependent on the specific features of the proposed program), the fact that such a tradeoff exists puts these interventions at a disadvantage.

In comparison, promotions and job perks have the advantage of being inexpensive. However, they run into problems of scale. While stowers already have the opportunity be promoted to roles including process guide, process assistant, and even area manager, the ratio of these roles to entry-level roles is low. For example, each process assistant may be responsible for as many as 25-50 stowers. Both stowers and managers understand that promotions are rare, and cannot be used to consistently motivate anyone but the very top performers in a group.

Finally, non-monetary compensation is an approach that seems to score well on all three dimensions. Unsurprisingly, it is already widely used at Amazon FCs. However, its ability to produce consistent performance gains across a large population is limited. The gifts and prizes that Amazon gives out simply do not have enough value to motivate most employees.

4.4.2 Gamification as an Alternative

Significant research and high-profile implementations have already identified other tools that are cheap, easy to implement, and have been proven effective at influencing task motivation for large and varied populations. The next section investigates the possibility that using game elements in non-game settings can influence users’ motivation to perform tasks without the need for formal incentive structures. This
approach is called gamification, and the remainder of this thesis makes the case that it is an ideal method for achieving large rate improvements in stow.

5 Gamification Literature Review

5.1 Motivation

There are many competing theories of human motivation. This thesis will focus on the widely-accepted self-determination theory (Ryan and Deci 2000). Not only does it have a history of experimental support, it is readily applicable to the concept of gamification discussed below.

This theory divides motivation along intrinsic and extrinsic lines. Extrinsic motivation is driven by factors outside the individual, like pay, work contracts, and social pressure. On the other hand, intrinsic motivation is based on three internal psychological needs: the need for competence, the need for autonomy, and the need for social relatedness.

The need for competence refers to feeling successful in deliberately influencing the environment. The need for autonomy refers to having the freedom and desire to complete a task of one’s own choosing. The need for social relatedness refers to feeling like one belongs and participates in a group of significant others. The more that these three basic needs are met, the more likely an individual will be intrinsically motivated to complete a task.

5.1.1 Why It Matters for Stow

Extrinsic motivators are widely used by companies to motivate their employees. Examples include salaries, mandatory work hours, and benefit packages. Significant research exists to support the fact that higher pay and higher pay growth increase employee performance and retention (Trevor, Gerhart, and Boudreau 1997). These motivators are effective but also expensive and complex to implement and maintain.

As discussed in Section 4.4.1 and summarized in Table 6, many of these methods have significant barriers to easy and effective implementation. For the methods that are effective and relatively cheap, Amazon is already using them, and continued development is likely to have diminishing returns. Influencing external motivation is therefore not an ideal method for improving throughput in stow.

On the intrinsic motivation side, we’ve seen that highly intrinsically motivated stowers outperform the average, sometimes by factors of 2 to 3. Yet, Amazon currently has no systematic approach for targeting intrinsic motivation.

Additionally, as a job, stow at Amazon Robotics facilities does a particularly poor job of satisfying the intrinsically motivating psychological needs discussed about. The work is extremely repetitive, employees have little freedom to choose their own task, and they are isolated from other workers.

Therefore, influencing intrinsic motivation represents a both a large and untapped opportunity to improve employee performance. They next question is: how can Amazon go about doing that in the fulfillment environment?
5.2 Gamification

5.2.1 What is Gamification?

The idea behind gamification is to apply the intrinsically motivating aspects of video games to non-gaming contexts. The hope is that doing so will make non-game activities more enjoyable and easier to complete for the user. While there is no single accepted definition for gamification, Sailer et al. provide the following: 'the process of making activities in non-game contexts more game-like by using game design elements' (Sailer et al. 2017).

5.2.2 How Does Gamification Motivate?

Gamification attempts to align intrinsic human motivators with some desired outcome. Of the underlying psychological needs of intrinsic motivation, there are game design elements that target all three (Hense et al. 2014; Rigby and Ryan 2011; Michael Sailer et al. 2014).

A list of some key game design elements is shown in Table 7, mapped to the intrinsic psychological needs that they influence.

<table>
<thead>
<tr>
<th>Psychological Need</th>
<th>Mechanism</th>
<th>Game Design Element</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need for competence</td>
<td>Granular feedback</td>
<td>Points</td>
</tr>
<tr>
<td></td>
<td>Sustained feedback</td>
<td>Performance graphs</td>
</tr>
<tr>
<td></td>
<td>Cumulative feedback</td>
<td>Badges</td>
</tr>
<tr>
<td></td>
<td>Cumulative feedback</td>
<td>Leaderboards</td>
</tr>
<tr>
<td>Need for autonomy (decision freedom)</td>
<td>Choices</td>
<td>Avatars</td>
</tr>
<tr>
<td>Need for autonomy (task meaningfulness)</td>
<td>Volitional engagement</td>
<td>Meaningful stories</td>
</tr>
<tr>
<td>Need for social relatedness</td>
<td>Sense of relevance</td>
<td>Teammates</td>
</tr>
<tr>
<td></td>
<td>Shared goal</td>
<td>Meaningful stories</td>
</tr>
</tbody>
</table>

Table 7: Psychological Needs and Their Corresponding Game Design Elements

5.2.3 Does it Work in Practice?

Over the last ten years, gamification has been deployed for a variety of applications, including in education, social networking, health and wellness, crowdsourcing, and sustainability (Seaborn and Fels 2015). In Seaborn and Fels' review, they found that on balance, gamification implementation led to positive results, but with plenty examples of mixed positive/neutral or positive/negative outcomes. Some examples of implementations and results are discussed below.

In healthcare, Cafazzo et al. (2012) used a mobile app with gamification elements to encourage children with Type I diabetes to measure blood glucose levels. They rewarded these behaviors with points in a user-friendly application. As a result, the average frequency of blood glucose measurement increased 50% (Cafazzo et al. 2012).
In education, Denny (2013) added badges to an online multiple choice question and answer portal for students. Badges increased both the number of answers submitted and the duration of engagement. Notably, response quality was not impacted. However, viewing of badges was not uniform across all students who participated, indicating that a variety of motivational factors were at play (Denny 2013).

In social media, Frith (2012) looked at the location based social network Foursquare, which employs badges, points, and leaderboards to encourage users to more frequently use the service. The research suggests that people enjoy these gamification features, but that they can lead to cheating. Some users employ techniques to gain additional points or badges without completing the system-desired behavior. However, when users were presented with surprising information about themselves by the system, it tended to increase engagement. Finally, leaderboards were found to be demotivational when the top spots were perceived to be unattainable (Frith 2012).

Despite strong theoretical grounding and widespread use, it is difficult to assess whether gamification works generally. This is complicated by the fact there is no agreed upon standard definition of gamification or on the theoretical framework that underlies it. That being said, when implemented in the right context, there is strong evidence that it can lead to performance improvements, and learning retention.

A notable example comes from a materials handling process analogous to stow. A study performed by Klevers et al. applied gamification elements to an order picking process in a warehouse environment. Using Deci and Ryan’s self-determination framework as a guide, they applied game elements including points, badges, a high score list, performance graphs, avatars, and a game narrative (Klevers 2017).

A game designed around these elements was trained and displayed in the warehouse. 103 participants were split into a test and control group. The test group significantly outperformed the control group in picking rate and number of errors. They also demonstrated higher fulfillment of the three intrinsic psychological needs based on a Likert scale and free-response format questionnaire. Finally, self-reported intrinsic motivation increased, as reported by participants on a survey.

While these results are encouraging, more study is required to determine whether these findings will apply at Amazon. First, the study was performed on new pickers who were trained for only ten minutes. They were then given twenty minutes to perform the task. The limited duration of time and the fact that the test subjects were new to the task leaves open the question of how it will work with a large population of experienced stowers for an extended period of time.
However, the fact that a simple gamification implementation in an extremely analogous environment to the stow process was able to produce 30%+ improvements in throughput and quality indicates that this is an area worthy of deeper study.

5.3 Gamification in the Manufacturing/Material Handling Environment

5.3.1 Few Attempts to Apply Gamification in Industrial Settings

With the exception of the study described above, essentially no research or empirical work has applied gamification to repetitive labor tasks in the industrial setting. Industrial environments have been slow to adopt user-focused approaches to performance improvement (Korn and Schmidt 2015). Most optimization approaches in material handling focus on the technical aspects of the task, to the exclusion of the human element (Coffey 1999). In industrial applications that use human operators, significant investment is made in human-machine interfaces. However, the focus has been on control of the work, rather than the supporting the worker or making the working more attractive (Korn 2012).

These facts are surprising for several reasons. First, the infrastructure required for gamification is already in place: individualized user interfaces, robust data tracking and reporting, and clearly defined repetitive objectives. Second, these industries rely on the continuous performance of a large body of hourly workers, so the benefits of increased intrinsic motivation are especially high.

5.3.2 Why use Gamification in this Setting?

The lack of gamification in these settings provides a unique opportunity. In an area where the technical aspects of the human-machine interface have been relentlessly optimized, there exist few cheap opportunities for continued improvement. Influencing intrinsic motivation through gamification techniques is an untapped opportunity for large throughput improvement gains.

6 Potential Impact of Gamification

Gamification has clearly demonstrated the potential to influence user behavior, but to what degree it could do so at Amazon? This section will attempt to put bounds on the size of the opportunity.
6.1 Cost Calculation Showing Potential Impact

Historically, after their first five weeks, stowers perform at roughly 85% of their peak throughput (measured as their max 4-day moving average rate). Since our arbitrary target for rate improvement is 10%, visualizing that against the distributions for current average rate and average maximum sustained rate gives a sense of how realistic the goal is. Figure 15 shows that a 10% rate improvement would bring the average rate up to 262.7 units per hour, raising the average as a percentage of maximum from 85% to 92%.

![Figure 15: Target Increase Versus Current Average Rates](image)

How does an improvement of this magnitude translate into cost reduction? Using some basic estimates for the all-in hourly cost of a stower and the number of units stowed in an Amazon Robotics sortable fulfillment center, it is possible to calculate a per-building annual savings for a given rate increase. Assuming there are six Amazon Robotics sortable fulfillment centers in 2017, and that number doubles every year, it is then possible to calculate Amazon-wide annual cost savings. Table 9 summarizes the results of such an analysis. Note that while reasonable assumptions are made below, actual Amazon figures have been omitted to protect trade secrets.

<table>
<thead>
<tr>
<th>Rate Increase</th>
<th>New Average Rate</th>
<th>Percent of Demonstrated Maximum</th>
<th>Per-Building Annual Savings</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5%</td>
<td>244.77</td>
<td>86.5%</td>
<td>$467,378</td>
<td>$2,804,265</td>
<td>$5,608,530</td>
<td>$11,217,061</td>
<td>$22,434,122</td>
</tr>
<tr>
<td>5.0%</td>
<td>250.74</td>
<td>88.6%</td>
<td>$912,499</td>
<td>$5,474,994</td>
<td>$10,949,988</td>
<td>$21,899,976</td>
<td>$43,799,952</td>
</tr>
<tr>
<td>7.5%</td>
<td>256.71</td>
<td>90.7%</td>
<td>$1,336,917</td>
<td>$8,021,503</td>
<td>$16,043,006</td>
<td>$32,086,011</td>
<td>$64,172,023</td>
</tr>
<tr>
<td>10.0%</td>
<td>262.68</td>
<td>92.8%</td>
<td>$1,742,044</td>
<td>$10,452,261</td>
<td>$20,904,523</td>
<td>$41,809,045</td>
<td>$83,618,090</td>
</tr>
<tr>
<td>12.5%</td>
<td>268.65</td>
<td>94.9%</td>
<td>$2,129,164</td>
<td>$12,774,986</td>
<td>$25,549,972</td>
<td>$51,099,944</td>
<td>$102,199,888</td>
</tr>
<tr>
<td>15.0%</td>
<td>274.62</td>
<td>97.0%</td>
<td>$2,499,454</td>
<td>$14,996,723</td>
<td>$29,993,445</td>
<td>$59,986,891</td>
<td>$119,973,782</td>
</tr>
<tr>
<td>17.5%</td>
<td>280.59</td>
<td>99.1%</td>
<td>$2,853,986</td>
<td>$17,123,917</td>
<td>$34,247,835</td>
<td>$68,495,670</td>
<td>$136,991,340</td>
</tr>
</tbody>
</table>

Table 9: Estimated Cost Savings Under Different Scenarios

The conclusion of this analysis is that even modest rate increases, if they can be sustained across all Amazon Robotics sortable FCs, can lead to large cost savings.
This is thanks to Amazon’s size and projected growth. Even the estimates shown above are conservative because they do not assume the possibility of applying gamification to other processes in the FC like pick, or to similar building types like Amazon Robotics non-sortable FCs.

To decide whether it is worth it to pursue such an initiative, we must also consider the magnitude of investment required to build a gamification system. They are many different ways of estimating software development cost. However, a simple approach is to use basic assumptions about the salary of software developers and projects of similar magnitude that have been in the past (Sommerville 2010).

More details on the proposed software development effort are described in detail in the next section. However, based on the author’s experience with similar development efforts that have taken place at Amazon, a project of the size proposed in this thesis could be completed by ten developers working for one year, with four developers required for ongoing system maintenance. Assuming an average developer salary of $120,000, with all-in costs totaling two times the base salary, such a project would require $2.4 million in year 1 and $960,000 in subsequent years.

Using this baseline cost estimate it is possible to perform a sensitivity analysis on the project’s net present value while varying cost overruns and potential rate increases. A discount rate of 8% was used on cash flows for 2017 through 2020. The resulting sensitivity table shows that this project is robust to large underestimates in development costs, even with modest rate increases.

<table>
<thead>
<tr>
<th>Cost Overrun</th>
<th>0%</th>
<th>5%</th>
<th>7.5%</th>
<th>10%</th>
<th>12.5%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>$28,286,192</td>
<td>$59,523,493</td>
<td>$89,307,897</td>
<td>$117,738,464</td>
<td>$144,905,451</td>
<td>$170,891,264</td>
</tr>
<tr>
<td>50%</td>
<td>$26,029,704</td>
<td>$57,267,006</td>
<td>$87,051,409</td>
<td>$115,481,977</td>
<td>$142,648,963</td>
<td>$168,634,776</td>
</tr>
<tr>
<td>100%</td>
<td>$19,250,241</td>
<td>$50,497,543</td>
<td>$80,281,947</td>
<td>$108,712,514</td>
<td>$135,879,500</td>
<td>$161,865,314</td>
</tr>
<tr>
<td>150%</td>
<td>$-1,048,147</td>
<td>$30,189,155</td>
<td>$59,973,559</td>
<td>$88,404,126</td>
<td>$115,571,113</td>
<td>$141,556,926</td>
</tr>
<tr>
<td>200%</td>
<td>$-68,742,773</td>
<td>$-37,505,472</td>
<td>$-57,721,068</td>
<td>$20,709,500</td>
<td>$47,876,486</td>
<td>$73,862,299</td>
</tr>
</tbody>
</table>

Table 10: Sensitivity Analysis for Project Cost

While the exact costs and outcomes of such a project cannot be known at the present time, this line of reasoning makes it clear that it is a worthwhile avenue of exploration. The key factor that ensures low risk is that, as with many software projects, the benefits tend to scale linearly with the user base, while the costs of development and maintenance remain relatively fixed.

7 Gamification Proposal

7.1 Design Framework

Using the intrinsic psychological needs design framework, specific game design elements were proposed that could be effective in Amazon’s warehouse setting. Table
11 shows specific game design elements that are proposed for the Amazon system. The following sections describe these elements in more detail and show examples of how they can be integrated into Amazon’s existing software.

<table>
<thead>
<tr>
<th>Psychological Need</th>
<th>Mechanism</th>
<th>Game Design Element</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need for competence</td>
<td>Granular feedback</td>
<td>Points for volume stowed and streaks without quality errors</td>
</tr>
<tr>
<td></td>
<td>Sustained feedback</td>
<td>Performance graphs of volume, rate, quality errors, time of task and other key indicators</td>
</tr>
<tr>
<td></td>
<td>Cumulative feedback</td>
<td>Badges for exceptional streaks and tenured performance</td>
</tr>
<tr>
<td></td>
<td>Cumulative feedback</td>
<td>Competition and lifetime leaderboards</td>
</tr>
<tr>
<td>Need for autonomy (decision freedom)</td>
<td>Choices</td>
<td>Avatars, ability to opt out of game system</td>
</tr>
<tr>
<td>Need for autonomy (task meaningfulness)</td>
<td>Volitional engagement</td>
<td>Meaningful customer stories</td>
</tr>
<tr>
<td>Need for social relatedness</td>
<td>Sense of relevance</td>
<td>Team battles, team avatars</td>
</tr>
<tr>
<td></td>
<td>Shared goal</td>
<td>Meaningful customer stories</td>
</tr>
</tbody>
</table>

Table 11: Psychological Needs and Their Corresponding Game Design Elements

7.2 Design Elements and Rationale

7.2.1 Points for Volume and Quality

The core of the gamification design is a system of assigning points for key performance metrics. The primary metrics are volume, rate, no-stow turn-away percentage (a measure of how often associates turn away a robotic pod without stowing any items), confirmed quality errors, and time off task. Standard definitions exist for these metrics, and they are widely reported across AR sortable fulfillment centers. Users would have access to this information in a variety of formats, as demonstrated in the mockups below.

7.2.2 Performance Graphs

In order to represent performance to employees, data should be available in a real-time format, accessible during the shift, and a long-term format that is preserved and accessible at any time.

Real-time data can be presented directly at the operator workstation. Long-term information would be available on a web portal accessible from home. The distinction between real-time and long-term information will appear throughout the gamification system design.

Below are mockups of performance graphs given on both the user’s station screen and through the web portal. The time scale of the first mockup is in hours, the second is in weeks.
7.2.3 Badges, Streaks, and Lifetime Statistics

Another satisfier of the need for competence is long-term cumulative feedback. One way this is often employed in game settings is through the use of badges. Badges are earned by completing a variety of tasks, and their collection indicates experience and seniority. Badges need not always be given strictly for increasing performance or tenure, but can also represent unusual or surprising challenges that the user has completed. Below are mockups showing streaks and badges on the web portal.
7.2.4 Competition and Lifetime Leaderboards

Competition not only satisfies the need for competence by supplying cumulative feedback, but team competition in particular adds to a sense of social relatedness. An additional benefit of team competitions is that they encourage everyone to perform at their maximum effort. In an individually competitive format, lower performers have no incentive to compete, since they are not capable of winning. In a team setting, everyone’s effort counts towards the final total, so even the lowest performing stowers play a part in (and reap the rewards) of victory.

Elements of competition are already used informally by managers at AR sortable fulfillment centers. Examples include games like ‘beat your manager’ where a manager will stow for a fixed period of time (usually 1-2 hours) and employees will try to beat their rate. Additionally, since employees already work in smaller groups divided up by sections of the inventory floor, some managers run ‘floor’ based competitions, where all the second floor stow stations will compete against all the third floor stow stations to stow the most number of units in a fixed time period.
A gamification system builds these elements into software and takes burden off managers. It can also add elements that would be difficult to do any other way. For example, head-to-head competitions between individuals are easily organized (perhaps even initiated by the participants themselves). The first mockup below shows a competition between two stowers. The second mockup shows how a manager could set up team competitions (in this case by floor, but other divisions are possible).

**Figure 20: Individual Competition Mockup at User Station**

**Figure 21: Team Competition Mockup at User Station**
7.2.5 Avatars and Opt-Out

Two key features are included in the system design to satisfy the need for autonomy. First, each employee will have a customizable avatar that they create at the beginning of their employment and modify thereafter. Avatars would be visible to the user whenever they logged into a station or the web portal, and would be visible to other employees during competitions or on the leaderboard. Special demarcations, like colors, articles of clothing, and badges could be earned for the avatar so that the visual representation of the user carries indications of status and tenure.

![Avatar Mockup in Web Portal](image)

A second key feature that satisfies the need for autonomy is the ability to opt out of the system. There are serious pitfalls to making the system mandatory. Doing so makes the gamification aspects seem onerous, and may paradoxically reduce engagement. The purpose of a gamification system is to make the activity enjoyable for the user. If the system is failing on that account, forcing people to use it will not improve performance.

Managers should make it extremely clear that participation in the gamified system is complete voluntary. Additionally, it must be stated that the decision not to participate will never adversely affect performance reviews, pay, or promotions.

7.2.6 Meaningful Customer Stories

Without some sort of background story, this system is simply a random collection of game design elements applied to an operator process. Incorporating meaningful stories into the program ensures that users feel there is some driving rationale for their continued play and performance. For example, during competitions, visualizations of Amazon trucks racing towards customers measure relative performance of the teams. The system could close out the day not with displays of how many Amazon deliveries were enabled, and how many customers were reached. In the future, development could include a rough map of the destinations of packages. All of this serves to bring the employee closer to the customer, and therefore adds purpose to the work. The
mockup below shows an end of shift summary report that focuses the user's attention on Amazon's customers.

![End of Shift Report](image)

**Figure 23: Compelling Customer Story Mockup at User Station**

7.2.7 Hardware and Other Considerations

Minimal additional hardware is necessary for this implementation. A CPU and screen exists at every stow station, and software updates can be easily pushed at a pilot or FC level. Hardware for accessing the web portal is supplied by the user (either through a mobile device or PC).

Finally, it is recommended that screens be placed at every inventory floor of the fulfillment center near the stow morning stand up meeting area. Screens should be placed in common areas. These screens will display leaderboards, ongoing competitions, and other notable statistics for the day.

8 Implementation and Next Steps

8.1 How Could This be Implemented?

While this is a very high return on investment improvement, it is not without cost and complexity. In order to minimize unnecessary cost and speed up the software learning curve, implementation should use a multi-step approach with clear decision points and opportunities for modifications between each step. The table below summarizes the four key phases.
<table>
<thead>
<tr>
<th>Phase</th>
<th>Description</th>
<th>Purpose</th>
<th>Participants</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Software Development</td>
<td>Developing prototype gamification software that can be deployed on a pilot basis.</td>
<td>Develop minimum viable product.</td>
<td>N/A</td>
<td>2 months</td>
</tr>
<tr>
<td>Alpha Test</td>
<td>Testing basic software functionality with a small group of employees in a highly controlled setting.</td>
<td>Address qualitative design concerns and eliminate bugs</td>
<td>&lt;50</td>
<td>2 Months</td>
</tr>
<tr>
<td>Beta Test</td>
<td>Testing refined gamification tool in a more realistic warehouse environment.</td>
<td>Demonstrate improvement in throughput.</td>
<td>300-1000</td>
<td>2 months</td>
</tr>
<tr>
<td>Production</td>
<td>Rollout to all AR Sortable Fulfillment Centers.</td>
<td>Fully rollout system to AR Sortable organization.</td>
<td>10,000+</td>
<td>Ongoing</td>
</tr>
</tbody>
</table>

Figure 24: Gamification Implementation Roadmap

The first phase would focus simply on building a minimum viable product that can be demonstrated to work in a limited sandbox, such as a single test station used by engineers and programmers.

Once the software passes initial quality and functionality requirements, the go/no-go decision to enter the next phase should be based on a review of the initial functionality with a cross functional group involving representation from the fulfillment centers (hourly employees and managers).

The second phase would involve an alpha test at a small robotic warehouse fulfilling actual customer orders. Approximately 50 employees would be exposed to the new software. The goal of this test would not be to demonstrate improvement in throughput, although that data would be collected. The goal would be to get qualitative feedback from operators and managers in a real-world environment. Changes at this stage could be made in real-time without compromising a controlled study. At the end of the two-month trial period, the software would be stable and ready for a larger rollout in a more representative facility.

The third phase involves deploying the software at a single AR Sortable fulfillment center, with several hundred stowers. This test would focus on quantifying the impact that the gamification has on throughput. No changes would be allowed during this test to preserve the integrity of the results.

After a two-month test, data should show whether the software system has an impact on throughput and if so to what degree. These results should be directly used to make a go/no-go decision on whether to roll out the software to the entire AR sortable network.
8.2 What Comes Next?
Depending on the results of the gamification system test, there are several things that Amazon will want to consider. First, there are many other processes at Amazon fulfillment centers that are similar to Amazon Robotics sortable stow. If the gamification system shows promise it should be possible with minor modification to port these changes to dozens of processes across robotic and non-robotic fulfillment centers.

9 Generalizability and Further Study

9.1 Generalizability
This thesis focused rather narrowly: the potential to improve throughput in Amazon Robotics sortable fulfillment centers by targeting the stow process path. However, there are much larger implications to the results derived here.

9.1.1 Within Amazon Robotics Facilities
The first and most obvious crossover application is in the pick process path. Like in Amazon Robotics stow, Amazon Robotics pickers operate at fixed stations on the perimeter of the robotic inventory field and perform a repetitive task with many of the same constraints as stowers. Gamification elements could be easily modified and applied to this process path.

Looking outside sortable fulfillment centers, non-sortable fulfillment centers containing products larger than 18" in length face many of these same problems. They also use operators in stow and pick, but because they handle larger products, the average rate in both processes is slower.

9.1.2 Within Amazon More Broadly
Almost all Amazon employees working in the order fulfillment network have jobs that are capable of employing similar gamification concepts. Amazon has a network-wide system that uses screens placed at stations or on handheld scanners to communicate with operators. Significant modification would be needed to tailor the approach described above to process paths like receiving, packing, and truck loading. However, the underlying principles and the expected benefits would remain the same.

9.1.3 Outside Amazon
More generally, the only requirements for using a gamification system like the one proposed in this thesis, are that operators have a task that is easily measureable and that they have access to a device capable of reporting their progress. Therefore, under the broadest possible application of these techniques, employees in fields ranging from manufacturing, maintenance, transportation, and retail could use this approach to increase employee motivation, job satisfaction and performance.

9.2 Further Areas of Study
Based on this thesis there are some key areas of further study that would strengthen the proposal and yield better results for Amazon. Both could take place during the implementation phases of the gamification system.
First, detailed analysis should be done on the effect of specific game elements on performance. Using methods like A/B testing can help isolate which specific game elements are effective in the Amazon context and which are not. This data could be used to build subsequent versions of the system that are more effective.

Second, more research should be done into the possibility of showing the user performance data that combines multiple metrics to achieve a balancing effect between competing priorities. High throughput has tradeoffs, often in the form of reduced quality. Developing and testing metrics that quantify ‘optimal performance’ will help to minimize the problem of an excessive focus on throughput to the detriment of all else.

10 Conclusions
This thesis examined ways to increase throughput in Amazon’s stow process, where human operators work in tandem with a robotic inventory system. The problem with such a system is that throughput is constrained by the average rate that operators can perform at a fixed number of stations. Since the number of stations cannot be easily increased in a given fulfillment center, the only way to increase total throughput is to increase average operator rate at each station.

The hypothesis of the thesis is that unlocking intrinsic motivation of human operators is a very high-potential area for throughput improvement, and using gamification elements is a cheap, easy, and effective way to do so.

First, a series of in-warehouse tests based on station and process modifications showed no significant improvement or even negative results. As a follow-on, learning curve analysis, surveys, and interview data showed that operators maintain large performance reserves and that intrinsic motivation represented an untapped area of opportunity for further exploration. A literature review of motivation research suggested that gamification was an easy way to access this intrinsic motivation for performance improvement.

Based on this finding, a cost analysis was performed to show that not only is the potential upside extremely large for the Amazon Robotic FC network, but also that the development costs are small and likely to be relatively fixed. These two factors ensure that a gamification implementation is likely to have a high return on investment.

A gamification system was proposed applying that tailored the core concepts of gamification and motivation specifically to the Amazon environment. Finally, an implementation plan gives a roadmap for creating such a system in a timely and cost effective manner.
Bibliography


