Inventory Diagnosis for Flow Improvement – A Design Science Approach

Martin Land, Matthias Thürer, Mark Stevenson, Lawrence Fredendall, and Kirstin Scholten

Name:	Dr. Martin J. Land
Institution:	University of Groningen
Address:	Department of Operations
	Faculty of Economics and Business
	University of Groningen
	9700 AV Groningen – The Netherlands
Email:	m.j.land@rug.nl

Name:	Professor Matthias Thürer
Institution:	Jinan University
Address:	School of Intelligent Systems Science and Engineering
	Jinan University (Zhuhai Campus)
	519070, Zhuhai – PR China
Email:	matthiasthurer@workloadcontrol.com

Name:	Professor Mark Stevenson
Institution:	Lancaster University
Address:	Department of Management Science
	Lancaster University Management School
	Lancaster University
	LA1 4YX – U.K.
Email:	m.stevenson@lancaster.ac.uk

Name:	Professor Lawrence D. Fredendall
Institution:	Clemson University
Address:	Department of Management
	101 Sirrine Hall
	Clemson, SC – United States
Email:	flawren@clemson.edu

Name:	Dr Kirstin Scholten
Institution:	University of Groningen
Address:	Department of Operations
	Faculty of Economics and Business
	University of Groningen
	9700 AV Groningen – The Netherlands
Email:	k.scholten@rug.nl

Keywords: *Buffers; inventory; capacity; flow; design science; operations management.*

Funding: This work was supported by the National Natural Science Foundation of China [grant number 71872072]; and the Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme 2017.

Inventory Diagnosis for Flow Improvement – A Design Science Approach

Abstract

Improving flow is a core Operations Management theme that is set to become even more important following contemporary developments in manufacturing, such as smart products and digital encapsulation that enable new control concepts such as multi-agent holonic control. But companies often struggle to realize flow improvements in practice, both with and without new technologies. While the literature agrees on the importance of flow, a structured and independent process that supports managers in identifying the root causes of why flow items wait in inventories instead of being processed is missing. Managers often use a single production management concept, such as lean production or the theory of constraints, when they seek to understand the reasons for a flow problem, which may lead to misdirected and unsuccessful interventions. In response, we use design science to develop a comprehensive approach to diagnosing flow problems that is independent from any production management concept. This diagnosis process results from successive iterations with five companies and supports the selection of appropriate analytical models and flow improvement solutions. It enables an organization to widen the focus of its flow improvement actions beyond the scope of a singular production management concept and complements the application of recent advances in technology, allowing smart products to quickly interpret what is happening in a location without first simulating and analyzing the whole system. Furthermore, the study expands buffer theories by showing that buffers have an internal hierarchy and can be absorbed by other buffers, whilst enhancing other theories related to coordination, material flow control and lean improvement.

Keywords: *Buffers; inventory; capacity; flow; design science; operations management.*

1. Introduction

This design science study develops an approach to diagnosing why in-process inventories, that indicate the presence of a flow problem, exist. The study provides a framework for flow problem diagnosis that extends existing inventory classifications, enabling a more granular diagnosis of the relationship between an individual inventory item and the variability against which it is buffering. We combine this framework, where all elements are compatible with existing literature, with a step-by-step process for diagnosing why an item waits in inventory. This diagnosis process is iteratively developed through application of the framework to five companies. Together, the framework and diagnosis process systematically guide managers to organize their thinking so that they analyze all possible causes of a flow disruption represented by each individual inventory item.

Our newly developed diagnosis approach reduces the risk that managers ignore any data that does not coincide with their preferred production management concept. In this way, it reduces managerial biases, such as confirmation bias (Pohl, 2004), that blind decision-makers from considering other flow inhibitors and potentially lead to a misdiagnosis of a company's flow problems. In fact, our design science study was triggered by a managerial misdiagnosis in one company, which resulted in a misdirected improvement investment that consequently did not yield the expected results. Our diagnosis helps to avoid confirmation bias by starting with the individual flow items that exist in inventory and stepwise assessing "why is this flow item waiting in inventory rather than being processed?" This engages managers in a systematic search process instead of allowing them to jump straight to analyzing only flow improvement actions that are aligned with their preferred production management concept. This reframes how we address flow problems and prompts managers to search for the true cause. It thus enables an organization to widen the focus of its flow improvement actions beyond the scope of singular production management concepts whilst complementing the application of recent advances in technology. Indeed, while flow improvement has been a core, enduring theme in Operations Management (OM), it may become even more important in the context of Smart Manufacturing technological advancements (Kusiak, 2018; Olsen & Tomlin, 2020), including smart products and digital encapsulation that enable new forms of production control (Meyer et al., 2011; Holmström et al., 2019). Managerial biases are not resolved via the use of new technology, rather they can easily be embedded into new technologies, such as digital twins (Grieves &

Vickers, 2017; Tao et al., 2018) that are being developed to improve our understanding of manufacturing processes.

Our diagnosis approach not only provides a practical tool for managers to develop context-specific solutions to variability, its development also enables us to contribute important insights that strengthen the Operations Management (OM) literature on flow improvement. First, our paper extends existing buffer theory (e.g. Newman et al., 1993; Lovejoy, 1998; Schwarz; 1998; Hopp & Spearman, 2004; Schmidt, 2005; Klassen & Menor, 2007) by identifying a form of buffer hierarchy, which explains how some buffers can be absorbed by other buffers meaning sources of variability remain hidden. Second, it provides a starting point for the use of simulation and queuing models to quantify the effects of current buffering solutions and to suggest alternatives. Third, it provides Smart Manufacturing technologies with a foundation by which they can quickly interpret what is occurring in a location without simulating and analyzing an entire system. Finally, it helps us to understand how planning and planning assumptions affect buffers.

The remainder of this paper is organized as follows. Section 2 reviews streams of literature that have been concerned with flow and buffering. It identifies the building blocks from the prior literature for developing our framework for diagnosis. The design science research approach adopted to develop our diagnosis process is then outlined in Section 3. In Section 4, the framework for flow problem diagnosis is deductively derived and an initial diagnosis process for its application is outlined. This diagnosis process is then iteratively developed through application to five cases in Section 5. A discussion of the main managerial and theoretical implications to emerge from the applications is presented in Section 6 before our paper concludes in Section 7.

2. Background

The importance of flow, i.e. the continuous movement of flow items, has been emphasized since the early conception of the OM field as Scientific Management (Taylor, 1919). In a fully synchronized process, demand, capacity and flow items become available at exactly the same time and a continuous flow is realized. However, in practice, variability often prevents full process synchronization meaning a perfect flow cannot be realized and one or more of the process inputs is forced to wait. Although all production management concepts focus on process synchronization, each has a different approach to managing the material flow cross operations.

For example, Statistical Inventory Control (SIC) seeks to minimize the number of items in inventory waiting for demand using reorder point and order-up-to-level methods. The Toyota Production System's (e.g. Monden, 1981; Ohno, 1988; Shingo, 1989) efforts to propagate one-piece flows and avoid inventories due to batching remain a major element of lean production (e.g. Hines et al., 2004; Holweg, 2007; de Treville & Antonakis, 2006; Shah & Ward, 2007). The logic underlying traditional Material Requirements Planning (MRP; e.g. Miller & Sprague, 1975; Orlicky, 1975) seeks to minimize the number of inventory items waiting for other items to arrive for assembly by coordinating the flows of parallel items contained within the bill-of-material. Finally, the Theory of Constraints (TOC; e.g. Goldratt & Cox, 1984; Rahman, 1998; Spearman, 1997; Watson et al., 2007) focuses on minimizing the buffer of orders waiting for capacity at the bottleneck whilst avoiding congestion at non-bottleneck resources. Thus, for arguably the four best-known production management concepts, we can identify a focus on four completely different causes of inventory – inventory waiting for demand, waiting for other batching items, waiting for other assembly items, and waiting for capacity.

Table 1 summarizes the above four production management concepts together with the associated flow problem diagnosis support that is available. We can observe from Table 1 that the control focus of the concepts is targeted towards specific causes of flow problems and that the diagnosis support within each concept is related to these specific causes. For a diagnosis to be correct, the manager must already know the cause of the inventory before selecting a production management concept to use for diagnosis. In contrast, we argue that an unbiased, independent approach to flow problem diagnosis should precede the adoption of any control solution. To provide the theoretical background for this new diagnosis approach we first review the literature on buffering against variability in Section 2.1, given that variability appears to be a main source of inventory. Section 2.2 then reviews existing classifications of inventory to provide the initial building blocks for our new diagnosis approach that follows in Section 4.

[Take in Table 1]

2.1 Theoretical Framing: Inventory and Buffering Against Variability

Multiple frameworks have been developed to explain the relationship between variability and its buffers. For example, Lovejoy (1998) and Schwarz (1998) postulated that all firms have an information/control/buffer portfolio that managers adjust to meet their competitive needs. Later, Schmidt (2005) used the concept of an OM triangle to model the relationship between capacity buffers, inventory buffers, and information, while Klassen & Menor (2007) built on Lovejoy's (1998) work using a process triangle with variability, inventory, and capacity utilization to explain how investments in one can substitute for the other(s). In general, the OM literature agrees on three broad types of variability buffer – inventory, capacity, and time (e.g. Newman et al., 1993; Hopp & Spearman, 2004). If variability exists then it will be buffered by one or a combination of these three variability buffers (Hopp & Spearman, 2004).

Each of the three buffer types relates to one of the three inputs required for an item to flow from one process step to another – (1) demand; (2) flow items; and, (3) capacity – as depicted in Figure 1. The flow items or transformed resources (Slack & Brandon-Jones, 2018) could be materials, information, or people and the capacity resources or transforming resources could be machines, operators, etc. In an ideal situation, the three inputs in Figure 1 are fully synchronized and become available simultaneously. Input availability however is often not synchronized due to variability. Inputs that are available before the others therefore have to wait, which leads to the three buffer types distinguished by Hopp & Spearman (2004).

[Take in Figure 1]

The input that is waiting for the other process inputs is a buffer to manage the lack of input synchronization. An inventory buffer implies that flow items are waiting for demand and a time buffer implies that demand is waiting for either flow items or capacity. An important distinction is that an inventory item represents a time buffer if demand has already been assigned to it. Here, inventory is not the buffering mechanism itself, but rather a consequence of time buffering. While this distinction between these two types of inventories is important, it has not been fully addressed in the literature. While the TOC literature has introduced the term "time buffering" it uses it in a more restricted sense. Time buffering within the Drum-Buffer-Rope approach refers to off-setting the release of flow items that have already been assigned to demand in terms of the bottleneck schedule (e.g. Watson et al., 2007), which controls the inventory of flow items waiting for capacity at the bottleneck.

More generally, time buffering results in flow items that are assigned to demand and then wait for either capacity or other flow items – these inventories will be referred to as *queues.* Queues are inventories waiting ahead of a process with known demand. In contrast, inventory buffers hold items that are produced in advance of demand, so that two process steps are decoupled by the inventory. This allows the first step to produce even though demand for the second step is unknown. We will logically refer to this type of inventory as *decoupling stock*. The position in the process where decoupling stock is located is called a *decoupling point*. This distinction between queues and decoupling stock – which was also made by Bertrand et al. (1990) – results in the first two of three definitions:

Definition 1: *A queue is the set of flow items between two process steps that is already assigned to demand from the downstream step.*

Definition 2: *A decoupling stock is the set of flow items between two process steps that is not yet assigned to demand from the downstream step.*

A particularly important type of decoupling point distinguished in the literature is the point where demand from an external customer is first assigned to flow items. This is the point where flow items wait until *external customer* demand for them has been specified. This particular type of decoupling point has been named the Customer Order Decoupling Point (CODP: e.g. Hoekstra & Romme, 1992; van Donk, 2001; Calle et al., 2016) or the Order Penetration Point (e.g. Sharman, 1984; Olhager, 2003). We adopt the term CODP in this paper, which we define as follows:

Definition 3: *The Customer Order Decoupling Point (CODP) is the decoupling point where flow items wait to be assigned to customer orders.*

The identification of the CODPs consequently has two important implications for the analysis of other inventories in an organization since all items downstream are assigned to (customer) demand:

Implication 1: *All inventory downstream of a CODP must be a queue.*

Implication 2: *Upstream of the CODP, both decoupling stocks and queues can exist.*

The position of the CODP will vary according to the type of product or service offering. For example, in make-to-order production environments the CODP occurs at the raw material stage while it is at the finished product stage in make-to-stock production environments. In assembly environments, where parallel flows exist, there may be multiple CODPs. Meanwhile, in procure-to-order production and many service contexts, there may not be a CODP at all within the internal processes of the company. The CODP position is highly relevant from a performance perspective. After the CODP the performance perspective of flow improvement generally relates to lead time reductions for customers, while before the CODP the internal inventory reduction itself might be the focus, which in turn relates to waiting time for the internal customer. Figure 2 indicates how different performance objectives relate to each other.

Figure 2 visualizes the difference between decoupling points and queues, and how these result from inventory buffering and time buffering, respectively. Cumulative representations over time, such as in Figure 2, have been used for many decades as they help to explain industrial dynamics (e.g. Forrester, 1961). The horizontal axis refers to time while the vertical axis indicates the cumulative number of flow items that are demanded and produced. When all of the depicted fluctuations in demand are handled by capacity buffers, production will follow demand. When time buffering is used, cumulative demand exceeds production and flow items assigned to demand wait to be produced. The vertical distance between the curves then represents the queue. Meanwhile, the horizontal distance indicates waiting time, which is the logical consequence of buffering by time. When inventory buffering is applied, cumulative production exceeds cumulative demand and the vertical distance between these two curves at any point in time represents the number of flow items in the decoupling stock. The horizontal distance is the runout time of the inventory, although it could also be seen as a negative waiting time.

Note that even if inventory buffering is the planned mechanism, cumulative production could still temporarily fall below demand. The vertical distance can then be seen as the backorder position instead of the queue, which is then equivalent to a negative inventory level (e.g. Zipkin, 2000). Following Little's result (Little, 1961) for stationary settings, we can thus always express both buffer types – inventory and time buffers – as either inventories or flow times. Notice that the other variable in Little's result, the throughput rate, would be represented by the slope of the curves in a stationary setting.

[Take in Figure 2]

The buffer theory outlined above explains why inventories occur. It represents the theory underlying our diagnosis approach. However, it only distinguishes in broad terms between inventories waiting for demand and inventories already assigned to demand. This level of granularity is arguably not enough to provide a comprehensive and systematic approach to inventory diagnosis. As a result, we next review existing inventory classifications in order to identify the building blocks for our more detailed diagnosis approach.

2.2 The Building Blocks: Existing Inventory Classifications

A rich literature has discussed the evils of inventory (e.g. Monden, 2011), such as hiding the real problems of an organization. Yet there has been only limited investigation of why particular in-process inventories occur in an organization. With the exception of Hopp et al. (2007), the inventory literature does not provide any structured diagnosis processes that allow for further distinctions between inventory types. Hopp et al. (2007) provided a systematic diagnosis tree for inventories in the specific context of production lines. This tree shows the type of inventory calculations that can be made in a specific branch, but the approach is restricted and only provides an overview of the underlying causes. Inventories can be the symptom of a flow problem, but before solving the problem we need to fully understand why an item exists in inventory.

Most textbooks distinguish between only a few basic reasons for inventory, e.g. cycle stocks, safety stocks, and anticipation or seasonal stocks (Cachon & Terwiesch, 2019; Slack & Brandon-Jones, 2019), but all of these reasons assume inventories are waiting for demand. Only some traditional textbooks (e.g. Bertrand et al., 1990) have discussed the essential distinction between whether an item is still waiting for demand (decoupling stock) or has already been assigned to demand (queue); but there has been little subsequent scientific investigation. Table 1 has already identified three reasons other than demand for why a flow item could be waiting in inventory, i.e. three causes of queues.

Based on the three inputs and using deductive reasoning, flow items waiting in *queues* can be waiting for capacity, waiting for similar flow items, or waiting for different flow items. Similar items are distinguished from different items because similar items arrive from the same preceding sequence of processing steps but have to wait for other items in the batch to enable setup reduction or joint processing (e.g. in an oven). Different items arrive through varying sequences of processing steps and have to wait for each other at a common point where they are jointly needed, for example, in an assembly or for transportation together to the next step. The three causes of queues relate to three waiting time types defined in traditional Dutch textbooks (e.g. Monhemius & Durlinger, 1985), i.e. service desk waiting time (or simply congestion), batch waiting time, and tour coach waiting time (or simply assembly waiting time). Monhemius & Durlinger (1985) also referred to platform waiting times, named after passengers waiting for a train on a railway station platform. Similar to congestion, platform waiting times relate to waiting for missing capacity. Here however the capacity is not continuously available, rather it comes in batches (e.g. a complete train).

In general, textbooks classify *decoupling stocks* as anticipation, safety, and cycle stocks (e.g. Cachon & Terwiesch, 2019; Slack & Brandon-Jones, 2019). This classification relates to the form of the variability that triggered production in advance of demand. To classify an inventory according to one of these three types relies on determining whether the production in advance of demand is due to uncertain fluctuations (safety stock), to predicted fluctuations (anticipation or seasonal stock), or to fluctuations caused by batching (cycle stock). These three possible forms of variability also apply to the three different queue types. While identifying the form of the variability is valuable, prior classifications do not specify whether the source of the variability is demand, supply or capacity. Nonetheless, the above has provided us with several ingredients to derive all possible root causes, which will be combined into our framework for flow problem diagnosis. The method used to develop this framework and the associated diagnosis process will be outlined next.

3. Research Method

We employ a Design Science Research (DSR) approach – where the goal is utility/effectiveness via description and explanation, as well as design and testing – in pursuit of *what can be* (van Aken et al., 2016). This approach was required as neither an exclusively inductive nor deductive logic, for defining a problem and making it traceable, were useful (Chandrasekaran et al., 2020). We concluded that the problem of understanding why inventories exist could not be solved using existing knowledge alone. Rather, we recognized the need to reframe the problem of how flow problems are addressed, adopting an item-level rather than top-down approach that is independent from production management concepts. Solving this newly framed problem required us to move from the problem to theory and then cycle in an abductive manner between theory, evidence and insight to arrive at a solution (Chandrasekaran et al., 2020). As such, DSR allowed us to consider a descriptive/explanatory component to cultivate a deeper understanding of the field problem while at the same time designing an artefact (van Aken et al., 2016). The artefact we develop consists of two elements: a framework for flow problem diagnosis that is developed deductively and robustly from the extensive literature and a diagnosis process that is developed inductively and iteratively through field tests.

The study broadly follows a rigor, relevance, and design cycle (Hevner, 2007) that proceeds through four phases of DSR, as suggested by Peffers et al. (2007) and Holmström et al. (2009a) and as is depicted in Figure 3. The literature was analyzed for ideas and the framework for flow problem diagnosis was developed as part of the rigor cycle (Hevner, 2007) (*Phase 1: Solution Incubation* in Figure 3). This framework was then 'translated' into a diagnosis process during the relevance cycle and extended during the design cycle (*Phase 2: Solution Refinement* in Figure 3). Phases 1 and 2 were conducted within an initial case organization (hereafter referred to as C0), where the problem was recognized, and Phase 2 was re-iterated in four other case organizations (C1-C4) to refine the artefact. Engaging with four additional organizations during the design cycle gave further insights for both the rigor and relevance cycles whilst ensuring pragmatic validity. Ultimately, the solution (framework and diagnosis process) allows us to provide key explanations of theoretical relevance (*Phases 3 and 4: Explanation* in Figure 3).

Overall, our study can be classified as an "improvement" DSR project (Gregor & Hevner, 2013) since it develops a new solution to a known problem.

[Take in Figure 3]

3.1 Phase 1: Solution Incubation

Solution incubation consists of identifying the problem, setting objectives for a solution as well as developing and designing the rudiments of a potential solution design (Holmström et al., 2009a). Our design science process was problem-centered (Peffers et al., 2007), with its initial motivation coming from the observation that company C0 misdiagnosed the root causes of its flow problem and implemented an intervention that did not solve this problem. Company C0 invested a significant amount of money in increasing the capacity of a key production step that it thought created a flow problem, without achieving an improved flow. The capacity intervention was aligned with the company's capacity-oriented planning focus, which in turn aligned with variable capacity requirements caused by a highly changeable product mix. Company C0 released orders when bottleneck resources had capacity available in order to achieve a smooth flow. When planning could structurally no longer avoid the accumulation of work at some bottleneck resources, the logical solution appeared to be an investment in capacity. While this logic had worked well for the majority of previous interventions, it did not work for the intervention that triggered this research since the root cause of the inventory was that items had to wait for other items to arrive ahead of final assembly. Rather than investing in capacity, as suggested by their capacity-oriented planning focus, the company should have focused on improving co-ordination, as would have been suggested by a material requirements focus. Hence, the initial objective of this research was to develop a flow problem diagnosis approach that would be independent of the adopted production management concept and that would enable the company to execute a correct diagnosis of why inventories exist, identifying not only inventory points where items wait but also the reasons behind why items wait.

We deductively developed a framework for flow problem diagnosis (Table 3 in Section 4) combining theories of buffering mechanisms (Hopp & Spearman, 1996) and the logic of queues and decoupling points (Sharman, 1984; Hoekstra & Romme, 1992; van Donk, 2001) with dimensions of inventories identified from some common classifications. This framework guides the user through an exhaustive list of potential causes that might explain why a flow item waits in in-process inventory instead of being processed. As explained by Wagner (1993), a generic diagnosis process will require such a list of potential problem causes (or explanations) to support problem-solvers in their search process. Following the development of this framework for flow problem diagnosis, an initial diagnosis process was designed to facilitate its application (see Section 5.1). The diagnosis process was developed inductively and refined through evaluations in five different companies during solution refinement. As such, the overall method employed in this paper is abductive and aligned with design science logic.

3.2 Phase 2: Solution Refinement

During the solution refinement phase, the rudimentary solution design is subjected to empirical testing (Akkermans et al., 2019). The initially developed artefact, consisting of the framework for flow problem diagnosis and the diagnosis process itself, was applied to C0 by two researchers with input from the CEO of the company and the production manager. The production process was first mapped and the different points where inventory accumulates were noted. Then, each researcher diagnosed the types of inventory present and the root causes of the inventory using the framework for flow problem diagnosis and the diagnosis process (evaluation). In the reconciliation of the independent diagnoses it was found that the researchers agreed about the types of inventory but that further information about the production process was required. The missing insights were gathered during follow-up interviews with the CEO and a process improvement manager. After another round of iteration with C0's management it was possible to derive a final diagnosis of the inventory types in the company. Problems encountered during the application of our diagnosis process were directly addressed through *ad-hoc* solutions to ensure the swift execution of the diagnosis process. After the diagnosis, the research team discussed problems identified with the approach. This testing and evaluation in C0 led to adjustments to the diagnosis process and as such helped to refine the artefact (see also Section 5.1).

Following the solution refinement in C0, we employed a multiple case study methodology to ensure the pragmatic validity of our design by re-iterating the solution refinement process several times (Hevner et al., 2004;

Peffers et al., 2007; van Aken et al., 2016; Akkermans et al., 2019). Therefore, we chose four additional organizations based on theoretical and literal replication drawing on the dimensions of Hayes & Wheelwright's (1979) volume/variety spectrum of production types. We expected the diagnosis process to be most useful in the middle of this spectrum. For job shops, the diagnosis was expected to be too complicated due to the variability in routings and processing times while it was anticipated that the usefulness of the approach would be diminished in continuous flow processes due to the limited opportunities for inventories. Accordingly, we selected two organizations (C2 and C3) that were similar to C0 in terms of their process and product structure and two organizations (C1 and C4) that had different structures (Seawright & Gerring, 2008). C1 was selected as an example of a company with a greater degree of customization and a job shop configuration; and C4 was selected as an example of a company with a greater degree of standardization and a line layout. The companies are all positioned across the diagonal depicted in Hayes & Wheelwright's (1979) classification in Figure 4. Together, they cover a large proportion of the companies that use conventional manufacturing techniques. None of the companies, except for the job shop company C1, use any specific digital approaches to allow high variety at high volume, while the positioning of customer-order-decoupling points varied both within and across the companies. The company characteristics and their purpose in terms of this research are summarized in Table 2.

[Take in Figure 4 and Table 2]

The researchers made use of existing company contacts to gain access to organizations that met the case selection criterion. In all four additional organizations (C1-C4), senior staff (e.g. CEOs, Production Managers, Operations Directors) were contacted in 2018 to explain the aim of the study. After confirming their willingness to participate and gain insights into their flow problems, the actual diagnosis followed a similar three-step approach. First, *pre-implementation*, where researchers discussed possible design changes to the diagnosis process given the company characteristics. For example, in C1 no clear map of the production process could be created since crisscross flows through the shop may occur and hundreds of different routings are registered. Second, *demonstration*, referring to the actual application of the diagnosis process. This consisted of a set of physical on-site observations. Two researchers independently applied the framework for flow problem diagnosis and took extensive notes about important problems or deviations from the planned diagnosis process. Third, during the *evaluation*, notes were compared, the main implications discussed, and conclusions revised if necessary. To ensure reliability, every diagnosis was conducted by two researchers. Meanwhile, to ensure consistency, one particular researcher was involved in all five applications of the artefact (C0-C4).

In a final step, and to increase the pragmatic validity of the designed solution, we returned to C0 approximately one year after the initial diagnosis with the iteratively refined diagnosis process and asked the production manager to reapply it to their production process. Following this diagnosis, three changes were proposed by C0's management to solve the company's flow problem. This demonstrates that the artefact could be used as part of a larger intervention design, as proposed by Oliva (2019). As such, this second diagnosis confirmed the applicability and relevance of the artefact for providing a solution to their problem of identifying the right causes of why inventories exist.

4. Design and Initial Development

Our artefact combines a deductively derived diagnosis framework for identifying flow problems with an inductively developed diagnosis process to apply this framework. The former is outlined in Section 4.1 while the latter is outlined in Section 4.2

4.1 Development of a Framework for Flow Problem Identification: Deriving 28 Causes of Inventory

The framework for flow problem diagnosis guides managers through an exhaustive list of possible causes of flow problems. The framework deduces and classifies all possible reasons for why a flow item exists in inventory by using existing definitions and dimensions provided by other classifications in the literature. Creating this classification of potential causes is essential to support problem solvers in their search process (Wagner, 1993). The overall logic can be captured in three "if … then …" statements: 1) if inventory exists, then there is a missing input; 2) if there is a missing input, then there is some source of variability; and, 3) if there is variability, then this variability takes one of three forms. The resulting framework is presented in Table 3. Horizontally, and from left to right, it shows the four possible missing inputs for which the flow items in the inventory can be waiting, i.e.

demand, capacity, and flow items (which can be either flow items of a different or similar type). The main vertical categories show the possible sources of variability. Independent from the identified missing input, these sources can be in each of the three inputs (demand, supplied flow items and capacity) that must be synchronized. Finally, within each source we can have the three forms of variability derived from the classical textbook distinction between decoupling stocks. Note that the column with "Similar Type Flow Items" as a missing input is not split across multiple rows since, in this case, the variability can only relate to batching for the next process. The inventory classification in Table 3 thus identifies 28 possible causes of inventory and associated inventory types – each of which is individually defined in Appendix 1.

[Take in Table 3]

Wherever possible, we have used existing naming conventions to label the different inventories in Table 3. When demand is the missing input, inventory is the key buffering mechanism, and the inventory is referred to as a type of stock. The inventories in the other three columns are labelled as waiting times since these inventories are identified as queues resulting from time buffering (see Definition 2 in Section 2.1 above). The labelling of the variability types follows the common distinction in the literature between safety, anticipation, and cycle stocks for uncertainty, predicted fluctuations, and batched movements, respectively. We claim that the framework's granular and systematic classification of the causes of inventory provides a useful basis for understanding why inventories exist. Further, we claim that this is important since each of the 28 causes will likely require a different management response to improve flow.

Prior inventory classifications have only included one of the three dimensions (missing input, source of variability, form of variability), which limits the search space. For example, decoupling inventory typologies (safety, anticipation, cycle stock) only consider the type of variability while queue typologies only look at the missing input that hinders the progress of a flow item (capacity or other flow items). If company C0 had used our framework during its diagnosis process it would have avoided the more general problem that triggered this study. The framework would have challenged the manager to consider that the observed queuing of flow items might be due to assembly requirements, i.e. waiting for other flow items to arrive, rather than congestion due to a bottleneck. This in turn would have engaged the manager in searching for a solution related to synchronizing the incoming flows. Yet, although the framework for flow problem diagnosis provides a first solution for identifying the causes of inventory in an organization, on its own this may be too abstract to be applied directly by managers. As such, solving flow problems is not just a matter of logical deduction based on the application of generic knowledge but requires further design (van Aken & Romme, 2012; Van Aken, 2014). Hence, a *diagnosis process* for applying the framework to organizations is also needed. We consequently continue to follow a design science approach (Holmström et al. 2009a; van Aken et al., 2016) to develop a diagnosis process that complements the framework and contributes to the overall aim of developing an artefact that supports practicing managers in identifying the root causes of inventory.

4.2 First Design of a Diagnosis Process

Three main requirements have to be fulfilled for the application of the framework developed above:

1. The diagnosis process should be applied to all points where flow items are delayed in the company's production process. As inventories may occur before and after each operation, a process map should be created that covers all operations in the physical flow through the company to identify all possible inventory points. This provides a contextual background to each piece of inventory.
2. A useful starting point should be derived for the diagnosis process. Rather than beginning at the most upstream point, Section 2 suggested that the CODPs may provide a more effective starting point. Since all inventory downstream of each CODP must be a queue, starting at the CODP simplifies the process of distinguishing other decoupling stocks and queues.
3. The framework should be translated into a diagnosis process. For each item in the inventory point the three dimensions of the inventory classification can be translated into three simple questions. By answering these questions the user arrives at the correct cell in the framework. A complete diagnosis process however relies on going one step further than identifying the right cell. Therefore, once the correct source of variability has been traced using the generic framework, the fourth and final question is context-specific, i.e. what causes this variability in this specific context?

All three requirements are reflected in the initial design of our diagnosis process, which focusses the user's attention on particular aspects of the production process by first mapping the process, specifying where the analysis should start, and what questions should be asked. The initial design was kept as simple as possible to ensure it is generic and can be easily applied. It can be summarized as follows.

1. Map the process to identify all inventory points.
2. Determine the customer order decoupling points (CODPs) and follow Step 3 for these points.
3. Ask the following four questions for the flow items in the inventory point:
    I. For which process input (missing input) are the flow items in this inventory waiting?
    II. Which input is the main source of variability that causes missing inputs?
    III. What form does this variability take?
    IV. What causes this variability in this specific context?
4. Repeat Step 3 for all remaining inventory points.

Step 3 forms the core of the diagnosis approach. It executes the diagnosis for a specific item at an inventory point using our framework for flow problem diagnosis. Meanwhile, steps 1 and 2 provide the 'driving instructions' on how to arrive at these inventory points in the smartest way. They apply a specific perspective to inventory diagnosis, which can be qualified as location-based. Each location where inventories occur will have its own diagnosis process step 3 (the core diagnosis) for the full set of inventorying items. Another perspective would have been to follow each item individually on its journey through the shop floor, which is the perspective of tracking and item-centric materials management (Holmström et al., 2009b). The results of item tracking could equally be used to identify commonalities between delays across flow items and then to point out the sources of variability. However, not all companies are yet able to facilitate item-centric materials management and so adopting this perspective would likely restrict the practical applicability of our diagnosis process at present.

5. Demonstration and Evaluation

5.1. Demonstrating and Evaluating the Design in Company C0

This section uses the initial diagnosis process presented above to apply the framework for flow problem diagnosis to C0. It begins with process mapping and the identification of CODPs before examining other inventory types within the shop. It concludes with an enhanced diagnosis process based on an evaluation of the demonstration in C0, which is taken forward into the other four cases.

5.1.1 Demonstrating Diagnosis Steps 1 and 2: Process Mapping and the Identification of CODPs

The production manager in C0 provided us with a detailed process map of the shop floor and product flows. This map was used as the basis to identify known CODPs and outline inventory policies, production planning procedures, etc. that may be useful for understanding the occurrence of inventories. The 20 production steps for most products are shown in Figure 5, with each production step connected by an arrow (*a* to *w*). Rods are produced in steps 1-6; rubber belts are produced in a different department in steps 7-9; and the two are assembled together by riveting in Step 10 before being coated and dried (Step 11). Some production steps are only needed for a subset of products, as indicated by arrows that circumvent a production step in Figure 5. The rolls and shafts are machined separately (steps 17 and 18, respectively) and the shaft is inserted into the roll in Step 19. Most belt-rod assemblies are delivered together with cogs, processed in steps 12-16, and the rolls are delivered with the inserted shafts (Step 19). The parts are then collected and stacked on pallets before being packed, which typically involves wrapping them in plastic (Step 20). The stacking and packing process may involve several products being grouped together for delivery to the same customer.

[Take in Figure 5]

The CODPs were identified as the production steps after inventories *a, h, p* and *t,* which do not commence until a customer order is available. The CODPs are represented by unshaded triangles in Figure 5.

5.1.2 Demonstrating Diagnosis Step 3 and Step 4: Core Diagnosis

The diagnosis process started at the CODPs. At any CODP, demand, more specifically customer demand, is the missing input for which the flow items are waiting. A main source of variability at all four CODPs was also demand, and more specifically demand uncertainty. Hence, all four inventories could be qualified as *demand safety stock*. This initial inventory diagnosis at the CODPs already greatly improved overall understanding of the complete flows and provided important insights into our design, as will be discussed in Section 5.1.3 below.

After focusing on the CODPs, the researchers went through the full production process on the shop floor – both physically and on paper – together with the production manager. This allowed them to visualize the process and observe all inventories in the system. The research team systematically stopped at each step in the process to understand the operation and its connections with upstream and downstream stations. We will now briefly discuss an example for each of the remaining inventory types not covered above that occurred in the company.

As customer orders are known, only queues waiting for capacity or other flow items can occur downstream of the CODPs (see Implication 1 in Section 3.2). Inventories upstream of a CODP can still be decoupling points, waiting for demand in the form of a production order for the next production step. It is important to reemphasize that the term *demand* in the diagnosis process does not necessarily relate to customer demand. The upstream inventories at *g, l, r* and *s* were waiting to be assigned to a production order when the raw materials arrived. They can consequently be classified as decoupling stocks. However, all of these decoupling stock inventories are relatively small since production steps 7, 12, 17, and 18 occur soon after the purchased items have arrived. The main source of variation is the batched supply, so these raw material inventories at *g, l, r* and *s* are classified as *supply cycle stock.*

The inventory at *m* and *n* is a queue, since a single work order is specified for the movement of the flow items through drilling (12), splitting (13), and reassembling (14), implying that demand from the downstream steps is known at both *m* and *n*. It was classified as *batched supply induced congestion* as the small queues occurring here relate to missing capacity due to the execution of the previous production step in small batches by the same operator. *Batched transformation induced congestion* occurs at *c*, which precedes the pressing step (3) that forms a capacity bottleneck, processing the rods in batches due to large set-up times. Inventories at *d* and *e* are also classified as *batched transformation induced congestion*, but for a different reason. Here the successive heat treatment processes at steps 4 and 5 have more than sufficient capacity. Yet, to avoid wasting energy, they are only operational for part of the week thereby causing an irregular batched outflow from inventories *d* and *e.*

The longest waiting times in the entire process occur at the point where rods and belts come together to be riveted (10). Rods (j_1) and belts (j_2) have unsynchronized arrivals, so one waits for the other because the two responsible departments optimize their own individual schedules. The inventories are partly classified as *batched supply induced assembly waiting time* (where the rods department purposely creates batches in preceding operations) and partly as *supply uncertainty induced assembly waiting time* (for the unplanned part). After riveting (10), the inventories at *k* are very limited due to a large capacity at coating and drying (11). Some inventory occurs before stacking and packing (20) because different parts and even different orders for the same customer may need to be combined (yet are not available at the same time). These customer orders will be supplied according to a plan, so this inventory is classified as *supply anticipation induced assembly waiting time* for the assembled belts at v_1, the cogs at v_2 and the rolls at v_3 when the shafts are inserted based on customer orders. The inventory at *f* is classified as *demand uncertainty induced congestion.* It precedes the optional step (6), which is only executed for part of the uncertain demand mix. This production step, encasing the steel rods in rubber sleeves, prevents bruising to harvested products. It is mainly executed for products delivered to the American market, where harvesting machines are used at higher speeds.

Finally, Table 4 shows the results of the diagnosis for all inventory points marked in the initial process map (Figure 5).

[Take in Table 4]

5.1.3 Evaluation: Resulting Design Adaptations to the Diagnosis Process

Following this first application and demonstration, the diagnosis process was evaluated. We found that the initial inventory diagnosis at the CODPs (Step 3) revealed five important lessons. First, <u>it is important not to immediately move to the next inventory point when a first source of variability has been found, as there can be multiple causes of inventory waiting for demand at a single inventory point</u>. For example, at the CODP labelled *h,* part of the inventory could also be classified as *demand anticipation stock*. The preceding punching operation (7) operated near full capacity, so it could not even respond to predictable fluctuations in demand. As a consequence, during periods of low demand it processes items meant for anticipated future demand. Likewise, not all inventory at *p* and *t* was due to demand variability. Since the raw material for the cogs and rolls were cast in large lots by the supplier, part of the inventories should be classified as *supply cycle stocks*. It was not however possible to quantify the parts related to different causes. As soon as inventory is created for one cause, it would also help to buffer against other causes that are not fully correlated (buffer pooling). Based on this lesson we refined Question 2 of Step 3 in the diagnosis process.

The second lesson for the diagnosis process came from examining the reasons why inventory existed at point *t.* We observed that point *t* was not always the CODP for the rolls because shafts were sometimes inserted in advance of customer demand. This meant that CODP v_3 also existed for some products (see the dashed lines in Figure 5). We learned that <u>care has to be taken in distinguishing all possible product-market combinations before the diagnosis is finalized</u>. For the diagnosis process this meant that we extended Step 1 and added Step 5.

The same issue resulted in the third lesson, because v_3 was not a CODP for the products processed at the time of the diagnosis. More generally, the company indicated that seasonal differences related to the harvesting periods could change the causes of inventories. Thus, we learned that <u>it may be important to repeat the diagnosis at periodic intervals to account for differences over time</u>. Accordingly, we added Step 6 to the diagnosis process. Lessons four and five related to the application of the diagnosis process, rather than affecting its steps.

The fourth lesson was that <u>flow items could remain at the same inventory point once a missing input became available, because they had to wait for another missing input</u>. This occurred at CODP *a*, where first coils of steel were waiting for demand. Once an order was placed that needed a certain type of steel, the coils still waited for other orders to be combined so that an entire coil could be used during one production run. At that stage, flow items are waiting for flow items of a similar type, so the inventory is due to *batch waiting time*. This means that CODP *a* can be visualized as a decoupling point followed directly by a queue of material waiting to be batched. We realized that it could even have been followed by another queuing stage if the inventory had to wait for capacity (*batch transformation induced congestion*) after the items to be batched were available. This however was not the case. The relationship between different inventory types further provides relevant lessons for buffer theory as both a hierarchy appears to exist and different inventory types hide each other.

The fifth important lesson learned was that causes for the position of a decoupling point do not necessarily relate to variability. The production manager stated that the CODP inventories were placed to avoid extending the customer waiting times due to long supplier lead times. Although not a source of variability, this cause seemed highly reasonable, explaining (1) the position of the CODP (see, e.g. van Donk, 2001) and (2) why decoupling (inventory buffering) is used instead of time buffering. In contrast, the demand variability during this supplier lead time explains the number of flow items waiting at the chosen decoupling points.

The main lessons learned from our first design cycle and their implications for the diagnosis process are summarized in Table 5. This evaluation helped to further develop the diagnosis process as follows, where changes to the initial design are highlighted using italics.

1. Map the process to identify all inventory points *for all product-market combinations (PMCs) and start Step 2 for one of the PMCs.*
2. Determine the customer order decoupling points (CODPs) and start Step 3 for these points.
3. Ask the following four questions for the flow items in the inventory point:
   (i) For which process input (missing input) are the flow items in this inventory waiting?
   (ii) Which input is *a* source of variability that causes missing inputs?
   (iii) What form does this variability take?
   (iv) What causes this variability in this specific context?

   In doing so, consider *the potential for multiple answers to the previous questions.*
4. Repeat Step 3 for all remaining inventory points.
5. *Repeat steps 2 to 4 for the remaining PMCs.*
6. *Periodically repeat steps 1 to 5 to check for changes over time.*

[Take in Table 5]

5.2 Demonstrating and Evaluating the Design in Companies C1-C4

The lessons learned during the first use of the initial design in C0 were embedded before the updated diagnosis process was applied to C1 to C4.

5.2.1 Demonstrating Diagnosis Steps 1 and 2: Process Mapping and the Identification of CODPs

In the second design cycle, companies were selected at the extreme ends of the volume/variety spectrum. This changes the perspective that needs to be taken for Step 1 and Step 2 to arrive at the inventory points for the core analysis (Step 3). In the high variety job shop company C1, crisscross flows through the shop may occur and hundreds of different routings are registered. Different to C0, where we took a *product-flow* perspective to the diagnosis due to its position in the middle of the volume/variety spectrum, we adopted a *resource-based* perspective, i.e. we structured the identification of inventory points around production resources instead of the material flow. For each of the functional departments of the company we reviewed the set of flow items waiting to be processed, categorized the items based on routing similarity, and for the main routing categories we analyzed the reasons for inventory, i.e. why flow items stop before being processed in this department. We followed the same resource-based perspective for C2. Although C2's process and product structure are similar to C0, it produces a larger variety of products. At the other end of the volume/variety spectrum, C4 had an extremely simple flow; most products were produced on a continuous line without opportunities for inventory within the line. Consequently, the main inventories were raw materials and finished goods. However, many differences appeared to exist within the spectrum of raw materials and finished goods. Therefore, we followed a *product-based* perspective, working through a list of all raw materials and finished goods produced in the company to identify sets of inventories as input for our core diagnosis. In C3, theoretically the same product-flow perspective could have been adopted as in C0. However, the size of the company (thousands of employees and hundreds of machines at the studied site) made it impossible to diagnose all inventories within the company. Instead, we decided to analyze one large inventory point between two key departments (mold injection parts and assembly) in-depth to determine whether that would lead to new insights for the design.

In Step 2, in C1 and C2 raw material stocks were usually the CODPs waiting for customer orders. In these companies nearly all activities were executed in response to a customer order, so that nearly all of the downstream inventories were queues waiting for either capacity or other flow items. In C4, we first identified whether the finished goods or raw materials formed the CODP, or whether raw materials were even purchased to order, before diagnosing the other inventories. In this company most products had only two inventory points (raw material and finished goods inventories).

5.2.2 Demonstrating Diagnosis Step 3 and Step 4: Core Diagnosis

Similar to the first design cycle in C0, the researchers went through the full production process in C1-C4 on the shop floor – both physically and on paper – together with an employee of the company to visualize the process and observe all inventories in the system. Across the companies, 22 different inventory types have been identified, including all three types of safety stocks, anticipation stocks, and cycle stocks. The two most common forms of inventory in the five companies are demand safety stock and supply safety stock, which were both observed in all five companies.

Note that six types of inventory were not found in these five companies. None of the three types of capacity induced assembly waiting time could be found since capacity variability at the 'assembly station' would in most cases affect all of the different flow items that need to be assembled to the same extent. As a consequence, the flow items will mostly be waiting together and not for each other. Similar reasons apply for the missing two types of demand variability induced assembly waiting time, which would only occur if demand variability affected some of the flow items to be assembled. Meanwhile, *capacity anticipation induced congestion* is most likely to occur when planned maintenance takes place; hence, this would only have been identified if maintenance was scheduled on the day of the diagnosis. Thus, overall, the five companies have enabled us to identify a broad range of inventory types, including all of the most common and some of the less common types.

5.2.3 Evaluation: Resulting Design Adaptations to the Diagnosis Process

After evaluating the second, third, fourth and fifth demonstration, six key lessons became evident that helped us to further refine the diagnosis process. First, <u>the perspective to the diagnosis process needs to be adapted to the</u>

shop type and its complexity so that it is either *resource-based* (job shop), *product-flow* focused (batch process) or *product-based* (continuous flow). After the design cycle in C0, which is positioned in the middle of the volume/variety spectrum, the diagnosis assumed the use of a process map as the starting point, to enable the user to follow the product flow stepwise through the company. In production processes with medium volume and variety, a clear product flow can generally be found, and such a *product-flow* perspective can be adopted. In other environments, however, that was not possible. To be able to map the process (Step 1) in C1 and C2, we had to structure our diagnosis around production resources instead of the material flow due to an increased amount of product routings. On the opposite side of the volume/variety spectrum, i.e. C4, we applied a *product-based* perspective due to the extremely simple flow where the main inventories were raw materials and finished goods in large volumes. Accordingly, we further differentiated Step 1 in the diagnosis process.

This difference in environment also had implications for the role of the CODPs and led to the second lesson: identifying CODPs is crucial for batch processes, but less crucial for job shop or continuous flow environments. In C0, and its dominant product flow, it was useful to identify CODPs in the process map to start the diagnosis of inventories as they had clear implications for preceding and succeeding inventory points. However, the CODP is defined by the product flow. If product flows are undirected and a high variety of different products is produced then no clear CODP can be identified and a different perspective for diagnosis needs to be adopted. This indicates that Step 2 of the diagnosis process might be skipped for job shop and continuous flow environments (i.e. in C1, C2, and C4).

The third lesson relates to the diagnosis Step 3 where we found the need to also repeat diagnosis Question 3 (iv): What causes this variability in this specific context? The diagnosis of the large inventory point in C3 revealed that a single point contained inventory caused by one type of inventory (*supply uncertainty induced congestion*) but related to two independent sources of variability. For part of the variability the source had to be found externally and related to quality problems with supplied materials, while the second source was found in internal process yields. Hence, the need to repeat diagnosis Question 3.

From the same example, we derived the fourth lesson as we found that the level of this particular supply uncertainty in C3 appeared to be rather low. To buffer against supply uncertainty the company planned to have the material available 24 hours before corresponding assembly capacity was scheduled. This highlighted to us that the buffer size may not relate to the real variability but rather to the planned variability. The highly variable environments of C1 and C2 had a similar issue. These companies mainly produce to customer orders and quote customer order-specific delivery dates that already determine the amount of waiting time that will occur within the company. Even if the realized waiting times in production were shorter than planned for in the delivery dates, the items to be delivered would generally wait for transportation to the customer in the final stage. Also within production, the waiting times would depend on the schedules created internally. These schedules included time buffers to allow for certain amounts of uncertainty or batching, which might either be larger or smaller than the amount needed. We might even see items waiting without the presence of variability, but rather because it had been planned for. As an example, most order routings in C1 started with a laser cutting operation, requiring different customer orders to be combined in a nesting at these machines. It was uncertain when the right set of orders would be available to avoid needless material spoilage at laser cutting. This prompted the planners to allow for a fair amount of *supply uncertainty induced congestion* in the planned waiting times for downstream machines. Since the full capacity was assigned to orders for a long period ahead, the orders would have to wait even though the finally realized nesting would lead to a smooth supply of orders from laser cutting. The studies in C1, C2, and C3 all confronted us with the role of planning. We consequently extended Question 2 of Step 3 to state 'Which input is a source of planned or actual variability that causes missing inputs?'

The *supply uncertainty induced congestion* diagnosed in company C1 also provides an example of the fifth lesson learned, which is important for the application of the diagnosis process: the variability related to the supply from the previous operation could have its root cause much further upstream. In C1 we observed the impact of nesting at the first operation in the form of supply variability for more than four operations further downstream. Similarly, customer demand variability that characterizes the make-to-order environment could also be observed as production order related demand variability at far upstream operations. For example, in C2 this relates to both *demand uncertainty induced congestion* and *demand uncertainty induced assembly waiting time,* which were observed even before the first operation in certain production departments. Based on this lesson, we extended the considerations for Step 3 in the diagnosis process.

Waiting times preceding the first operation led to a sixth lesson when this waiting time related to the order book instead of physical items. Particularly in C1, a significant part of the waiting times occurred before the first operation and before material was assigned to an order. As such, no flow items, but rather just the order, was waiting before being released. The planner would not release the order if capacity at some of the operations to be executed would not be available in the short term. The waiting time of the order could be qualified as *demand uncertainty induced congestion* as capacity was missing and the unpredictable arrival of customer orders was the main source of variability. However, avoiding the waiting time of flow items on the shop floor and replacing it with waiting time before release could also be considered positive for the flow of this company.

The main lessons learned from the additional applications of the framework for flow problem diagnosis to C1-C4 and the implications for the diagnosis process are summarized in Table 6. These lessons learned allowed us to re-evaluate the diagnosis process and develop it further. Changes to the previous design from Section 5.1.3 are again marked in italics.

1. *Identify all inventorying items according to the type of shop*:
    a. *In a batch process (with families of routings), map the process* to follow a product through the process to identify all inventory points for each of the product-market combinations (PMCs). Start Step 2 for one of the PMCs.
    b. *In a continuous flow process (with a single short routing),* adopt a *product-based perspective*, creating a list of all PMCs to identify their raw material and finished goods inventories. Start Step 2 for one of the PMCs.
    c. *In a job shop process (with a high variety of routings)*, adopt a *resource-based perspective* to map the overall shop and identify the items inventorying at each resource.
2. Determine the customer order decoupling points (CODPs) and start Step 3 for these points.
3. Ask the following questions for the flow items in the inventory point:
    (i) For which process input (missing input) are the flow items in this inventory waiting?
    (ii) Which input is a source of *planned or actual* variability that causes missing inputs?
    (iii) What form does this variability take?

(iv) What causes this variability in this specific context?

In doing so, consider the potential for multiple answers to the previous questions and *root causes for supply variability that can be further upstream, for demand further downstream*.

4. *In shop types identified as a job shop process in Step 1c,* repeat Step 3 for all remaining items at the same inventory point.
5. Repeat steps 3 to 4 for all remaining inventory points.
6. *In shop types identified as a batch process in Step 1a or a continuous flow process in Step 1b,* repeat steps 2 to 5 for the remaining PMCs.
7. Periodically repeat steps 1 to 5 to check for changes over time.

[Take in Table 6]

5.3 Pragmatic Validity

To increase pragmatic validity and test our final design, C0 was revisited and the diagnosis repeated by the manager using the refined diagnosis process approximately one year after the initial diagnosis. The repeated diagnosis provided the same result for nearly all of the 21 positions in the flow where inventories could build up. This was to be expected as the product mix and the main approach to production management had not changed. Inventories at two points were attributed to a different cause, which appeared to be related to company-specific circumstances that had changed over time. Another inventory was attributed to the same root cause, but the form of supply variability was now classified as batched instead of predicted fluctuations. This related to the inventory points of rods (j_1) and belts (j_2) preceding the riveting of rods and belts to assemble them. Here belts are waiting for rods supplied in batches. The supply had a batched character, i.e. *batched supply induced assembly waiting time* for belts, even though at some point it could be predicted when the batches of rods would become available. Originally it was classified as *supply anticipation assembly waiting time* of belts. This would have occurred if, for instance, certain orders had combined a high workload for belt production with a low workload for rod production, and belt production had been required to precede rod production to have the last belts available at the same time as the last rods. On the one hand, this change in inventory classification demonstrated the accidental complexity of determining the right cell in the framework for flow problem diagnosis. On the other hand, it showed that the thought process was more important and led to the right context-specific source of variability after Question 3 (iv) of the diagnosis process. Based on the second diagnosis, performed by the manager himself, management proposed three interventions to improve the flow of the shop. The interventions focused particularly on reducing the assembly waiting times of rods and belts prior to riveting since this was the largest inventory. In the past, the company had made the mistake of trying to reduce inventories at exactly this point by increasing capacity, which formed the trigger for this research. Two of the proposed changes now focused on synchronizing the production planning of the two supplying flows. The third intervention was a set-up time reduction to facilitate this synchronization by requiring less batching in the rod production steps.

The first change to synchronize planning was to introduce a signal from the rods department to the belt department as soon as the rods department started production for a certain order. This signal was used to trigger the production of belts. At the time of the diagnosis, the internal planning department was a black box to other departments. This simple rule would at least avoid a needlessly early start to belt production, which normally requires less time than the rods. The planning system could not support an integrated planning approach across the two departments, which led to this alternative of synchronization via a 'release signal'.

The second change to synchronize planning was to reduce the granularity of planned order completion dates. Originally planned completion dates were specified in terms of week numbers. The successive planning logic then provided each of the supply departments with several orders with the same planned end date. As the departments could freely sequence their orders as long as they realized these end dates, the two departments could make completely different choices in terms of which orders to schedule in batches during the first part of a week and which to schedule towards the end of the week. As a result, some parts could have to wait nearly a week for their counterparts from the other department. By issuing more precise end dates, the production sequences of the two departments could be better aligned. Aligning the sequences initiated the third intervention, because reducing the sequencing possibilities also reduces batching possibilities. This in turn caused the need for a set-up reduction project in order to avoid either capacity losses or an increase in inventories upstream.

The above highlights the usefulness of the developed artefact in addressing the problem of why inventories exist. At the same time, it shows that when it comes to the actual realization of flow improvements, a company will have to consider the required effort to improve the flow and the ’cost’ of the current buffer within its specific setting. This is not an easy task as these costs often go beyond the expense of having these items in inventory. Further costs may relate to, for example, reduced responsiveness to customer orders due to long waiting times. When analyzing the improvements proposed by C0 we observed that, originally, the company wanted to take some relatively inexpensive measures to improve coordination between the flows, based on some simple changes to their planning approach. This however in turn triggered the need for set-up time reductions, which required much larger efforts in this specific context. This shows how contextual factors also play an important role in the realization of flow improvements based on our diagnosis. Tenhiälä (2011) showed that similar contingencies apply to capacity management approaches that aim to reduce capacity buffers instead of inventory and time buffers.

6. Discussion

This study designed a generic diagnosis approach for identifying problems with flow in organizations. It first used existing literature to deductively develop a framework that classified flow items based on why they existed. A diagnosis process for applying this framework was then iteratively developed in five companies. During the iterative design process, several important implications emerged, including the need to refine the existing knowledge base of inventories as flow inhibitors and buffers. These will be discussed next.

6.1 Theoretical Embedding of the Artefact

The developed artefact can be used to inform subsequent analytical modelling work. It should therefore not be seen as an alternative but rather as a precursor to more quantitative approaches. Bertrand & Fransoo (2016) referred to Mitroff et al.’s (1974) research cycle when positioning the contribution of simulation modelling to the field of OM. As shown in Figure 6, this cycle can also demonstrate the utility of our diagnosis approach and how it relates to modelling work. Our diagnosis approach answers the question regarding why flow is disrupted and inventories build up. This understanding supports the creation of a conceptual model, which in turn guides the selection of key inputs to the modeling process.

Scientific models assume that the system being modelled is known, meaning they are reliant on the right abstraction from reality being made in the conceptualization phase (Robinson et al., 2010). If the analytical model is not built on the correct underpinning assumptions then the user may arrive at the wrong solution. Our diagnosis process translates a problem situation to a conceptual model, exposing the correct sources of variability that disrupt flow and cause queues. This provides the right inputs to quantitative models, which is particularly important given that the conceptual model organizes thoughts in global, intuitive terms while the modelling process typically relies on formal, analytical skills (Sagasti & Mitroff, 1973; Mitroff et al., 1974). Clearly, these two elements rely on very different thought processes and skills that need connecting. Following the diagnosis, a diagnosis-informed modelling approach is advocated followed by a modelling-informed solution design. After implementation, the diagnosis process can be used again to evaluate the chosen solution.

[Take in Figure 6]

Our framework for flow problem diagnosis also provides an indication on the applicability of particular quantitative models ("3. Scientific Model" in Figure 6). For example, it distinguishes between decoupling point inventories waiting for demand from the next production step and queues for which the production order has already been specified. For decoupling point inventories, a suitable inventory management approach could be selected in a successive diagnosis process (de Vries, 2007). The most appropriate quantitative model could then be selected from the rich inventory modelling literature after our artefact has specified the right source and type of variability to model. For queues, the framework for flow problem diagnosis distinguishes between waiting for capacity and waiting for other flow items. Queueing models can in particular provide support when flow items have been diagnosed as waiting for capacity, while simulation models can be constructed when, for example, items are waiting for other items in complex assembly situations. For queues, the final solutions to improve flow ("4. Solution" in Figure 6) may then be found by applying an appropriate material flow control approach (Graves et al., 1995).

A specific form of real-time modelling that has received attention recently is the so-called digital twin (Grieves & Vickers, 2017; Tao et al., 2018). The use of digital twins recognizes the need to respond to emerging problems quickly. Digital twins however require accurate digital models. These models are typically large-scale system models that are updated in real time by sensor data. As a consequence, the digital twin concept has thus far been largely applied to industries where an accurate digital model of equipment can be used, for example, to support predictive maintenance (Khajavi et al., 2019). In contrast, useful applications that aim to improve the flow of individual flow items would rely on rapid interpretations of what is happening at particular locations without the development of large-scale models of the entire production system. Indeed, operations and decision-making processes that are triggered and controlled by the product itself result in higher quality and greater efficiency when compared to standard operations and external control (Kubler et al., 2016).

Smart products are cognizant of their local context and can negotiate with local manufacturing resources (Bussmann & Schild, 2000, Meyer et al., 2011). Meanwhile, digital encapsulation shifts the loci of design and life-cycle information associated with these products to the level of the individual flow item (Holmström et al. 2019). This provides a distributed and localized context in which our diagnosis approach can be used to support smart products in quickly self-diagnosing any flow disruptions at the item level without the need for large-scale modelling. Holmström et al. (2011) showed how item dwell times can be used to enhance traditional forms of flow diagnosis, simply by distinguishing between slow and fast-moving items at a certain location. Our approach extends Holmström et al. (2011) by allowing smart products to actually self-diagnose the root causes of large item dwell times. Interventions can then be triggered in combination with smart resources and resource agents, such as in Bussmann & Schild (2000) and McFarlane & Bussman (2003), or higher-level inventory and material flow control agents. Again, the appropriate choice of intervention is guided by our diagnosis.

Finally, our diagnosis process is most closely related to the study by Hopp et al. (2007), who focused on diagnosing production line performance. But it can also augment value stream mapping (Hines & Rich, 1997) and improve the resulting conceptual model. Value stream mapping provides a general understanding of value adding and non-value adding operations from the perspective of lean production, whereas our artefact identifies the specific source of variability that hinders the flow of items. This extension is important since it provides a link to the broader lean literature that has provided solution approaches ("4. Solution" in Figure 6) for different sources of variability. In fact, all of the internal lean constructs identified by Shah & Ward (2007) relate to flow improvement solutions within production. For example, *pull production* avoids waiting for missing demand from the next production step while *set-up time reduction* reduces batching-related variability, *preventive maintenance* avoids flow disruptions related to uncertain capacity availability, and *statistical process control* avoids supply availability problems triggered by upstream process steps.

6.2 Insights into Buffer Theory Gained from the Design Process

Our study elaborates on Hopp & Spearman's (2004) statement that all variability in a system is somehow buffered by inventory, capacity and time, and that if variability exists then it will be buffered by one or a combination of these three variability buffers. This interrelatedness of buffers is frequently noted in the literature (e.g. Spearman, 2014) and was also observed in our study.

Our field tests extend existing theory in three ways. First, the field tests established that there is a hierarchical structure between multiple causes of inventory that exist at the same point in the flow. When multiple inputs are missing, the demand input must arrive first, followed by the arrival of flow items that were missing, finally followed by capacity becoming available. A clear example was found in C2, where raw material stock first waited for an order and then, after the order was available, for capacity to become available. This means that missing flow items and capacity can remain hidden until demand occurs. Meanwhile, missing flow items can hide missing capacity, but this does not apply the other way around. This offers a much more granular understanding of the relationship between variability, inventory, capacity, and flow time at a process location than is provided by the OM triangle literature (e.g. Klassen & Menor, 2007), and it has implications for the execution of the diagnosis by either managers or smart objects. Both have to consider the possible answers to the first question of Step 3 – For which process input (missing input) are the flow items in this inventory waiting? – in a strict sequence. First, check for missing demand, and if demand is available then check for missing other flow items, and if these are all there, then capacity must be missing.

A second important finding from our field test is that flexibility in one process input can absorb variability in another input and hide this variability from management. For example, capacity flexibility can be used to avoid congestion caused by variable demand or the variable supply of flow items. Likewise, the flexible supply of flow items may avoid assembly waiting times or batch waiting times when either demand, the supply of other items or capacity is variable. Similarly, the flexible release of production orders may allow demand from the next production step to respond to fluctuations in supply or capacity. Finally, the source of variability can be hidden due to variability propagation (Hopp & Spearman, 1996) through several successive stations. For example, in C1 the source of variability was four stations upstream of the actual inventory occurrence.

The third important finding from our field tests emerged when we found buffers that existed without a missing input; a type of buffer classified as “obvious waste” in Hopp & Spearman (2004). In these cases, there appeared to be a discrepancy between the variability perceived and planned for by management and the variability that is actually realized. We argue that this inventory should only be classified as obvious waste if the buffer is structurally planned to be much larger than required to manage the variability that is realized. Incidentally, buffers will always be too high if they relate to uncertainty (safety stock or safety waiting time) since they have been designed to account for the worst case scenario that is anticipated. In this sense, they are dependent on the risk perception and propensity of management. The difference between realized variability and the variability that has been planned for is also important for the improvement phase. If there is a planned source of variability, the first question should be 'Is the planned amount necessary, given the real variability'? If not, then it is obvious waste and the solution will be relatively simple – to reduce the planned values.

6.3 Implications for Practice

Our framework discriminates between 28 types of inventory that differ from each other in terms of: 1) the missing input; 2) the type of variability creating the missing input; and, 3) the form of variability. These 28 types of inventory and their classification were deduced from the existing literature. As shown in Table 7, which provides a complete overview of the inventory types that have been diagnosed in the five companies, we identified 22 of these 28 inventory types in one or more of the five companies in which the framework was developed. This included identifying all three types of safety stocks, anticipation stocks, and cycle stocks. The two most common forms of inventory were demand safety stock and supply safety stock, which were both observed in all five companies.

[Take in Table 7]

Most importantly, the in-process inventories we identified in each case typically related to causes that the company's existing production management concept did not consider. For example, company C0 had a production management concept focused on releasing amounts of work that fitted available capacity, and we did not observe any items waiting for capacity, as shown in Table 7. Meanwhile, company C3 used an MRP-based production concept to synchronize its assembly flows. Consequently, the flow items in the inventory at the main assembly point were not waiting due to assembly waiting time. Instead, they appeared to be mostly waiting for capacity. This highlights the need to look beyond the chosen production management concept to avoid confirmation bias.

The artefact developed in this study overcomes the confirmation bias introduced by a top-down perspective that is informed by a company's pre-selected production management concept. It not only identifies each possible cause of inventory, focusing on the individual flow item level, but also provides a starting point for developing generic, widely applicable production management concepts (Thürer *et al*., 2020) that consider all possible causes of inventory to realize the best possible level of flow in an organization. For example, MRP, which focuses on synchronization, is often combined with Kanban (see, e.g. Graves *et al*., 1995; Lage Junior & Godinho Filho, 2010) or POLCA (Paired-cell Overlapping Loops of Cards with Authorization; see e.g. Suri, 1998; 2018; Riezebos, 2010) systems to avoid congestion. Similarly, the Theory of Constraints could be adapted to expand its suitability to assembly environments by extending the set of constraints considered to include those related to synchronization. Meanwhile, our diagnosis also provides an important means of assessing criticality, as used to determine which items to buffer in Demand-Driven MRP systems (e.g. Ptak & Smith, 2016).

Finally, all of the companies made exclusive use of traditional manufacturing techniques except for C1, which also made use of some of the opportunities provided by digitalization. Customers could upload drawings for their unique products before automatically receiving a planned delivery date. Meanwhile, the drawing could be directly translated into code for nesting and laser cutting. While the influence of digitalization on the diagnosis process was found to be limited in this company, it is inevitable that digital manufacturing developments will greatly affect the future of OM (Holmström et al. 2019). This in turn may affect the occurrence of certain inventories in the presented framework for flow problem diagnosis. For example, model-based kitting (Khajavi et al., 2018) may reduce the occurrence of supply variability induced assembly waiting times. Meanwhile, the increased application of additive manufacturing may affect the positioning of the CODP (Hedenstierna et al., 2019), decrease the need for batching induced variability, and reduce the number of successive operations (and therefore the number of in-between inventories) required to produce one part. More generally, uncertainty related variability may decrease. Yet producing a variety of products on a production line would still require buffers to synchronize supply, capacity availability, and demand even if variability is more predictable with digitalization.

6.4 Insights from the Design Science Approach

Our research has followed a design science research approach where each research step links into the design science phases proposed by Peffers et al. (2007) and Holmström et al. (2009a). However, while conventional DSR typically starts with a context-specific design, which then becomes more generic through iterative solution refinements, we started with a generic design that was then applied to different contexts for refinement. This is motivated by the existence of prior knowledge, which allowed for deductively deriving our framework for flow problem diagnosis. The refinement of our artefact, consisting of this framework for flow problem diagnosis and an inductively derived diagnosis process, then focused on ensuring that the artefact will work after contextualization in different process and product settings along Hayes & Wheelwright's (1979) volume/variety spectrum of production types. This highlights that the use of DSR does not have to be limited to the generalization of context-specific solutions; it also applies to the contextualization of generic solutions.

7. Conclusions

Problems with flow manifest themselves in the form of inventory, which is symptomatic of variability and a lack of synchronization between the availability of capacity, demand, and flow items. In practice, flow issues are ubiquitous in all types of OM settings and are a predominant concern of many OM theories. Improving flow has consequently been a core OM theme, and one that is set to become even more important following contemporary developments in manufacturing, such as smart products and digital encapsulation that enable new control concepts such as multi-agent holonic control. Yet, despite the importance of flow, the literature has lacked a structured and independent approach to diagnosing the root causes of inventories based on the underlying buffer mechanisms and forms of variability. Rather, much diagnosis of flow problems is strongly influenced by a manager's *a priori* choice of production management concept. This confirmation bias may lead to a misguided intervention, as in our initial case company. Therefore, and to widen the focus of flow improvement actions beyond the scope of a singular concept, a comprehensive approach to diagnosing flow problems from the item level that is independent from any production management concept was needed. This study has used a DSR approach to address this shortcoming.

First, this study used the concept of variability buffers to develop a framework for flow problem diagnosis that distinguishes between 28 different inventory types. The framework is embedded in the dimensions of traditional inventory classifications. To facilitate the application of this framework for flow problem diagnosis we then developed a diagnosis process. The generic design of this diagnosis process can be used to identify the root causes of inventories across a broad set of traditional production environments. It builds on systematically answering a sequence of four key questions: For which process input (missing input) are the flow items in this inventory waiting? Which input is a source of variability that causes missing inputs? What form does this variability take? And, what causes this variability in this specific context? This generic design was used in five cases to determine its usability and validity across a broad set of traditional production environments.

This paper contributes to OM theory in multiple ways. Our designed artefact supports conceptual modelling by providing a base from which to choose appropriate analytical models to quantify the impact of potential solutions. It guides the choice of appropriate quantitative models and provides the foundations for extending the use of digital twins for rapid localized diagnosis without the need for large-scale modelling. The diagnosis process can, for example, be encapsulated into smart products that use digitally encapsulated information for self-diagnosis. Meanwhile, the artefact also provides a basis for the development of generic production management systems that combine the foci of multiple existing systems. Finally, the artefact's application in five companies provided a detailed understanding of buffering, which significantly extends existing theory on buffering mechanisms while providing a missing ingredient for flow improvement practices.

7.1 Limitations and Future Research

Our study has presented a new generic design for diagnosing the root causes of inventories and flow problems in organizations, which should be broadly applicable to a wide range of industry sectors. However, we remained entrenched in conventional manufacturing contexts. Future research could study the implications of digitalization in manufacturing and extend the framework to service contexts, including healthcare where Johnson et al. (2020) recently conducted a more solutions oriented DSR study. In most services, flow items are the customers themselves, so flow items that wait for demand will not occur. But all other missing inputs, variability sources, and causes presented in our framework would still apply. Nonetheless, a different emphasis might be needed. For example, in emergency departments, delays are more likely to be caused by waiting for multiple parallel capacity resources to arrive (e.g. doctors, nurses, etc.) rather than waiting for multiple flow items. The current framework does not consider parallel capacity resources. Our process could also be used to diagnose flow problems at a broader supply chain level. However, a supply chain represents a larger and more complex system. For example, this study assumed inventory falls under the ownership of a single company. In a more complex supply chain, ownership may change from supplier to buyer while inventory may accumulate in transit. Although we believe our diagnosis process can be applied to these contexts, future research is required to identify necessary refinements to the current design.

Another limitation is our sole focus on the process of diagnosis. While this is an important step, future research could focus on connecting it with other parts of Figure 6, including the design of solutions. For example, each cell in our framework for flow problem diagnosis refers to a specific cause of inventory. The high granularity of our framework, compared to previous classifications, could enable the development of a generic design for the construction of solutions drawing, for example, on those provided in the lean and material flow control literature.

Meanwhile, our categories for the form of variability appeared to be sufficient. However, given the importance of planning, a greater degree of granularity might be useful. For example, in the field of healthcare, Litvak & Long (2000) distinguished between artificial variability and natural variability. Artificial variability is defined as variability due to dysfunctional management, i.e. caused by controllable human actions, while natural variability is further subdivided into clinical, flow, and professional variability according to three categories of uncontrollable causes in care environments. Similar variability types could be developed for production environments to further guide the diagnosis of root causes of inventory and more precisely indicate solutions for flow improvement.

Finally, our study has focused on improving flow, but it is important to acknowledge that flow may not be the only priority for an organization, and it may be necessary to trade-off flow against other considerations, such as quality. Thus, future research could investigate how the design can be elaborated to interact with other system priorities.

References

Akkermans, H., Oppen, W. Van, Wynstra, F., & Voss, C. (2019). Contracting outsourced services with collaborative key performance indicators. *Journal of Operations Management*, 65, 22–47.

Bertrand, J. W. M., & Fransoo, J. C. (2016). Modelling and simulation. In *Research methods for operations management* (pp. 306-346); Karlsson, C. (ed), Routledge.

Bertrand, J. W. M., Wortmann, J. C., & Wijngaard, J. (1990). *Production control: A structural and design oriented approach*. Elsevier Science Inc.

Bussmann, S., & Schild, K. (2000). Self-organizing manufacturing control: An industrial application of agent technology. In *Proceedings Fourth International Conference on Multi Agent Systems* (pp. 87-94). IEEE.

Cachon, G., & Terwiesch, C. (2019). *Matching supply with demand: An introduction to operations management*, McGraw-Hill, 4th edition.

Calle, M., González-R, P. L., Leon, J. M., Pierreval, H., & Canca, D. (2016). Integrated management of inventory and production systems based on floating decoupling point and real-time information: A simulation based analysis. *International Journal of Production Economics*, 181, 48-57.

Chandrasekaran, A., de Treville, S., & Browning, T. (2020). Intervention-based research (IBR)—What, where, and how to use it in operations management. *Journal of Operations Management*, in print

de Treville, S., and Antonakis, J. (2006). Could lean production job design be intrinsically motivating? Contextual, configurational, and levels-of-analysis issues. *Journal of Operations Management*, 24, 99-123.

de Vries, J. (2007). Diagnosing inventory management systems: An empirical evaluation of a conceptual approach. *International Journal of Production Economics*, 108(1-2), 63-73.

Forrester, J.W. (1961). *Industrial Dynamics*, MIT Press, Cambridge MA.

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

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, 37(2), 337-355.

Grieves M., & Vickers J. (2017). Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In: Kahlen FJ., Flumerfelt S., Alves A. (eds) *Transdisciplinary Perspectives on Complex Systems*. Springer.

Goldratt, E. M., & Cox, J. (1984). *The goal: Excellence in manufacturing*. Great Barrington, MA: North River Press.

Hayes, R.H., & Wheelwright, S.C. (1979). Link manufacturing process and product life cycle. *Harvard Business Review,* 57(1), 133-140.

Hedenstierna, C. P. T., Disney, S. M., Eyers, D. R., Holmström, J., Syntetos, A. A., & Wang, X. (2019). Economies of collaboration in build-to-model operations. *Journal of Operations Management*, 65(8), 753-773.

Hevner, A.R. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, 19(2), 87–92.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information system research. *MIS Quarterly*, *28*(1), 75–105.

Hines, P., & Rich, N. (1997). The seven value stream mapping tools. *International journal of operations & production management*, 17(1), 46-64.

Hines, P., Holweg, M., & Rich, N. (2004). Learning to evolve: a review of contemporary lean thinking. *International Journal of Operations and Production Management*, 24(10), 994-1011.

Hoekstra, S., & Romme, J. (1992). *Integral logistic structures: developing customer-oriented goods flow*. Industrial Press Inc.

Holmström, J., Ketokivi, M., & Hameri, A. P. (2009a). Bridging practice and theory: A design science approach. *Decision Sciences*, 40(1), 65-87.

Holmström, J., Kajosaari, R., Främling, K., & Langius, E. (2009b). Roadmap to tracking based business and intelligent products. *Computers in Industry*, 60(3), 229-233.

Holmström, J., Tenhiälä, A., & Kärkkäinen, M. (2011). Item dwell time in project inventories: a field experiment. *Computers in Industry*, *62*(1) 99-106.

Holmström, J., Holweg, M., Lawson, B., Pil, F. K., & Wagner, S. M. (2019). The digitalization of operations and supply chain management: Theoretical and methodological implications. *Journal of Operations Management*, 65(8), 728-734.

Holweg, M. (2007). The genealogy of lean production. *Journal of Operations Management*, 25(2), 420-437.

Hopp, W., & Spearman, M. (1996). *Factory physics: foundations of factory management*. Chicago: Irwin.

Hopp, W. J., & Spearman, M. L. (2004). To pull or not to pull: what is the question? *Manufacturing and Service Operations Management*, 6(2), 133-148.

Hopp, W. J., Iravani, S. M. R. and Shou, B. (2007). A Diagnostic Tree for Improving Production Line Performance. *Production and Operations Management*, 16, 77–92.

Johnson, M., Burgess, N., & Sethi, S. (2020). Temporal pacing of outcomes for improving patient flow: Design science research in a National Health Service hospital. *Journal of Operations Management*, 66(1-2), 35-53.

Khajavi, S. H., Motlagh, N. H., Jaribion, A., Werner, L. C., & Holmström, J. (2019). Digital twin: vision, benefits, boundaries, and creation for buildings. *IEEE Access*, *7*, 147406-147419.

Khajavi, S. H., Baumers, M., Holmström, J., Özcan, E., Atkin, J., Jackson, W., & Li, W. (2018). To kit or not to kit: Analysing the value of model-based kitting for additive manufacturing. *Computers in Industry*, 98, 100-117.

Klassen, R. D., & Menor, L. J. (2007). The process management triangle: An empirical investigation of process trade-offs. *Journal of Operations Management*, 25(5), 1015-1034.

Kubler, S., Holmström, J., Främling, K., & Turkama, P. (2016). Technological theory of cloud manufacturing. In *Service orientation in holonic and multi-agent manufacturing* (pp. 267-276). Springer, Cham.

Kusiak, A., (2018). Smart manufacturing. *International Journal of Production Research*, *56*(1-2), 508-517.

Lage Junior, M., M. Godinho Filho. (2010). Variations of the kanban system: Literature review and classification. *International Journal of Production Economics*, 125, 13-21.

Little, J.D.C. (1961), A proof for the queuing formula: L= λ W, *Operations Research*, 9(3), 383-387.

Litvak, E., & Long, M. C. (2000). Cost and quality under managed care: Irreconcilable differences. *The American Journal of Managed Care*, 6(3), 305-12.

Lovejoy, W.S. (1998). Integrated operations: A proposal for operations management teaching and research. *Production and Operations Management*, 7(2), 106-124.

McFarlane, D. C., & Bussmann, S. (2003). Holonic manufacturing control: rationales, developments and open issues. In *Agent-based manufacturing* (pp. 303-326). Springer, Berlin, Heidelberg.

Meyer, G. G., Wortmann, J. C., & Szirbik, N. B. (2011). Production monitoring and control with intelligent products. *International Journal of Production Research*, *49*(5), 1303-1317.

Miller, J. G., & Sprague, L. G. (1975). Behind growth in material requirements planning. *Harvard Business Review*, 53(5), 83-91.

Mitroff, I. I., Betz, F., Pondy, L. R., & Sagasti, F. (1974). On managing science in the systems age: two schemas for the study of science as a whole systems phenomenon. *Interfaces*, 4(3), 46-58.

Monden, Y. (1981). What makes the Toyota production system really tick. *Industrial Engineering*, 13(1), 36-46.

Monden, Y. (2011). *Toyota production system: an integrated approach to just-in-time*. CRC Press.

Monhemius, W.V., & Durlinger, P. P. (1985). *Logistiek management*. Kluwer, The Netherlands (In Dutch).

Newman, W. R., Hanna, M., & Maffei, M. J. (1993). Dealing with the uncertainties of manufacturing: flexibility, buffers and integration. *International Journal of Operations and Production Management*, 13, 19-19.

Ohno, T. (1988). *Toyota Production System: Beyond Large-Scale Production*, 1st Ed., Productivity Press.

Olhager, J. (2003). Strategic positioning of the order penetration point. *International Journal of Production Economics*, 85(3), 319-329.

Oliva, R. (2019). Intervention as a research strategy. *Journal of Operations Management*, 65(7), 710-724.

Olsen, T.L. & Tomlin, B., (2020). Industry 4.0: opportunities and challenges for operations management. *Manufacturing & Service Operations Management*, *22*(1), 113-122.

Orlicky, J. (1975). MRP: Material Requirements Planning. In *The New Way of Life in Production and Inventory Management*. McGraw Hill.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45-77.

Pohl, R. ed. (2004). *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory*. Psychology Press.

Ptak, C., & Smith, C., (2016). *Demand Driven Material Requirements Planning (DDMRP)*. Norwalk, CT: Industrial Press.

Rahman, S. U. (1998). Theory of constraints: a review of the philosophy and its applications. *International Journal of Operations and Production Management*, 18(4), 336-355.

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

Robinson, S., Brooks, R., Kotiadis, K., & Van Der Zee, D. J. (Eds.). (2010). *Conceptual modeling for discrete-event simulation*. CRC Press.

Sagasti, F. R., & Mitroff, I. I. (1973). Operations research from the viewpoint of general systems theory. *Omega*, *1*(6), 695-709.

Schmidt, G.M. (2005) The OM Triangle. *Operations Management Education Review*, 1(1), 87-104.

Schwarz, L.B. (1998). A new teaching paradigm: The information/control/buffer portfolio. *Production and Operations Management*, 7(2), 125-131.

Seawright, J. & Gerring, J. (2008). Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options. *Political Research Quarterly*, 61(2), 294-308.

Shah, R., & Ward, P. T. (2007). Defining and developing measures of lean production. *Journal of Operations Management*, 25(4), 785-805.

Sharman, G. (1984). The rediscovery of logistics. *Harvard Business Review*, 62(5), 71-79.

Shingo, S. (1989). *A Study of the Toyota Production System from an Industrial Engineering Viewpoint*, Cambridge, MA: Productivity Press.

Slack, N., & Brandon-Jones, A. (2019). *Operations management*, Prentice Hall, 9th Edition.

Spearman, M. L. (1997). On the theory of constraints and the goal system. *Production and Operations Management*, 6(1), 28-33.

Spearman, M. (2014). Of physics and factory physics. *Production and Operations Management*, 23(11), 1875-1885.

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

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.

Tao, F., Zhang, H., Liu, A. & Nee, A.Y. (2018). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, *15*(4), 2405-2415.

Taylor, F. W. (1919). *The principles of scientific management*. Harper & brothers.

Tenhiälä, A. (2011). Contingency theory of capacity planning: The link between process types and planning methods. *Journal of Operations Management*, 29(1-2), 65-77.

Thürer, M., Fernandes, N.O., and Stevenson, M. (2020). Production Planning and Control in Multi-Stage Assembly Systems: An Assessment of Kanban, MRP, OPT (DBR) and DDMRP by Simulation. *International Journal of Production Research*, (in print).

van Aken, J. (2014). Two Hurdles to Take for Maximum Impact of Design Science Research in the IS-Field, in Helfert, M., Donnellan, B. and Kenneally, J. (Eds.), *Design Science: Perspectives from Europe*, Springer International Publishing, pp. 28–40.

van Aken, J.E. & Romme, G. (2012). A Design Science Approach to Evidence-Based Management, in Rousseau, D.M. (Ed.), *The Oxford Handbook of Evidence-Based Management*, Oxford University Press.

van Aken, J., Chandrasekaran, A., Halman, J., & Ketokivi, M. (2016). Conducting and publishing design science research Inaugural essay of the design science department of the Journal of Operations Management. *Journal of Operations Management*, *47–48*, 1–8.

van Donk, D. P. (2001). Make to stock or make to order: The decoupling point in the food processing industries. *International Journal of Production Economics*, 69(3), 297-306.

Voss, C.A. (2007). Learning from the first Operations Management textbook. *Journal of Operations Management*, 25(2), 239-247.

Wagner, C. (1993). Problem solving and diagnosis. *Omega*, *21*(6), 645-656.

Watson, K. J., Blackstone, J. H., & Gardiner, S. C. (2007). The evolution of a management philosophy: The theory of constraints. *Journal of Operations Management*, 25(2), 387-402.

Zipkin, P. H. (2000). *Foundations of inventory management* (Vol. 2). McGraw-Hill.

Appendix 1: Definitions of the Types of Inventory and Associated Route Cause from Table 3

Demand Safety Stock

Flow items have been produced in advance because of unplanned fluctuations in demand.

Demand Anticipation Stock (also Seasonal Stock)

Flow items have been produced in advance because of planned fluctuations in demand.

Demand Cycle Stock

Flow items have been produced in advance because demand (for the next step) arrives in batches.

Supply Safety Stock

Flow items have been produced in advance because of unplanned fluctuations in the supply of flow items.

Supply Anticipation Stock

Flow items have been produced in advance because of planned fluctuations in the supply of flow items.

Supply Cycle Stock

Flow items have been produced in advance because of batching at the preceding step.

Capacity Safety Stock

Flow items have been produced in advance because of unplanned fluctuations in available capacity.

Capacity Anticipation Stock

Flow items have been produced in advance because of planned fluctuations in available capacity.

Capacity Cycle Stock

Flow items have been produced in advance because they are produced in batches at the next step.

Demand Uncertainty Induced Congestion

Flow items wait as capacity is not sufficient to handle the unplanned peaks in demand.

Demand Anticipation Induced Congestion

Flow items wait as capacity is not sufficient to handle the planned peaks in demand.

Batched Demand Induced Congestion

Flow items wait as capacity is not sufficient to handle batched demand.

Supply Uncertainty Induced Congestion

Flow items wait for capacity because of unplanned fluctuations in their supply.

Supply Anticipation Induced Congestion

Flow items wait for capacity to become available because of planned fluctuations in their supply.

Batched Supply Induced Congestion

Flow items wait for capacity because they became available as a batch.

Capacity Uncertainty Induced Congestion

Flow items wait because of unplanned fluctuations in available capacity.

Capacity Anticipation Induced Congestion

Flow items wait because of planned fluctuations in available capacity.

Batch Transformation Induced Congestion

Flow items wait for the next 'event' of capacity becoming available.

Demand Uncertainty Induced Assembly Waiting Time

Flow items wait to be assembled with other items, as their supply follows unplanned demand fluctuations more closely.

Demand Anticipation Assembly Waiting Time

Flow items wait to be assembled with other items, as their supply follows planned demand fluctuations more closely.

Batched Demand Induced Assembly Waiting Time

Flow items wait to be assembled with other items, as their supply follows batched demand more closely.

Supply Uncertainty Induced Assembly Waiting Time

Flow items wait to be assembled with other items, because of uncertainty in supply.

Supply Anticipation Assembly Waiting Time

Flow items wait to be assembled with other items, because of planned differences in the timing of their supply.

Batched Supply Induced Assembly Waiting Time

Flow items wait to be assembled with other items, because some items are supplied in batches.

Capacity Uncertainty Induced Assembly Waiting Time

Flow items wait to be assembled with other items, as their supply follows unplanned capacity fluctuations more closely.

Capacity Anticipation Assembly Waiting Time

Flow items wait to be assembled with other items, as their supply follows planned capacity fluctuations more closely.

Batched Transformation Induced Assembly Waiting Time

Flow items wait to be assembled with other items, as their supply follows the batched processing in the next step more closely.

Batch Waiting Time

Flow items wait for each other because the demand from the next step takes place in batches.

Table 1: Summary of Production Management Concepts: Control Focus, Targeted Flow Problems, and Diagnosis Support

Traditional Concept	**Control Focus**	**Causes of Flow Problems Targeted by This Focus**	**Diagnosis Support**	**Gap**
Statistical Inventory Control (SIC)	Determining stocks to cover future demand	Decoupling stocks (flow items waiting for demand)	Inventory models to determine optimized parameter settings	Concepts focus on a limited group of inventory causes and do not identify the root causes determining why each individual flow item waits in inventory
Lean Production	Realizing one-piece flow	Batch waiting times (flow items waiting for other flow items of the same type)	Value stream mapping to identify inventory points and sizes	
Material Requirements Planning (MRP)	Synchronizing the availability of parts to assemble	Assembly waiting times (flow items waiting for other flow items of a different type)	Input/output models to monitor lead time offsets	
Theory of Constraints (TOC)	Exploiting bottlenecks with minimal buffers	Congestion (flow items waiting for capacity)	Buffer management to indicate appropriate buffer sizes	

Table 2: Summary of Organizations C0 to C4 and their Role in the Study

	Company				
	C0	**C1**	**C2**	**C3**	**C4**
Attributes					
Company Size	$14m turnover, 80 employees	$14m turnover, 60 employees	$15m turnover; 100 employees	$2.85bn; 90,000 employees	$5.4m; 30 employees
Shop Type	Batch	Job shop	Batch	Assembly line	Continuous flow
Dominant Operation Mode(s)	Make-to-order	Make-to-order	Make-to-order & assemble-to-order	Assemble-to-order & make-to-stock	Make-to-stock & make-to-order
Products	Conveyor belts for harvesting equipment	Customer-specified parts produced from sheet metal	Outdoor play equipment (e.g. swings and climbing frames)	Air Conditioners	Bags for bag-in-box systems (e.g. dairy and wine bags)
Key informant	Production Manager	Supply Chain Manager	Technical & Operations Director	Director of Industrial Engineering	Plant Manager
Location	The Netherlands	The Netherlands	UK	China	U.S.A.
Purpose					
Initial motivation for the research	X				
Test the initial diagnosis process	X				
Validate the framework for flow problem diagnosis	X	X	X	X	X
Implement context-specific improvements	X				
Further develop the diagnosis process		X	X	X	X

Table 3: Framework for Flow Problem Diagnosis

<table>
<tr><th colspan="2"></th><th colspan="4">Missing Input</th></tr>
<tr><th rowspan="2">Main Source of Variability</th><th rowspan="2">Form of Variability</th><th rowspan="2">Demand</th><th rowspan="2">Capacity
[Service Desk or Platform Waiting Time]</th><th colspan="2">Flow Items</th></tr>
<tr><th>Different Type
[Assembly Waiting Time]</th><th>Similar Type
[Batch Waiting Time]</th></tr>
<tr><td rowspan="3">**Demand**</td><td>**Uncertainty**</td><td>Demand Safety Stock</td><td>Demand Uncertainty Induced Congestion</td><td>Demand Uncertainty Induced Assembly Waiting Time</td><td rowspan="9">Batch Waiting Time</td></tr>
<tr><td>**Predicted Fluctuation**</td><td>Demand Anticipation Stock</td><td>Demand Anticipation Induced Congestion</td><td>Demand Anticipation Induced Assembly Waiting Time</td></tr>
<tr><td>**Batched**</td><td>Demand Cycle Stock</td><td>Batched Demand Induced Congestion</td><td>Batched Demand Induced Assembly Waiting Time</td></tr>
<tr><td rowspan="3">**Supply of Flow items**</td><td>**Uncertainty**</td><td>Supply Safety Stock</td><td>Supply Uncertainty Induced Congestion</td><td>Supply Uncertainty Induced Assembly Waiting Time</td></tr>
<tr><td>**Predicted Fluctuation**</td><td>Supply Anticipation Stock</td><td>Supply Anticipation Induced Congestion</td><td>Supply Anticipation Induced Assembly Waiting Time</td></tr>
<tr><td>**Batched**</td><td>Supply Cycle Stock</td><td>Batched Supply Induced Congestion</td><td>Batched Supply Induced Assembly Waiting Time</td></tr>
<tr><td rowspan="3">**Capacity**</td><td>**Uncertainty**</td><td>Capacity Safety Stock</td><td>Capacity Uncertainty Induced Congestion</td><td>Capacity Uncertainty Induced Assembly Waiting Time</td></tr>
<tr><td>**Predicted Fluctuation**</td><td>Capacity Anticipation Stock</td><td>Capacity Anticipation Induced Congestion</td><td>Capacity Anticipation Induced Assembly Waiting Time</td></tr>
<tr><td>**Batched**</td><td>Capacity Cycle Stock</td><td>Batch Transformation Induced Congestion</td><td>Batch Transformation Induced Assembly Waiting Time</td></tr>
<tr><td colspan="2"></td><td>Decoupling stocks</td><td colspan="3">Queues</td></tr>
</table>

Table 4: Summary of Inventory Points and Inventory Types in C0

	Inventory Type								
	Demand Safety Stock	Demand Anticipation Stock	Supply Cycle Stock	Demand Uncertainty Induced Congestion	Batch Supply Induced Congestion	Batch Transformation Induced Congestion	Supply Uncertainty Induced Assembly Waiting Time	Supply Anticipation Induced Assembly Waiting Time	Batch Waiting Time
Missing Input	Demand	Demand	Demand	Capacity	Capacity	Capacity	Flow Items	Flow Items	Flow Items
Source of Variability	Demand	Demand	Flow Items	Demand	Flow Items	Capacity	Flow Items	Flow Items	Flow Items
Appearance of Variability	Uncertainty	Predicted Fluctuation	Batch	Uncertainty	Batch	Batch	Uncertainty	Predicted Fluctuation	Batch
Inventory Point									
a	x								x
b					x				
c						x			
d						x			
e				x		x			
f				x					
g			X						
h	x	x							
i					x				
j_1							x	x	
j_2							x	x	
k									
l			X						
m					x				
n					x				
p	x		X						
q					x				
r			X						
s			X						
t	x		X						
v_1								x	x
v_2								x	x
v_3	x		X					x	x
W						x			

Table 5: Summary of Evaluation from the Design Cycle in C0

Step where Learning Occurred	**Lesson Learned**	**Consequences for the Diagnosis Process**
Diagnosis Step 3	Lesson 1: It is important not to automatically move to the next inventory point when a source of variability has been found as there can be multiple causes of inventory waiting for demand at a single inventory point.	Added a note to the end of Step 3 to allow for multiple answers to each question.
	Lesson 2: Care has to be taken in distinguishing all possible product-market combinations before the diagnosis is finalized.	Extended Step 1 and added a 5th (repeat) step to the diagnosis process.
	Lesson 3: It is important to repeat the diagnosis to account for seasonal and non-seasonal differences over time.	Added a 6th (periodic repeat) step to the diagnosis process.
	Lesson 4: Flow items can remain at a given inventory point even after a missing input has become available because they have to wait for another missing input (i.e. multiple inputs were missing).	Requires careful treatment of findings: one inventory type may hide another.
	Lesson 5: The causes for the position of a decoupling point do not necessarily relate to variability.	Requires careful treatment of information, as causes of the existence of inventory should be distinguished from causes of the position of inventory.

Table 6: Summary of Evaluations from the Design Cycles in C1 to C4

Step where Learning Occurred	**Lesson Learned**	**Consequences for the Diagnosis Process**
Diagnosis Step 1	Lesson 1: The approach taken when identifying inventory points for subsequent diagnosis needs to be adapted to the shop type and its complexity so that it is either resource-based (job shop), product-flow focused (batch process), or product-based (continuous flow).	Differentiated Step 1.
Diagnosis Step 2	Lesson 2: Identifying CODPs is crucial for batch processes, but less so for job shop or continuous flow environments.	Step 2 can be skipped for job shop and continuous flow environments.
Diagnosis Step 3	Lesson 3: One type of variability at a single inventory point may relate to two different sources.	The context-specific Question 3 (iv) should allow for multiple answers.
	Lesson 4: An observed buffer may not relate to the real variability but to the planned variability.	Extension to Question 2: 'Which input is a source of *planned or actual* variability that causes missing inputs*?*'
	Lesson 5: Variability can have its root causes at operations that are further upstream or downstream.	Extension to the note at the end of Step 3 to look further upstream and downstream.
	Lesson 6: A significant part of the waiting times may occur as part of the order book before the first operation and before material is assigned to an order.	Requires careful treatment of findings because waiting as part of the order book can be positive.

Table 7: A Summary of the Inventory Types Identified in C0 to C4

<table>
<tr><th rowspan="3">Main Source of Variability</th><th rowspan="3">Form of Variability</th><th colspan="4">Missing Input</th></tr>
<tr><th rowspan="2">Demand</th><th rowspan="2">Capacity
[Service Desk or Platform Waiting Time]</th><th colspan="2">Flow Items</th></tr>
<tr><th>Different Type
[Assembly Waiting Time]</th><th>Similar Type
[Batch Waiting Time]</th></tr>
<tr><td rowspan="3">**Demand**</td><td>**Uncertainty**</td><td>C0, C1, C2, C3, C4</td><td>C1, C2, C4</td><td>C2</td><td rowspan="9">C0, C1, C2, C4</td></tr>
<tr><td>**Predicted Fluctuation**</td><td>C0, C4</td><td>C4</td><td></td></tr>
<tr><td>**Batched**</td><td>C0, C2</td><td>C2, C3, C4</td><td></td></tr>
<tr><td rowspan="3">**Supply of Flow items**</td><td>**Uncertainty**</td><td>C1, C2, C4</td><td>C1, C2, C3, C4</td><td>C0, C1, C2</td></tr>
<tr><td>**Predicted Fluctuation**</td><td>C1, C2</td><td>C1</td><td>C2</td></tr>
<tr><td>**Batched**</td><td>C0, C1, C2, C3, C4</td><td>C1, C2, C4</td><td>C0</td></tr>
<tr><td rowspan="3">**Capacity**</td><td>**Uncertainty**</td><td>C0, C1</td><td>C2, C4</td><td></td></tr>
<tr><td>**Predicted Fluctuation**</td><td>C0</td><td></td><td></td></tr>
<tr><td>**Batched**</td><td>C0, C1</td><td>C1, C2, C4</td><td></td></tr>
<tr><td></td><td></td><td>Decoupling stocks</td><td colspan="3">Queues</td></tr>
</table>

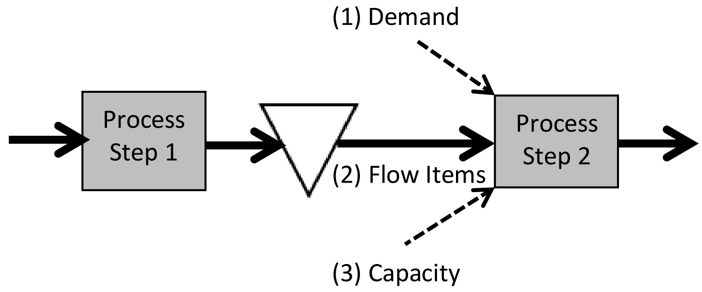

Figure 1: The Three Inputs to a Process – Flow Items, Capacity, and Demand

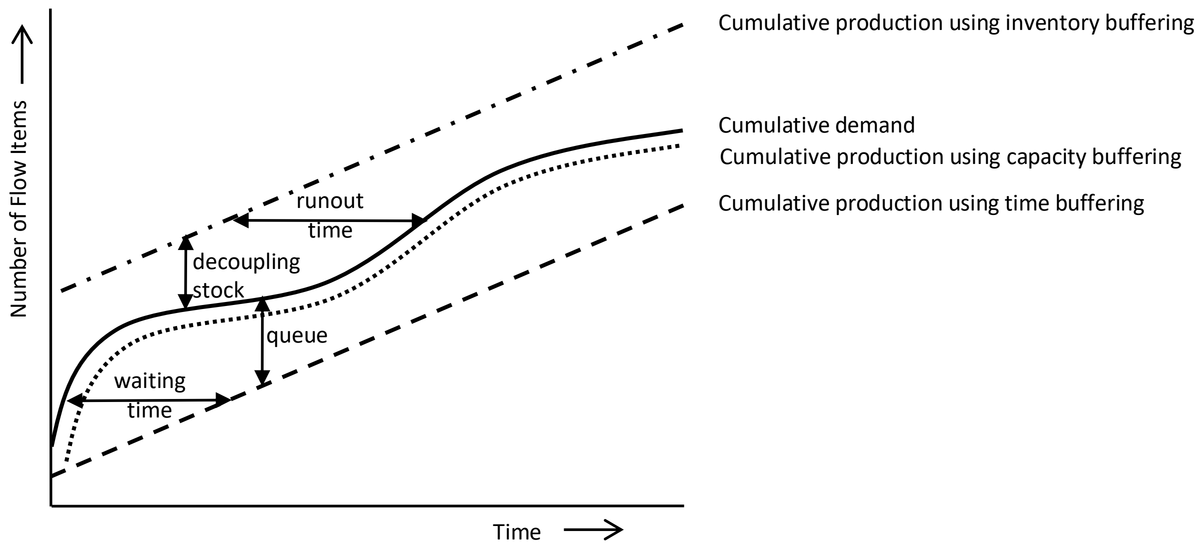

Figure 2: Inventory Buffering and Time Buffering Resulting in Decoupling Stocks and Queues

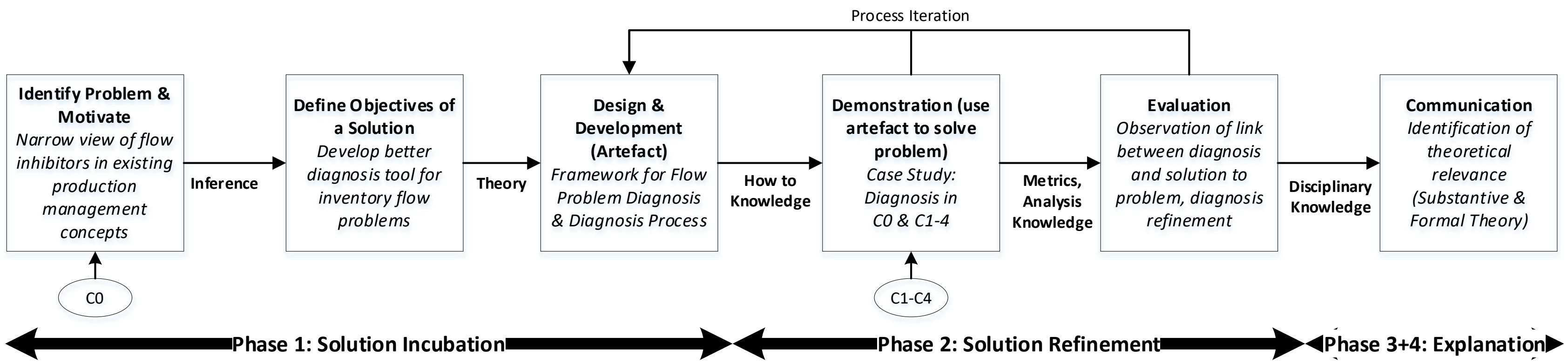

Figure 3: Applied Design Science Process, drawing on Peffers et al. (2007) and Holmström et al. (2009)

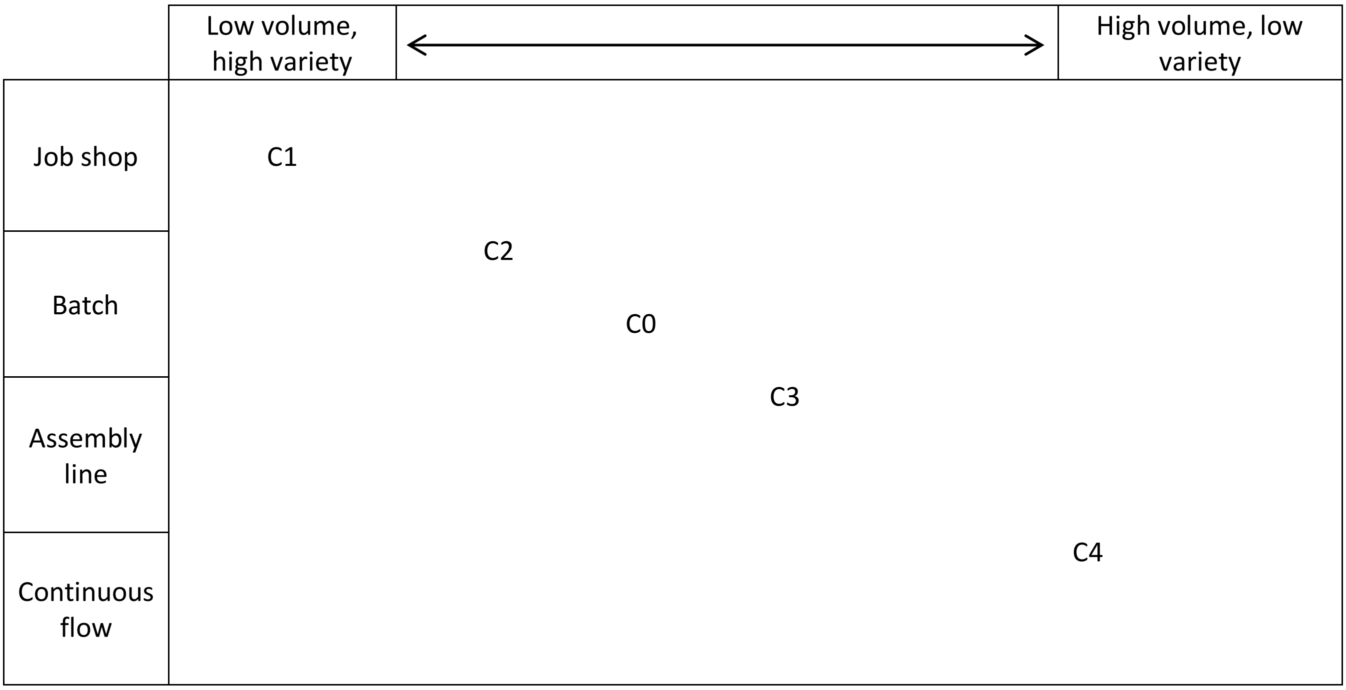

Figure 4: Positioning C0 to C4 on Hayes & Wheelwright's (1979) Classification

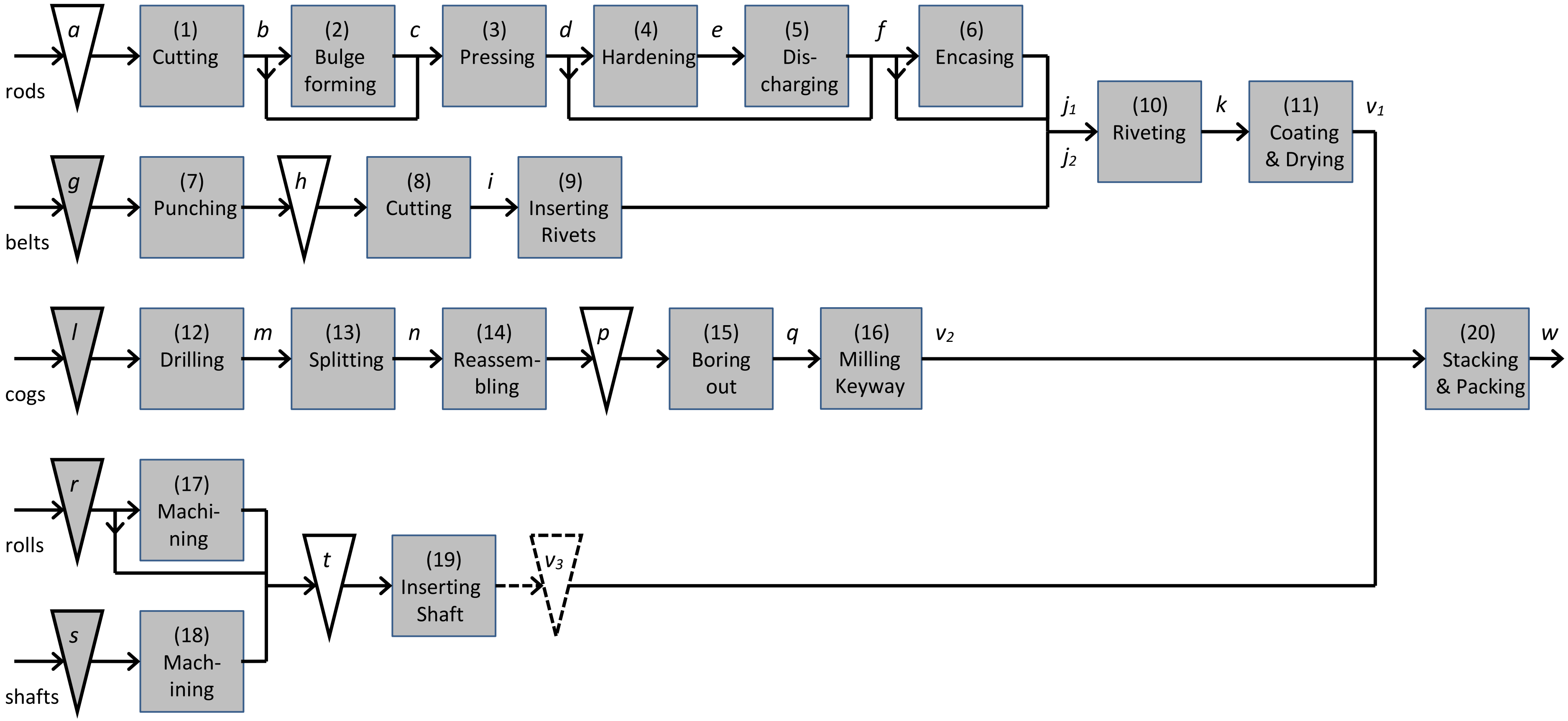

Figure 5: Process Map of the Field Test in C0

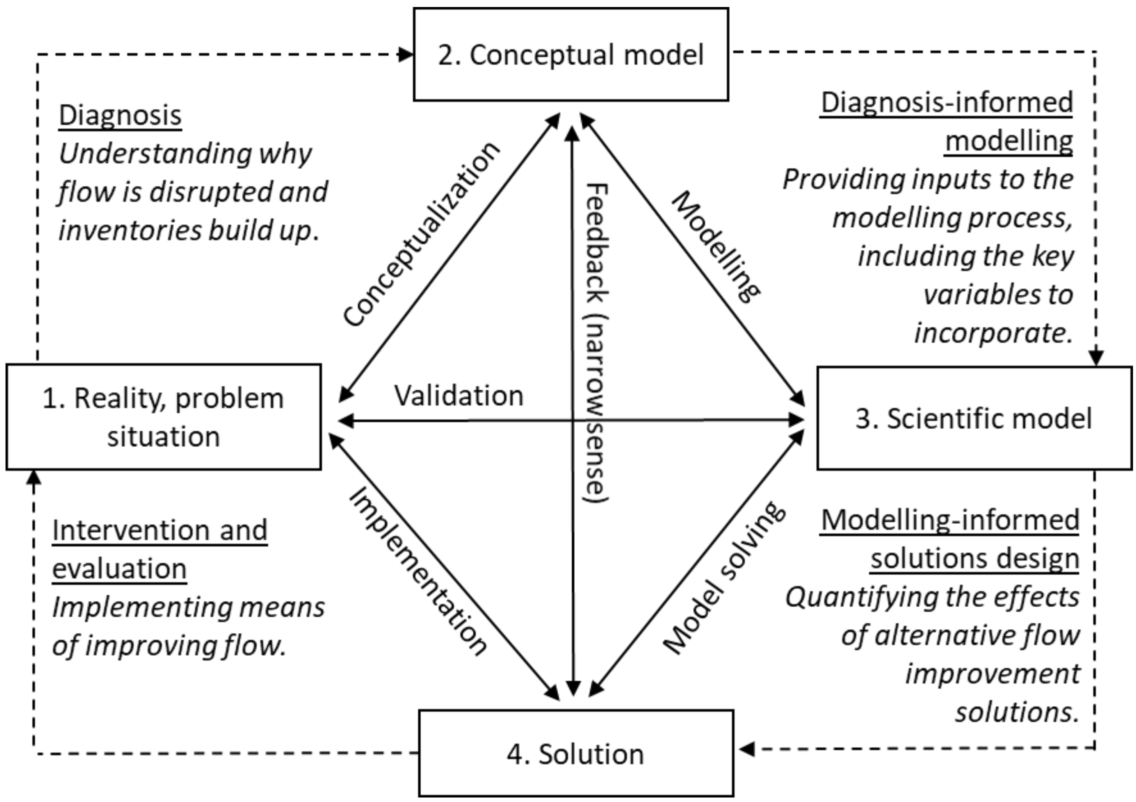

Figure 6: Positioning the Contribution of the Diagnosis Process (adapted from Mitroff et al., 1974)