On the Integration of Card-Allocation and Dispatching Decisions in POLCA Systems: An Assessment by Simulation

Nuno Octavio Fernandes, Matthias Thürer, and Mark Stevenson

Abstract

POLCA is an important card-based control system for low volume, high variety production contexts. A job can only be produced at an upstream station if it has acquired a POLCA card that has returned from its downstream station. A common assumption in the POLCA literature is that cards are allocated to jobs as soon as they return to the upstream station. This dissects the queue in front of a station into jobs that have a card (and can be produced) and those that do not have a card (and cannot be produced). This artificially and prematurely constrains the dispatching decision, i.e., the decision concerning which job to produce next at a station. In response, this paper proposes integrating the card-allocation and dispatching decisions such that the allocation of POLCA cards to jobs is postponed until the dispatching decision is made. Simulation results demonstrate that this integrated approach does not improve performance under simple ERD dispatching, as is commonly applied in the POLCA literature. But when a more powerful rule is applied, percentage tardy and mean tardiness performance improve by more than 75% and 50%, respectively, for an integrated decision. Most importantly, results suggest that in production environments like the one considered in this study, the integrated approach dispenses with the use of POLCA altogether if a suitable priority rule is used.

Keywords: *POLCA; Card-Allocation Decision; Dispatching Decision; Discrete Event Simulation.*

1. Introduction

This paper seeks to clarify how card-allocation and dispatching should be integrated within Paired-cell Overlapping Loops of Cards with Authorization (POLCA; e.g. Suri, 1998, 2010 and 2018). POLCA is a card-based control system of specific importance to low volume and high variety production environments that has been widely implemented in practice (Vandaele *et al*., 2008; Krishnamurthy & Suri, 2009; Riezebos, 2010). It is likely to gain even more importance in practice in the future given that many sectors are evolving towards mass customization, which goes hand-in-hand with lower volumes and higher varieties. POLCA links pairs of stations with overlapping loops of cards to control the movement of material through the shop floor. Only if a POLCA card from a downstream station is available can a job be processed at the upstream station. The POLCA card consequently signals that there is capacity at the downstream station, which authorizes a job to be produced at the upstream station. POLCA consequently does not take an explicit decision on whether to release a job to the shop floor. Rather, its control is restricted to production authorization at individual station queues. This is different from release methods such as Constant Work-in-Process and Workload Control, which only control the release of orders to the shop floor (Thürer *et al*., 2021).

Since its inception, POLCA has received considerable research attention (e.g., Lödding *et al*., 2003; Fernandes & Carmo-Silva, 2006; Germs & Riezebos, 2010; Farnoush & Wiktorsson, 2013; Harrod & Kanet, 2013; Braglia *et al*., 2014; Thürer *et al*., 2017; Severino & Godinho Filho, 2019; and Thürer *et al*., 2021). A common assumption in this literature is that the POLCA card is attached to the most urgent job in an upstream queue as soon as it returns to this upstream station in its loop (e.g., Fernandes & Carmo-Silva, 2006; Riezebos, 2010; Harrod & Kanet, 2013; Thürer *et al*., 2017). Since there is typically more than one card in a POLCA loop, and since there is typically more than one job queuing for capacity at a station, the queue in front of a station is divided into: (i) jobs that have a POLCA card, and therefore have authority to be produced, and (ii) jobs that have no card and are not authorized to be produced. But this restricts the decision concerning which job to produce next at a station, i.e., the so-called dispatching decision (Blackstone *et al*., 1982). In a POLCA system, the dispatching rule can only be used to choose between the set of authorized jobs. Meanwhile, Suri (2010, p., 135) advocated that each station should have a bulletin board where it organizes all the POLCA cards currently available. While this appears to suggest that POLCA cards are only attached when the job is actually dispatched at a station, the decision on which job to allocate a card is executed periodically (daily, every shift). So, there may again be more than one job in the station queue that has received a POLCA card, and the dispatching decision is restricted by the card allocation decision.

It is argued here that POLCA cards should not be assigned to jobs when they arrive at the station (or periodically), but only when the dispatching decision is being executed. Instead of first executing the so-called card-allocation decision (Thürer *et al.*, 2017) and then the dispatching decision, as is typically implemented in the POLCA literature, the two should be integrated such that the dispatching decision is executed followed by a POLCA card being attached to the selected job if available. So, instead of two decisions, i.e., card-allocation and dispatching, only one decision is executed. This decision can be understood as a dispatching decision with the constraint that a card needs to be available. In practice, this can simply be realized by creating a pool of POLCA cards at each station (the bulletin board in Suri (2010)) instead of attaching the cards to waiting orders when a card arrives. Using this procedure, the selection permutations for the decision concerning which job to process next will be significantly extended.

Note that this change is arguably more powerful than the introduction of a specific card-allocation rule, such as in Thürer *et al*. (2017). It integrates the card-allocation and dispatching decisions and postpones their execution to the latest possible moment in time, i.e. when the job is about to enter the station. In contrast, for the procedure typically applied in the literature, the dispatching decision is constrained as soon as the POLCA card arrives from the downstream station, which may be long before the dispatching decision is actually executed. This is a major issue since, for example, a job that is more urgent than the one to which the card was attached may arrive after the POLCA card was attached (and consequently will not receive priority). Discrete event simulation will be used to assess the impact of the proposed design change.

The remainder of this paper is structured as follows. Section 2 outlines how POLCA has been implemented and the simulation model used. Results are then presented and discussed in Section 3. Finally, conclusions are summarized in Section 4, where managerial implications and future research directions are also outlined.

2. Methodology

This study started by asking:

Can POLCA performance be improved by postponing the assignment of POLCA cards to jobs from the point in time when the card arrives at a station to the moment when the dispatching decision is taken?

To answer this question, we explore the performance of POLCA using a simulation model of a general flow shop (Oosterman *et al*., 2000). Make-to-order companies that produce a high variety of products often use a functional layout and operate as some form of job shop. Enns (1995, p.2804) further argued that '*routeing in most real job shops lies somewhere between the pure job shop and pure flow shop extremes.*' This 'in-sequence with bypassing flow' is characteristic of the general flow shop (Aneke & Carrie, 1986), which is the environment that is considered in our study. The model characteristics largely follow Thürer *et al*. (2017) to allow for comparability. We first describe how the shop was modeled in Section 2.1 before we describe how we implemented POLCA in Section 2.2. Finally, Section 2.3 summarizes our experimental set-up and the main performance measures considered.

2.1 Shop and Job Characteristics

A simulation model of a high-variety make-to-order shop has been implemented using ARENA simulation software. Table 1 summarizes the main shop and job characteristics modelled in our study.

[Take in Table 1]

To capture the high routing variability, processing time variability, and arrival variability that defines high variety shops in practice, job routings, processing times, inter-arrival times and due dates are modelled as stochastic (random) variables. The shop contains six stations, where each station is a single constant capacity resource. The routing length varies uniformly from one to six operations. All stations have an equal probability of being visited and a station is required at most once in the routing of a job. Since we consider a general flow shop, the resulting routing vector is sorted so there are typical upstream and downstream stations. Operation processing times follow a truncated 2-Erlang distribution with a maximum of 4 time units and a mean of 1 time unit before truncation. Set-up times are considered as part of the operation processing time. Meanwhile, the inter-arrival time of orders follows an exponential distribution with a mean of 0.642, which, based on the number of stations in the routing of an order (3.5), deliberately results in a utilization level of 90%. Due dates are set exogenously by adding a random allowance factor, uniformly distributed between 30 and 50 time units, to the job entry time. These values were set arbitrarily to result in a percentage tardy that is neither too high nor too low. The percentage tardy should not be too high in order to avoid certain adverse effects, since rules that reduce the variance of lateness across jobs may even lead to an increase in the percentage tardy when due date allowances are too tight on average. Likewise, the percentage tardy should not be too low to avoid our results being affected by incidental effects, as very few jobs would be responsible for the performance of the shop.

2.2 POLCA

This section outlines how we implemented the different elements of POLCA. Section 2.2.1 discusses POLCA's card-based material flow control element. The card-based element of POLCA controls the material flow, i.e., it decides *whether* a job should be authorized for production at a station (Graves *et al*. 1995). It does not decide *which* job should be authorized. The decision concerning which job should be authorized is executed by the card-allocation and dispatching decisions. POLCA's card allocation and dispatching element is discussed in Section 2.2.2. Note that we do not model POLCA's authorization element, which would restrict the number of eligible jobs at a station to jobs for which an earliest release date has been reached, given its direct detrimental impact on performance in Thürer *et al*. (2019). The card-allocation decision is therefore executed continuously whenever a card becomes available (for the separated decision) or whenever an operation is complete (for the integrated decision), and all jobs in the queue are considered.

2.2.1 Card-based Material Flow Control Element

POLCA links the different stations in the routings of jobs using card loops between pairs of stations. Each pair of consecutive stations in the routing of a job has a specific POLCA card that identifies the two stations. These POLCA cards are job anonymous, i.e., they are assigned to station pairs and not to jobs, unlike in the case of Kanban systems (Riezebos, 2010; Ziengs *et al*., 2012). For example, a POLCA 1-2 card is used to signal from Station 2 to Station 1 and *vice versa*.

The card-based element of POLCA is illustrated in Figure 1 for a shop that produces jobs that move from Station 1 to Station 2 and then on to Station 3, following the framework used in Liberopoulos & Dallery (2000) and Riezebos (2010). There are three elements: (i) a queue A_{ik}^{POLCA} that contains the POLCA cards for each control loop *i-k*, where *i* and *k* are successive stations in the routing of the jobs; (ii) a queue PA_i that contains the jobs waiting at each station for a card from queue A_{ik}^{POLCA}, being PA_1 the backlog of the system; and (iii) a queue DA_i that contains the jobs queuing at each station that have received a card from queue A_{ik}^{POLCA}. When a customer places an order, a new job is created and enters queue PA_1 (being station 1 the first in the job routing). The job waits in queue PA_1 until a POLCA 1-2 card is available in queue A_{12}^{POLCA} and it is the highest priority job. Then the POLCA card is attached and the job moves to queue DA_1 waiting to be processed. Once processed it moves to the queue PA_2 of the next station with the POLCA 1-2 card still attached. The job waits in queue PA_2 until a POLCA 2-3 card is available in queue A_{23}^{POLCA}. After processing at Station 2, the POLCA 1-2 card is freed and moves back to queue A_{12}^{POLCA} , and the job moves to the queue of the next station PA_3 with the POLCA 2-3 card attached. Thus, card loops are overlapping since the POLCA 1-2 card is only released after the operation at Station 2 has been completed.

[Take in Figure 1]

In this study two versions of POLCA will be considered: POLCA (i.e. the original system) and POLCA with Starvation Avoidance (SA). Including POLCA SA reflects recent developments in the POLCA literature (Thürer *et al*. 2017) and ensures that arguably the best-performing version of POLCA is included in the experimental design. On some occasions, a station may be starving although there is work in the queue, e.g., when all available POLCA cards that authorize production at that station are at the downstream stations in the loops. This form of premature idleness (Kanet, 1988; Land & Gaalman, 1998) can be resolved by attaching a starvation avoidance card to a job, thereby allowing it to be processed at the starving station (Thürer *et al*., 2017). Using a starvation avoidance card means that the work-in-process cap or limit will be exceeded in the loop. In order to restore the limit, POLCA cards do not become available after being detached from jobs as long as starvation avoidance cards remain in use on the shop floor. Only after all starvation avoidance cards have been returned can standard POLCA cards be used again. While premature station idleness can also be resolved by using more POLCA cards, this would be at the cost of POLCA's capability to limit the level of work-in-process.

According to Little's Law, the number of POLCA cards has a direct impact on realized flow times on the shop floor since POLCA cards limit the work-in-process. However, the total throughput time, which includes the time before a job is released to the first station in its routing, is not necessarily reduced by lowering the work-in-process. Rather, it depends on POLCA's load balancing capabilities (Germs & Riezebos, 2010). We cannot therefore predict in advance which setting of POLCA cards will lead to the best performance. In alignment with previous simulation studies on POLCA (e.g., Lödding *et al*., 2003; Fernandes & Carmo-Silva, 2006; Germs & Riezebos, 2010; Farnoush & Wiktorsson, 2013; Thürer *et al*., 2017; and Thürer *et al*., 2021), we therefore treat the level of POLCA cards as an experimental variable, i.e., five levels for the number of cards are considered: 8, 10, 12, 14, and 16 cards per loop. This spectrum of settings was chosen based on preliminary simulation experiments such that we capture the best performance across the performance measures considered in this study. The same number of cards is used within each loop in each experiment, which is justified by the balanced shop considered in our study. As a baseline measure, experiments with an infinite number of POLCA cards have also been executed.

2.2.2 Card Allocation and Dispatching Element

The same decision rule is used for the card allocation and dispatching decision. Two rules are considered: the Earliest Release Date (ERD) and Modified Earliest Release Date (MERD) rules. ERD was chosen since it is the card allocation and dispatching rule advocated in POLCA. The ERD's for each operation are calculated by backward scheduling from the job due date using planned operation throughput times for each operation in the routing of a job. This is similar to the calculation of operation due dates for each operation in the routing of a job (instead of a start date), which was shown to perform well in shops with high-variety routings (Kanet & Haya, 1982). As in previous POLCA literature (e.g., Riezebos, 2010) a constant allowance for the planned operation throughput time is used. This allowance was defined based on preliminary simulation experiments. The MERD rule combines the ERD with the SPT (Shortest Processing Time) rule. MERD divides the set of waiting jobs into two subsets: a subset of urgent jobs, for which the ERD has already passed and a subset of non-urgent jobs. Urgent jobs always receive priority over non-urgent jobs, whereby urgent jobs are selected for processing according to SPT and non-urgent jobs are selected according to ERD. This is similar to the modified operation due date rule (Baker & Kanet 1983). MERD was chosen as a more powerful dispatching rule alternative (Land *et al*., 2015).

The card-allocation and dispatching decisions may be *separated*, as is typically assumed in the literature, or they can be *integrated* as suggested in this study. If the decision is separated, then the card is attached to a job at arrival using the card-allocation rule. The dispatching rule can then only choose from jobs that have a card attached and are authorized to produce (i.e., queue DA_i). If the decision is integrated, then arriving cards remain in a card pool (the bulletin board). Whenever an operation is complete and the station becomes idle, the dispatching decision is executed by considering all jobs in the queue PA_i, i.e., it is decided which job should be produced. Only at this point it is checked whether this job can be produced, i.e., whether a POLCA card is available and can be attached to the job. As a result, a job that enters queue DA_i can be processed directly, i.e. queue DA_i becomes redundant and it can be eliminated. While for 'separated' a decision on the sequence in queue PA_i (card-allocation) and queue DA_i (dispatching) must be made, for 'integrated' only a decision on the sequence in the queue PA_i

(integrated card-allocation and dispatching decision) must be made since the card-allocation decision is postponed.

2.3 Experimental Design and Performance Measures

The experimental factors are: (i) the two types of card-based element (POLCA and POLCA SA); (ii) the two types of POLCA card-allocation and dispatching rule elements (separated and integrated); (iii) the five levels for the number of cards per control loop; and (iv) the two card acquisition and dispatching rules (ERD and MERD). A full factorial design was used with 40 (2x2x5x2) scenarios, where each scenario was replicated 100 times. All results were collected over 13,000 time units following a warm-up period of 3,000 time units to minimize initialization bias. These parameters allow us to obtain stable results while keeping the simulation run time to a reasonable level.

Since we focus on a make-to-order context, our main performance criterion is delivery performance. In this study, delivery performance will be measured by three main performance measures as follows: mean *total throughput time* – the mean of the completion date minus the entry date across jobs; *percentage tardy* – the percentage of jobs completed after the due date; and *mean tardiness* – that is, $T_j = \max(0, L_j)$, with L_j being the lateness of job j (i.e., the completion date minus the due date of job j). The percentage tardy provides the most general indication of delivery performance while the total throughput time indicates the mean lateness (given by the total throughput time minus the due date allowance). Meanwhile, both the mean tardiness and the standard deviation of lateness can be used to measure the dispersion of lateness across jobs. We decided to measure the mean tardiness since the standard deviation of lateness is more sensitive to extreme values than the mean tardiness. Finally, in addition to the three main performance measures, we also measure the average shop throughput time as an instrumental performance variable. While the total throughput time includes the time that an order waits (in a backlog) before being authorized at the first station in its routing, the shop throughput time only measures the time after an order has entered the queue at the first station in its routing.

3. Results

Detailed performance results are given in Table 2 with ERD card-allocation and dispatching. The results for MERD will be presented and assessed below. The results are given together with the 95% confidence intervals to indicate the statistical significance of the performance results. In addition to the results for POLCA, we also include the results for an infinite number of POLCA cards. Meanwhile, for POLCA SA, the maximum number of SA cards during a simulation run, averaged across replications, is also given in brackets.

[Take in Table 2]

The following can be observed from the results:

- *POLCA vs. POLCA SA*: When the number of POLCA cards is reduced, less work is authorized on the shop floor and the shop floor throughput times decrease. But this does not necessarily mean that tardiness performance improves or that total throughput times decrease, since the unauthorized work is still in the backlog. POLCA's workload balancing capability is not effective, i.e. there is only a limited reduction in the shop floor throughput time and even an increase in the total throughput time when compared to the use of infinite cards (Germs & Riezebos, 2010). Because of these increased throughput times, the mean tardiness and percentage tardy performance also deteriorates with a reduction in the number of cards. In contrast, POLCA SA allows for improvements across all four performance measures considered in this study for a separated card-allocation and dispatching decision.
- *Separated vs. Integrated*: There is an extensive reduction in shop floor throughput times, but there appears to be no significant improvement in terms of the percentage of tardy jobs, mean tardiness or total throughput times. Rather, performance even may deteriorate for POLCA SA. Meanwhile, if the card-allocation and dispatching decisions are integrated, then POLCA and POLCA SA perform statistically equivalent.

While the above results confirm previous literature on POLCA and POLCA SA (Thürer *et al*. 2017), they question our proposed change to the design of POLCA. Integrating the card-allocation and dispatching decisions does not yield the desired positive performance effect. This is however realized if a more powerful dispatching rule is used. This can be observed from the results when using MERD as the card-allocation and dispatching rule, as shown in Table 3. Again, the results are provided together with the 95% confidence intervals to indicate the statistical significance of the performance results.

[Take in Table 3]

Counterintuitively, MERD increases the total throughput times and mean tardiness when compared to ERD for the POLCA system and separated card-allocation and dispatching decision. It appears that MERD's SPT element introduces additional premature station idleness specifically during periods of high load. Consider the following situation: jobs are processed according to SPT at the upstream station in a POLCA loop, but when arriving at the downstream station in the POLCA loop, card allocation is according to ERD since jobs are still not urgent at this station. As a result, POLCA cards are retained longer at the downstream station and block the start of new jobs at the upstream station.

While MERD deteriorates performance compared to ERD for the original POLCA system that uses separate card-allocation and dispatching decisions, it significantly improves performance for POLCA SA, and most importantly for both POLCA and POLCA SA using integrated card-allocation and dispatching decisions. Compared to a separated decision and ERD (from Table 2), both percentage tardy and mean tardiness performance improve for POLCA (with 16 cards) by more than 75% (from 11.31% to 2.64%) and 50% (from 1.06 to 0.5 time units), respectively for an integrated decision if MERD is used.

Finally, for the separated card-allocation and dispatching decisions, the use of an infinite number of cards is equivalent to not exercising material flow control. To enter the queue DA_i (in Figure 1), an order only requires a POLCA card, which is always available. For the integrated decision, there is the extra condition that the station must be idle. This condition means a job is not released directly to the first station. In other words, for an integrated decision both a release decision and a production authorization decision are executed. The POLCA system with an integrated decision therefore becomes equivalent to combining POLCA with a work center workload trigger (see, e.g. Melnyk & Ragatz, 1989) that releases orders onto the shop floor whenever the queue is empty. In our experiments, executing the release decision only appears to be a better choice than executing the release decision and the production authorization decision providing that the right sequencing rule is chosen. In this case, the integrated approach dispenses with the use of POLCA altogether.

4. Conclusions

Our study started by asking: *Can POLCA performance be improved by postponing the assignment of POLCA cards to jobs from the point in time when a card arrives at the station to the moment when the dispatching decision is taken?* If ERD dispatching is used then our results confirm previous studies – integrating the card-allocation and dispatching decisions does not improve performance compared to separated card-allocation and dispatching decisions. But if a more powerful rule, such as MERD, is used then performance actually deteriorates when the decisions are separated and improves for the integrated decision, as suggested by our proposed design change. Our results further highlight that, for an integrated decision and MERD dispatching, an infinite number of cards – where the system transforms into a work center workload trigger that releases orders onto the shop floor whenever the queue is empty – leads to the best performance. In other words, in production environments like the one considered in this study, the integrated approach dispenses with the use of POLCA altogether if a suitable priority rule is used.

4.1 Managerial Implications

Our results indicate that companies that have implemented POLCA may be missing out on significant performance improvements by maintaining separate card allocation and dispatching decisions. This separation may occur by attaching a free POLCA card to a job in the queue as soon as it becomes available, as suggested in the literature, or by periodically attaching free POLCA cards from a so-called bulletin board that organizes all the POLCA cards currently available. In both cases, cards are attached to jobs before they are actually selected for processing. While this is a viable approach if jobs are prioritized according to ERD, better performance can be obtained by using MERD and an integrated decision. In this case, the use of POLCA is questioned altogether. In general, our study re-emphasizes managerial common sense, i.e. a decision that constrains a future decision should be postponed if there is no need for it to be taken early.

4.2 Limitations and Future Research

A main limitation of this study is that we only focus on two card-allocation and dispatching rules: ERD and MERD. While using two rules is arguably sufficient to answer our research question, more powerful rules may be available. This includes so-called look-ahead rules that take the queue at downstream stations into consideration. Meanwhile, future research could also incorporate other decisions, such as concerning machine assignment. Finally, POLCA was designed for production cells, so different decision rules within the cell may be constrained by the card-allocation and dispatching rule. Future research should assess whether this constraint is really necessary and whether performance can be improved by integrating the different decisions within a POLCA-controlled cellular manufacturing system.

References

Aneke, N.A.G., and Carrie, A.A., 1986, A design technique for the layout of multi-product flow lines, *International Journal of Production Research*, 24, 3, 471-481.

Baker, K.R., and Kanet, J.J., 1983, Job shop scheduling with modified operation due-dates, *Journal of Operations Management*, 4, 1, 11-22.

Blackstone, J.H., Phillips, D.T., and Hogg, G.L., 1982, A state-of-the-art survey of dispatching rules for manufacturing job shop operations, *International Journal of Production Research*, 20, 1, 27-45.

Braglia, M., Castellano, D., and Frosolini, M., 2014, Optimization of POLCA-controlled production systems with a simulation-driven genetic algorithm, *International Journal of Advanced Manufacturing Technology*, 70, 385 – 395.

Enns, S.T., 1995, An integrated system for controlling shop loading and work flows, *International Journal of Production Research*, 33, 10, 2801-2820.

Farnoush, A., and Wiktorsson, M, 2013, POLCA and CONWIP performance in a divergent production line: an automotive case study, *Journal of Management Control*, 24, 159 – 186.

Fernandes, N.O. and Carmo-Silva, S., 2006. Generic POLCA - A production and materials flow control mechanism for quick response manufacturing, *International Journal of Production Economics*, 104, 1, 74-84.

Germs, R., and Riezebos, J., 2010, Workload balancing capability of pull systems in MTO production, *International Journal of Production Research*, 48, 8, 2345-2360.

Graves, R.J., Konopka, J.M., and Milne, R.J., 1995, Literature review of material flow control mechanisms, *Production Planning & Control*, 6, 5, 395-403.

Harrod, S., and Kanet, J.J., 2013, Applying work flow control in make-to-order shops, *International Journal of Production Economics*, 143, 620-626.

Kanet, J.J., and Hayya, J.C., 1982, Priority dispatching with operation due dates in a job shop, *Journal of Operations Management*, 2, 3, 167-175.

Kanet, J.J., 1988, Load-limited order release in job shop scheduling systems, *Journal of Operations Management*, 7, 3, 44 – 58.

Krishnamurthy, A. and Suri, R., 2009, Planning and Implementing POLCA: a card-based control system for high variety or custom engineered products, *Production Planning & Control*, 20, 7, 596-610.

Land, M.J., and Gaalman, G.J.C., 1998, The performance of workload control concepts in job shops: Improving the release method, *International Journal of Production Economics*, 56 – 57, 347 – 364.

Land, M.J., Stevenson, M., Thürer, M., and Gaalman, G.J.C., 2015, Job Shop Control: In Search of the Key to Delivery Improvements, *International Journal of Production Economics*, 168, 257-266.

Liberopoulos, G., and Dallery, Y., 2000, A unified framework for pull control mechanisms in multi‐stage manufacturing systems, *Annals of Operations Research*, 93, 1-4, 325-355.

Lödding, H., Yu, K.-W., and Wiendahl, H.-P., 2003, Decentralized WIP-oriented manufacturing control (DEWIP), *Production Planning & Control*, 14, 1, 42-54.

Melnyk, S.A., and Ragatz, G.L., 1989, Order review/release: research issues and perspectives, *International Journal of Production Research*, 27, 7, 1081 – 1096.

Oosterman, B., Land, M.J., and Gaalman, G.J.C., 2000, The influence of shop characteristics on workload control, *International Journal of Production Economics*, 68, 1, 107 – 119.

Riezebos, J., 2010, Design of POLCA material control systems, *International Journal of Production Research*, 48, 5, 1455-1477.

Severino, M.R., and Godinho Filho, M., 2019, POLCA system for supply chain management: simulation in the automotive industry, *Journal of Intelligent Manufacturing*, 30, 3, 1271-1289.

Suri, R., 2018, *The Practitioner's Guide to POLCA The Production Control System for High-Mix, Low-Volume and Custom Products*, Routledge Productivity Press, Boca Raton, FL.

Suri, R., 2010, *It's about time: The competitive advantage of quick response manufacturing*, Productivity Press.

Suri, R., 1998, *Quick Response Manufacturing: A Companywide Approach to Reducing Lead Times*, Productivity Press, Portland, OR.

Thürer, M., Fernandes, N.O., and Stevenson, M., 2021, Material Flow Control in High-Variety Make-to-Order Shops: Combining COBACABANA and POLCA, *Production & Operations Management*, 29, 9, 2138 – 2152.

Thürer, M., Fernandes, N.O., Stevenson, M., Silva, C., and Carmo-Silva, S., 2019, POLC-A: An Assessment of POLCA's Authorization Element, *Journal of Intelligent Manufacturing*, 30, 6, 2435-2447.

Thürer, M., Fernandes, N.O., Carmo-Silva, S., and Stevenson, M., 2017, Improving Performance in POLCA Controlled High Variety Shops: An Assessment by Simulation, *Journal of Manufacturing Systems*, 44, 143-153.

Vandaele, N., Van Nieuwenhuyse, I., Claerhout, D., and Cremmery, R., 2008, Load-Based POLCA: An Integrated Material Control System for Multiproduct, Multimachine Job Shops, *Manufacturing & Service Operations Management*, 10, 2, 181-197.

Ziengs, N., Riezebos, J., and Germs, R., 2012, Placement of effective work-in-progress limits in route-specific unit-based pull systems, *International Journal of Production Research*, 50, 16, 4358-4371.

Table 1: Summary of Simulated Shop and Job Characteristics

Shop Characteristics	Routing Variability	Random routing; directed, no re-entrant flows
	No. of Stations	6
	Station Capacities	All equal
	Station Utilization Rate	90%
Job Characteristics	No. of Operations per Job	Discrete Uniform [1, 6]
	Operation Processing Times	Truncated 2–Erlang; (mean ≈ 1; max = 4)
	Due Date Determination Procedure	Due Date = Entry Time + *d*; *d* U ~ [30, 50]
	Inter-Arrival Times	Exp. Distribution; mean = 0.642

Table 2: Simulation Results for ERD Card-Allocation and Dispatching

	Cards	*POLCA*			
		STT[1)]	TTT[2)]	T[3)]	P[4)] (%)
Separated	8	22.06 ± 0.36	24.77 ± 0.57	1.51 ± 0.23	14.22 ± 1.28
	10	22.57 ± 0.37	24.07 ± 0.49	1.23 ± 0.18	12.42 ± 1.10
	12	22.90 ± 0.38	23.80 ± 0.46	1.12 ± 0.16	11.74 ± 1.02
	14	23.12 ± 0.39	23.69 ± 0.44	1.08 ± 0.15	11.45 ± 0.97
	16	23.28 ± 0.40	23.65 ± 0.43	1.06 ± 0.15	11.31 ± 0.96
	Infinity	23.62 ± 0.43	23.62 ± 0.43	1.05 ± 0.14	11.21 ± 0.93
Integrated	8	14.68 ± 0.26	24.83 ± 0.57	1.52 ± 0.23	14.34 ± 1.28
	10	14.89 ± 0.27	24.11 ± 0.49	1.24 ± 0.18	12.50 ± 1.11
	12	14.99 ± 0.28	23.81 ± 0.46	1.13 ± 0.16	11.77 ± 1.02
	14	15.04 ± 0.28	23.70 ± 0.44	1.08 ± 0.15	11.45 ± 0.97
	16	15.09 ± 0.28	23.65 ± 0.44	1.06 ± 0.15	11.32 ± 0.96
	Infinity	15.17 ± 0.28	23.62 ± 0.43	1.05 ± 0.14	11.21 ± 0.93
	Cards (Safety Cards)	*POLCA SA*			
		STT[1)]	TTT[2)]	T[3)]	P[4)] (%)
Separated	8 (3.0)	21.45 ± 0.32	22.93 ± 0.40	0.95 ± 0.13	10.32 ± 0.88
	10 (2.5)	22.17 ± 0.34	23.17 ± 0.40	0.97 ± 0.14	10.58 ± 0.89
	12 (2.0)	22.64 ± 0.36	23.32 ± 0.41	0.99 ± 0.14	10.77 ± 0.90
	14 (1.7)	22.96 ± 0.37	23.43 ± 0.41	1.01 ± 0.14	10.91 ± 0.91
	16 (1.4)	23.17 ± 0.39	23.50 ± 0.42	1.02 ± 0.14	11.00 ± 0.92
	Infinity	23.62 ± 0.43	23.62 ± 0.43	1.05 ± 0.14	11.21 ± 0.93
Integrated	8 (1.8)	14.80 ± 0.26	24.38 ± 0.52	1.34 ± 0.20	13.18 ± 1.18
	10 (1.3)	14.94 ± 0.27	23.93 ± 0.47	1.17 ± 0.17	12.07 ± 1.06
	12 (1.2)	15.02 ± 0.28	23.75 ± 0.45	1.10 ± 0.16	11.62 ± 1.00
	14 (1.1)	15.07 ± 0.28	23.67 ± 0.44	1.07 ± 0.15	11.37 ± 0.97
	16 (1.0)	15.10 ± 0.28	23.64 ± 0.43	1.06 ± 0.15	11.29 ± 0.95
	Infinity	15.17 ± 0.28	23.62 ± 0.43	1.05 ± 0.14	11.21 ± 0.93

95% confidence intervals on: [1)] shop throughput time; [2)] total throughput time; [3)] tardiness; and [4)] percent tardy

Table 3: Simulation Results for MERD Card-Allocation and Dispatching

	Cards	*POLCA*			
		STT[1)]	TTT[2)]	T[3)]	P[4)] (%)
Separated	8	22.74 ± 0.51	26.52 ± 1.15	2.80 ± 0.76	13.95 ± 1.55
	10	22.98 ± 0.46	24.64 ± 0.67	1.64 ± 0.34	11.68 ± 1.24
	12	23.11 ± 0.43	24.00 ± 0.52	1.29 ± 0.22	10.53 ± 1.00
	14	23.22 ± 0.42	23.73 ± 0.46	1.15 ± 0.17	10.04 ± 0.88
	16	23.30 ± 0.40	23.61 ± 0.43	1.08 ± 0.15	9.86 ± 0.82
	Infinity	23.51± 0.42	23.51± 0.42	1.02 ± 0.14	9.90 ± 0.83
Integrated	8	13.95 ± 0.19	23.07 ± 0.41	0.76 ± 0.13	3.42 ± 0.25
	10	14.09 ± 0.20	22.54 ± 0.34	0.55 ± 0.07	2.85 ± 0.23
	12	14.16 ± 0.20	22.41 ± 0.33	0.51 ± 0.06	2.71 ± 0.21
	14	14.21 ± 0.20	22.37 ± 0.32	0.50 ± 0.06	2.65 ± 0.20
	16	14.23 ± 0.20	22.36 ± 0.32	0.50 ± 0.06	2.64 ± 0.20
	Infinity	14.25 ± 0.20	22.35 ± 0.32	0.50 ± 0.06	2.63 ± 0.20
	Cards (Safety Cards)	*POLCA SA*			
		STT[1)]	TTT[2)]	T[3)]	P[4)] (%)
Separated	8 (2.8)	21.20 ± 0.30	22.60 ± 0.36	0.75 ± 0.10	9.29 ± 0.79
	10 (2.5)	22.01 ± 0.32	22.96 ± 0.37	0.83 ± 0.11	10.13 ± 0.84
	12 (2.1)	22.53 ± 0.34	23.19 ± 0.39	0.89 ± 0.11	10.54 ± 0.87
	14 (1.7)	22.89 ± 0.36	22.33 ± 0.40	0.94 ± 0.12	10.75 ± 0.89
	16 (1.4)	23.13 ± 0.38	23.43 ± 0.41	0.98 ± 0.13	10.91 ± 0.91
	Infinity	23.51 ± 0.42	23.51 ± 0.42	1.02 ± 0.14	9.90 ± 0.83
Integrated	8 (1.4)	14.04 ± 0.36	22.71 ± 0.36	0.61 ± 0.08	3.07 ± 0.27
	10 (0.9)	14.13 ± 0.20	22.46 ± 0.33	0.52 ± 0.06	2.76 ± 0.22
	12 (0.6)	14.19 ± 0.20	22.40 ± 0.32	0.50 ± 0.06	2.68 ± 0.21
	14 (0.5)	14.22 ± 0.20	22.36 ± 0.32	0.50 ± 0.06	2.64 ± 0.20
	16 (0.3)	14.24 ± 0.20	22.36 ± 0.32	0.50 ± 0.06	2.64 ± 0.20
	Infinity	14.25 ± 0.20	22.35 ± 0.32	0.50 ± 0.06	2.63 ± 0.20

95% confidence intervals on: [1)] shop throughput time; [2)] total throughput time; [3)] tardiness; and [4)] percent tardy