Available online at www.sciencedirect.com
ScienceDirect
Journal of Operations Management 25 (2007) 1–13
JOURNAL OF OPERATIONS MANAGEMENT
www.elsevier.com/locate/jom

Bipolarity in reactions to operational 'constraints': OM bugs under an OB lens

Elliot Bendoly [a,*], Daesik Hur [b,1]

[a] *Emory University, 1300 Clifton Road, Atlanta, GA 30307, United States*
[b] *Yonsei School of Business, Yonsei University, 134 Seodaemun Gu, Shinchon Dong, Seoul 120-749, Korea*

Received 1 May 2005; received in revised form 1 August 2005; accepted 12 August 2005
Available online 22 December 2005

Abstract

In the 1980 inaugural issue of the *Journal of Operations Management*, Powell and Johnson stressed the need to introduce behavioral factors into research models of operational processes and performance. While some progress has been made since then, contemporary authors argue that limited dialogue between operations management and behavioral researchers continues to restrict the interpretability of such studies. Drawing on expertise from both operations and behavioral research, and motivated by a case example, we intend to further this dialogue. In doing so, we outline interdisciplinary commonalities and several methodological pitfalls that those studying the links between resource constraints and operational performance should take into account when developing new research. These issues range from appropriate model conceptualization to operationalization for empirical studies. Recommendations for appropriately confronting methodological concerns related to these issues are provided.

Keywords: Behavioral theory; Operations; Constraints; Inverted-U

1. Introduction

Powell and Johnson (1980) stressed that if workers have "one iota of discretion" regarding the performance of productive systems, their behaviors and the determinants of these behaviors must be incorporated in the development of meaningful research models. Citing criticisms made by previous management researchers (e.g. Cummings, 1977), they emphasized that sufficient behavioral considerations have been long overdue in the operations management literature. Notable efforts to address this limitation have been made since that time through studies of the links between operating decisions, worker behavior, and performance. Authors in the field of operations management continue to stress that additional effort on this front is needed (Amundson, 1998; Hayes and Hill, 2001; Rungtusanatham, 2001; Schultz et al., 2003).

A fundamental question remains, however, as to how additional research aimed at spanning operational and behavioral issues should be conducted. Specifically, when studying the linkages between elements common to both operations management and behavior research, such as resource constraints and operational performance, what are the appropriate factors, relationships and methods of analysis that should be used to provide meaningful and lasting contributions? This key research question on practice remains insufficiently addressed in part due to its necessarily complex and interdisciplinary nature. Answering it for researchers in operations management requires a willingness to consider

* Corresponding author. Tel.: +1 404 727 7138.
E-mail addresses: elliot_bendoly@bus.emory.edu (E. Bendoly), dhur@yonsei.ac.kr (D. Hur).
[1] Tel.: +82 2 2123 5487; fax: +82 2 364 7828.

0272-6963/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jom.2005.08.004

established research in alternate fields such as organizational behavior and psychology. It also requires that researchers pursuing studies that span these disciplines are aware of and take measures to avoid the difficulties in interpretation that past research on these topics has encountered.

The present note is designed to outline some of the interdisciplinary commonalities and methodological pitfalls encountered in studying the links between constraints and operational performance. Our objective is to contribute to research practice by describing these typically overlooked issues and ultimately suggesting tactics for addressing them. These issues range from appropriate model conceptualization to operationalization and analysis in empirical studies. Recommendations for appropriately confronting methodological concerns related to these issues are provided.

The rest of the paper presents the study. In Section 2, we begin by presenting a case example of constraints management at a large urban hospital to motivate a discussion of basic conceptual interpretations in the performance constraints domain. In discussion, we draw on specifications from operations management and behavioral research and briefly review notions of both the behavioral and mechanistic (non-behavioral) dynamics of performance constraints. By simultaneously considering established views of operational dynamics and behavioral theory, Section 3 then discusses the operationalization of performance constraints, particularly, workload and potential problems with current practice. A demonstration of the non-monotonic dynamics between workload and performance is illustrated through the use of data drawn from the hospital's ERP system. The final section concludes with the recommendations for future research.

2. Conceptual interpretations

Through their meta-analytic review, Villanova and Roman (1993) claim that the ability of research to inform practice in the management of constraints is "critical given management's interest in situational changes that can influence worker contributions". Equally critical is the ability of researchers in alternate fields to inform each other on this topic and hence facilitate the practicality of research findings and conclusions. Notably, the field of operations management also has a long tradition of identifying, "managing around" and seeking to eliminate constraints in the workplace, though this literature is seldom cited in behavioral studies of constraints. Regardless, the Theory of Constraints (TOC), based on the ideas popularized by the Goal (Goldratt and Cox, 1986), has had one of the greatest impacts on the management community both in terms of theoretical suggestions and practical application (Herroelen and Leus, 2001; Fry et al., 1992). According to TOC, a "constraint" is anything that limits a system's ability, or the ability of the individual workers within that system, to attain higher levels of performance. The management of constraints focuses on the identification, adjustments around, and modification of such constraints.

2.1. Managing constraints: a case

To emphasize the impact that managerial views of "constraints" can have on operating policies, and as an example of the role that worker behavior can have on how these policies impact operational performance, we consider the case of the materials management department at a large urban hospital. The first author spent 15 months collecting field data primarily through interviews with the performance audit team, post hoc interviews with the staff and clearance to internal (i.e. ERP archived) data and reports to verify the interview findings, wherever possible.

The main responsibilities of the materials management department involve the management of incoming and outgoing inventories of one-time use items (e.g. latex gloves and saline solution) as well as multi-use items such as surgical gowns for which cleaning was outsourced. The department is also responsible for equipment decontamination/sterilization, surgical kit assembly (i.e. the assembly of the variety of sterilized surgical tools required for individual patient cases), and the recovery of "lost" items due to either to emergent cannibalization of sterilized goods or non-returns after use. Due to complaints from the nursing and surgical staffs regarding poor levels of accuracy and consistency in the services provided, an internal audit of the performance of the materials department was conducted by the hospital more than a year prior to this paper. The internal audit team comprised of hospital administration staff that practiced TOC concepts and methods.

Operations managers for the materials department had argued that they were constantly faced with high levels of turnover and thus chronically low levels of experience among their staff regarding the procedures in place. Further investigation by the audit team discovered not only the low levels of consistency and accuracy previously reported, but also an environment that seemed to foster relatively high levels of stress. These self reported levels of stress were not only higher than average levels at other materials management

departments in the municipal area, but were also judged to be higher than at previous environments the current workers had been employed at. According to the audit team's report, individuals felt that they were faced with a chronic excess of work to accomplish in the limited time provided by their work schedules. Furthermore, the workers of the department reported this excess level of work served to limit any satisfaction they had with completing tasks, and ultimately limited their overall job satisfaction. Exit interview data emphasized the fact that individuals simply felt that too much was expected of them, that their work was never seen as ''good enough'' and that that the nursing staff used this sentiment to create an environment that was unappreciative and even occasionally ''hostile'' to the materials staff. The audit team concluded that this all contributed to relatively low levels of job satisfaction and work motivation, and ultimately their decisions to leave.

In a program designed to improve these conditions, and ultimately increase both the retention and the effective (i.e. ''high quality'') productivity of the department, the audit team in conjunction with the department's operation managers suggested a permanent increase in the materials department staff (from 31 to 34 FTEs), in order to reduce the workload per staff member. The argument made by the audit team was that the current staff level served as the primary ''binding'' constraint on quality (cf. Yano and Rachamadugu, 1991) and due to in no small part associated rework, the primary constraint on the overall productive capacity (total throughput) of the department. Since the proposed staff increase was relatively small, with respect to the current staffing level and the hospital's overall operating budget, the hospital's higher-level administration approved a first wave of materials staff hiring to meet the new steady-state level proposed. A second audit conducted 6 months after instituting the new staffing level showed sustained increases in measures of morale (e.g. job satisfaction and motivation), quality and productivity across the board (especially as measured among the original 31 FTEs for which no layoffs or specific changes in work content had occurred). However, certain tasks and staff assignments still seemed to be associated with chronically excessive workloads and the operations managers in the department pressed the audit team and the administration to authorize a second wave staff level increase, slightly smaller than the first (34 to 36 FTEs).

Six months after the second staff level increase, a third audit showed some less than desirable effects among the original 31 FTEs. Although the average level of quality did show signs of slight, but not statistically significant, increases over a variety of tasks, and although stress levels reported by the staff showed another significant decrease, measures of productivity showed noticeable declines, as did more general measures of morale and motivation. *Productivity was measured based on the amount of uninterrupted time spent from the initiation of tasks to their completion*. In monitoring work as in prior audits and assisted by data captured by work completion records in the hospital's ERP system, the audit team was careful to filter out any time the individual workers spent on other activities within this period, to help ensure the amount of time used in productivity measures was task-specific value-added time. Therefore, the productivity declines could not be attributed purely to idleness. No social conflict issues due to new hires were reported and task content again appeared unchanged among the original 31 FTEs. No significant increases in task-interruptions had been observed when compared to the previous audits, and the general skill levels of the average worker did not seem to been significantly lower than in previous reviews. In fact due to the lack of attrition following the first staff increase, skill levels had seemed to continue a gradual increase.

That is, the audit report showed that while accuracy and consistency did not see noticeable increases after the second staffing increase, with the total amount of work completed remaining steady, the average *pace* of work had dropped significantly along with other worker utilization measures. As noted by one of the audit team members, ''*It seemed like people had simply slowed down*''. Interviews with workers also revealed an overall decline in job satisfaction and motivation compared to the second audit. The unfortunate result of this second staff change and third audit was a resurgence in criticism towards the management of the materials department for ''losing control'' and not making effective use of their staff. In their defense, the managers cited their intention to promote the hospital's mantra of ''continuous improvement'' by reducing the negative impact of what had been assessed as the primary ''constraint'' to their operation, i.e. insufficient staffing. (And yes, these managers as well as much of the hospital administration were similarly versed in the concepts behind TOC.) Yet, while ''over-correction and idleness was always a possibility . . .'' anticipated by the managers in their staffing decisions, the marked slow-down in value-added work-speed was not.

Why did workers adjust their work pace? Why did job satisfaction and motivation decline? Besides the unsettling implications this case may have for practitioners, it also poses issues for researchers interested in studying the impacts of operational policies regarding

resource constraints. Questions such researchers need to ask include:

- Would they, and their research methods, appropriately have anticipated and captured similar effects in this setting?
- If not, would such an oversight reflect an *inadequacy in the conceptualization* of links between constraints and performance?
- And, even if appropriately anticipated and conceptualized, would the chosen research *model and analysis allow* such effects to be demonstrated?

Where these doubts exist, major gaps between research findings from such studies and real world implications may be unavoidable. It is interesting to note that in this case the audit team and management representatives were taking great care in accounting for issues often cited as critical by researchers (e.g. making distinctions between non-value and value-added time, inquiring on work and social content factors, basing longitudinal assessments on a constant element of the work force, etc.). Yet at the same time, they also maintained an assumption typical of research that can prove inappropriate in practice: those changes in an operating parameter (such as staff size) can result in at least consistently *directional* (e.g. positive) changes in performance. There are a number of theoretically supported arguments why such an assumption can be problematic. As we will discuss, these arguments call for a change in the methods used by both researchers and practitioners alike in assessing work scenarios and subsequently prescribing solutions for improvement. To initiate discussion, the next section reviews past literature on constraints in operations management and behavioral fields.

2.2. Traditional views on constraints

From an operations management perspective, the Law of Bottlenecks, which Schmenner and Swink (1998) attribute to TOC, states that system-wide productivity is improved by eliminating or better managing bottlenecks (i.e. elements of a system's value-added process which constrain the extent to which other sub-processes are productively utilized). The identification of bottleneck stages as rate-limiters is typically associated with a lack of sufficient work inputs or process resources given the time required to complete tasks at subsequent stages. In the terminology of normative analytical modeling, these critical limiting constraints are typically distinguished as ''binding'', such that although other constraints can be potentially meaningful, under the scenario investigated, only modifications to these binding constraints can effectively benefit overall system costs and performance. In contrast, the elimination of alternate non-binding constraints (non-bottleneck forming constraints) is typically viewed as an ineffective management.

The history of research undertaken by behavioral theorists has both parallels and distinctions from the interpretations of operations management researchers. Some of the earliest references to ''constraints'' in the behavioral literature come from the comprehensive work of Campbell et al. (1970). As a general point, these authors refer to ''situational variables'' as properties of task environments that interact with individual differences and other environmental factors to influence behavior and work performance. Although they use the term ''constraints'' to help drive arguments that apply specifically to work hindrance, the authors do not formalize a distinction between the terms ''situational constraints'' and ''situational variables''. In particular, throughout the discussion of the extended model of managerial effectiveness presented in the conclusion of their text, the terms ''situational constraints'', ''situational variables'', ''situational characteristics'', and ''situational effects'' are all used equivalently with reference to the same work-moderating phenomena (Campbell et al., 1970, pp. 475–476). Over time such discretion yielded to a growing belief by behavioral researchers that a distinction between ''situational constraints'' and other ''situational variables'' was worth formalizing. Peters and O'Connor (1980) allude to this in their early work on the implications of situational constraints:

> ''Persons who are both willing and able to accomplish a task successfully may be inhibited in or prevented from doing so by situational variables that are beyond their control .... To the extent to which situational constraints hinder utilization of ability, performance should be reduced.'' (p. 392).

Several implications can be drawn from this statement. One of the most direct is an implied monotonicity (i.e. consistent directionality such as ''positive'') regarding the impact of situational constraints on performance. Specifically, situational ''constraints'' are referred to as issues that limit or hinder, rather than aid, performance. Such monotonic specification mirrors the general view of previous scholars (e.g. Dachler and Mobley, 1973; Schneider, 1978). Behavioral research into situational constraints has in particular taken this view to help explain differences in performance among workers in alternate work

environments. Similar specification is typically not associated with the broader category of situational variables of which situational constraints are a subset.

In an attempt to further clarify the distinction between situational variables and constraints, O'Connor et al. (1982) formally define situational constraints as factors that ''inhibit or block individuals from achieving valued work goals''. Definitions used by later authors also share this conceptual focus (Phillips and Freedman, 1984). For example, researchers have more recently referred to situational constraints as ''characteristics of a work environment that interfere with employees' work performance'' (Mathieu et al., 1992; Tesluk and Mathieu, 1999). Still more explicitly, Kane (1993) contrasts situational constraints and situational ''facilitators'' such as incentive wages (Guzzo and Gannett, 1988), referring to each as diametric subset of situational variables in general. This sub-classification scheme is extremely appealing in that it emphasizes what situational constraints are specifically *not intended* to represent (Fig. 1).

Of course, the interpretation of even basic definitions and frameworks for delineating ''constraints'' from other situational factors may be less than effective if the terms used to describe these concepts are not consistently interpreted in research. For example, in order to apply the concept of ''constraints'' within empirical studies, researchers must ultimately confront what exactly they intend to mean by ''hindering performance''. In part this requires the assumption of certain processes by which hindering occurs. Two general schemes have already been posed in past research. One of the two suggests that situational constraints influence performance directly through mechanistic relationships such as traditional work-processing and queuing models (Kendar et al., 2001; Icmeli and Erenguc, 1996; Campbell and Pritchard, 1976), while the other assumes negative behavioral effects of constraints on motivation and the subsequent impact of motivation on performance (e.g. Lawler, 1973). Fig. 2 illustrates the traditional monotonic directionalities of effect-models associated with these interpretations. Typically these monotonic linkages are even more simply assumed to be linear.

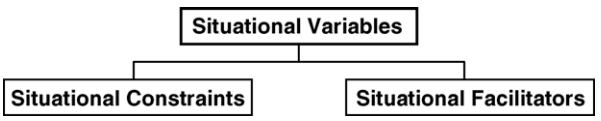

Fig. 1. Kane's categorization of situational variables.

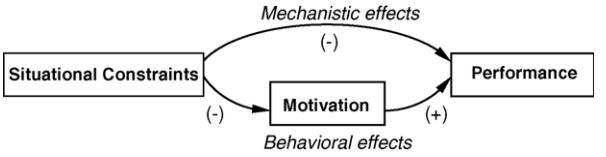

Fig. 2. Typical effect relationships considered in OB situational constraints research.

Specifically, in the presence of work-hindering constraints, researchers propose that employees may become frustrated with their inability to fulfill goals, particularly if they feel that their inherent abilities would allow them to more easily achieve established goals in less constrained environments (Peters and O'Connor, 1980). This frustration can lead to a breakdown in motivation and associated productivity by reducing the expectancy perceptions held by these workers (Peters et al., 1985; Vroom, 1964). The phenomenon behind this chain of effects is often referred to as part of the Expectancy Model of Goal Commitment, which basically states that commitment to a task will fall if workers expect that they will not be able to generate positive enough task-outcomes (Lawler and Suttle, 1973). The impact of constraints on affective criteria such as frustration and satisfaction, along with the effects on motivation, commitment, and work performance, has been supported by a number of studies (e.g. O'Connor et al., 1984; Steel and Mento, 1986).

While indirect effects suggest a range of behavioral reactions to the recognition of constraints by individuals, direct effects are viewed in large part to be mechanistic functions of the inherent capabilities of workers and the extent to which their work environment allows them to make use of those capabilities. As a simple example, if a worker is given 10 tasks to complete over an hour, but only given enough supplies or information to complete 9 of them, then the upper limit of tasks completed cannot be greater than 9, even if the individual would otherwise be able to complete all 10 tasks during that period. This effect is directly the result of a situational constraint, before any behavioral considerations are taken into play. In contrast, indirect behavioral effects suggest that this lack of resources might de-motivate the worker to the point at which the ''pace'' or ''rate'' of task completion is reduced. If this results in an effective work rate of less than nine tasks per hour, then this indirect phenomenon would be expected to dominate over a mechanistically imposed direct effect on work completed. In fact, such a dominating effect was thought, by the audit team, to be in place after the second wave of hiring in the hospital materials management case.

Ultimately the practical interpretation of work data may therefore rely on the ability to disentangle and distinguish such unique effects on performance variables. To that end, the selection of indicator measures used to sufficiently characterize ''constraints'' and distinguish their role in various contexts is critical in empirical studies on the topic. Equally critical in this process is the distinction between (a) factors that are thought to consistently fit expectations regarding the dynamics traditionally attributed to *constraints* and (b) factors whose dynamics do not fit or are expected to vary over the contexts examined in a particular research study. In the next section, specifically, we will have a closer look at one of the latter factors, workload.

3. Constraint analysis—the *workload* example

The appropriate selection of indicator metrics is seldom a simple or obvious task. The first essential question for empirical researchers facing this task is whether or not specific indicators appropriately match the definition of the concept of interest given the context of study (Malhotra and Grover, 1998; Kerlinger, 1986; Nunnally, 1978). For studies intent on discussing ''constraints'' as higher level constructs, as in most behavioral research on ''constraints'', part of this question involves whether individual indicators can be viewed consistently as hindrances, or if they should be instead viewed within a markedly different category (or categories) of situational variables. More to the point, although alternate situational variables might also impact performance either directly or indirectly, they may not capture the monotonic impact of situational constraints and thus may not allow for theoretical or practical interpretation of related results.

The selection issue is first and foremost one of face or content validity and applies regardless of whether these items are used formatively (i.e. as causal-indicators) or reflectively (i.e. as effect-indicators) (Bollen and Lennox, 1991; Hinkin, 1995; Hensley, 1999). Furthermore, if items selected do not represent the consistently monotonic effects on performance associated with existing conceptual definitions, these misspecifications can extend into deficiencies in structural and overall construct validity (Flynn et al., 1994; Ahire et al., 1996). As a result they may give rise to spurious analytical results and further interpretation problems (Schriesheim et al., 1993). This issue of the appropriateness of item selection is emphasized by the weak pattern of findings discussed in Villanova and Roman's (1993) meta-analytic review of the situational constraint literature. These authors speculate that the equivocality of findings across past literature is due to the variety of methodologies used and questionability regarding indicator selection, aggregation techniques, and general model constructions in specific contexts.

One of the potential sources of difficulty comes from the fact that many of the past scales intended to capture the concept of ''constraints'' have included *workload* indicators (O'Connor et al., 1984; Villanova, 1996). In both the behavioral and operations literature, workload is often depicted in terms of the number of tasks of a specific class (e.g. tools to sanitize, emergency calls to respond to, etc.) over a specific time frame or the number of standard work-hours assigned to a work group (Bertrand, 1983; Mangiameli and Krajewski, 1983; Mabert, 1985; De Vries, 1987). As a result, if employee performance is linked to time–work measures such as utilization, average completion rates, or daily output, as it commonly is (Schmenner, 1988), standard models of the form depicted in Fig. 2 face the undesirable challenge of clearly distinguishing ''effects'' from ''basic overlaps in formulaic definitions'' (e.g. the common use of ''amount of work'' measures in assessing both outcomes and antecedents used in models). When these definitional overlaps occur, correlations become essentially built-in and the observation of ''effects'' may be little more than tautological.

However, even if ''quantity of work'' (or ''standard hours-of-work'') is not simultaneously a fundamental definitional component of performance measures, there are other practical reasons for expecting workload decisions to demonstrate complex dynamics with respect to performance. With specific reference to the traditional definition of situational constraints stated earlier, these concerns focus on the appropriateness of monotonic effect assumptions that have become commonplace in empirical studies in this area. Arguments against of this ''traditional'' use of workload indicators, and other indicators indicative of workload, in capturing the constraining nature of work environments come from both direct mechanistic and indirect behavioral perspectives.

3.1. A closer look at theorized dynamics

3.1.1. Mechanistic considerations

Purely mechanistic, non-behavioral perspectives tend to focus on the notion that an *insufficient* quantity of work ''to be completed'' can limit the amount of work that ''can be completed''. However, workload items have almost always been treated with *reverse-directionality* in empirical research on performance

constraints (e.g. O'Connor et al., 1984; Klein and Kim, 1998). That is, rather than a deficit of work, an *excess* of work per unit time has traditionally been used as a measure of a performance constraint and aggregated equivalently (i.e. without a change in sign or direction) with other constraint indicators. A notable exception to this treatment is work of Adkins and Naumann (2001). Given the service context of their study (a call center), this is analogous to not having sufficient inputs. Yet, as alluded to in the materials management case, it is perfectly reasonable to expect that excess workloads even in this service environment would eventually lead to high stress work-constraining conditions—a point central to past behavioral interpretations.

However, even operations management research has suggested such reverse directionality in certain cases. Recent experience in practice and research into JIT systems has demonstrated that overly excessive build-ups of work can lead to reductions in work accomplished (i.e. productivity over time). Non-behavioral arguments have been provided to explain this result. One line of reasoning stems from the idea that excess work-in-process (WIP) inventories interfere with efficient material flow, mask the severity of other system constraints, impede standard checks and balances designed to correct for such constraints and thus limit productivity (Sakakibara et al., 1997). Regardless of whether this leads to behavioral reactions as well, this lack of visibility can derail both automated and manned activities. Specifically, the effectiveness of automated job prioritization may suffer from misleading WIP accounts if inappropriately scheduled jobs force restrictions in resource availability. Excess WIP also implies real physical visibility and work flexibility repercussions that negatively impact productivity since such WIP necessarily occupies work-space at some level. Lieberman and Demeester's findings (1999) may in fact be partially emblematic of such relationships between WIP reductions and productivity gains. While it may be argued that both WIP levels and productivity are dependent on over-riding process management decisions rather than on one another, the potential for additional direct causal relationships along these lines cannot be simply ignored.

Another argument from the operations management literature comes from experiences in the semi-conductor industry. Bertrand and Wortmann (1981) observed that certain process measurements can take place several steps after the sources of quality failure. With higher workloads, the amount of time between failure and measurement increases. This results in a reduction in the likelihood that appropriate resources will be available for quality recovery efforts, and ultimately leads to lower productivity yields. Such an observation has prompted later theoretically driven suggestions for new workload-based order release policies (Bertrand, 1983; Bertrand and van Ooijen, 2002). Specifically, release of orders based on time-phased workload information reduces system congestion, resulting overall system performance. Once again, while the complete line of causality even in this observation may be debated, the issue of an inverse relationship of some form nevertheless remains. Again, regardless of whether behavioral reactions are also implied, the sheer reduction in timely visibility and immediacy in accounting give rise to negative results even in highly automated and regulated settings. Coupled with simultaneous propositions of positive mechanistic relationships with productivity, workload therefore poses serious consistency issues for those intending to model these effects as monotonic relationships.

3.1.2. Behavioral considerations

From an indirect behavioral perspective, cautions against the view of workload as an indicator of situational constraints are also very compelling. Specifically, while other items that have been used in operationalizing constraints are consistently viewed as monotonic with regards to any behavioral indirect effects on motivation, a considerable amount of the literature suggests that workload/time may have a non-monotonic inverted-U effect on this issue. Basically this is the result of two very well established phenomena. The first phenomenon is based on Parkinson's Law which states that "work expands so as to fill the time available for its completion" (Parkinson, 1958). Parkinson explains the "expansion" of work to mean that individuals facing with a greater amount of time to accomplish a fixed amount of work will tend to spread that same amount of work across that amount of time, without necessarily adding appreciable value in its completion. A similar individual with either a greater workload for that length of time, or a shorter length of time with which to accomplish that work, may be able to generate completed work of indistinguishable quality, though operating at a markedly greater work pace. Later authors have attempted to elaborate on the rationale behind this general effect, by attributing it to behavioral mechanisms.

Specifically these shifts in work pace have been tied theoretically to the concept of motivation. According to organizational goal-setting theory (Locke, 1968; Locke and Latham, 1990), motivation can be positively

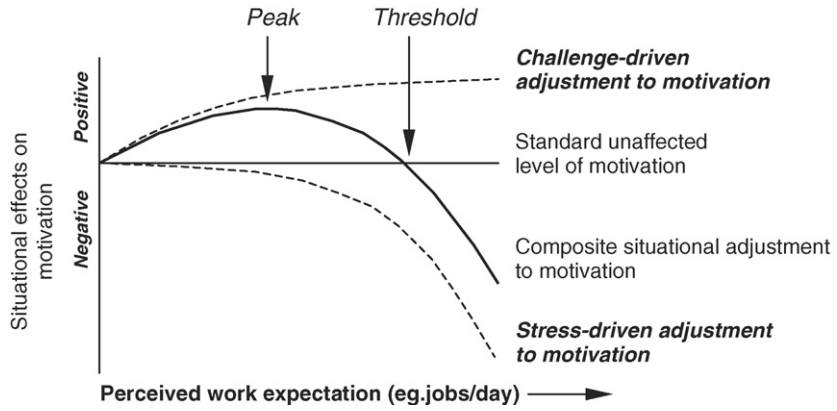

Fig. 3. Non-monotonic effects of workload.

influenced by the existence of challenges faced by workers. Such a phenomenon has been recently applied by Linderman et al. (2003) in their study of six-sigma practices. In specific support of the current discussion, an increase in workload can be seen as a challenge and thus a motivational force for work (Deci et al., 1989). This result captures the dynamics of both Parkinson's Law and goal-setting theory, and is likewise supported by the implications of parallel concepts such as the value of achievement (Hogan and Hogan, 1996; Ravlin and Meglino, 1987; Hollenbeck and Klein, 1987).

However, a large body of behavioral research suggests that motivation can also be negatively influenced by stress created by mismatches in the capabilities of a worker and the work environment (e.g. too much work or not enough work to illustrate a worker's true potential). Karasek (1979) suggests that high work demands may not be able to engender greater effort if workers are subject to environments that limit such actualisation. Instead, frustration, anxiety, de-motivation, and subsequently, the potential for lower than typical productivity may arise (Parker and Sprigg, 1999; Theorell and Karasek, 1996). Most recently, De Treville and Antonakis (2006) have also discussed such de-motivational and 'anti-productive' effects due to extreme operational policies with regards to lean manufacturing objectives. The tradeoffs between these two phenomena suggest that situational effects on productivity are multifold and perhaps emblematic of non-linearities such as inverted-U dynamics suggested by Villanova and Roman (1993) and more recently alluded to by Elsass (2000) and van Yperen and Hagedoorn (2003). Fig. 3 depicts this composite effect on outcome measurements such as motivation.

In this figure, the composite impact of challenge-driven effects on motivation suggested by goal-setting theory and stress-driven adjustments result in an inverted-U form with a characteristic motivational *peak* and a *threshold* beyond which negative effects on motivation dominate. Such a view is also synonymous with the classical Yerkes–Dodson Law, which first formally suggested an inverted-U relationship between arousal and performance (Benson and Allen, 1980; Yerkes and Dodson, 1908; Shaw and Weekley, 1985). This inverted-U dynamic has also more recently been corroborated in operations management studies that have recognized the role of behavior in leading to results contrary to purely mechanistic conclusions (van Ooijen and Bertrand, 2003; Schmenner, 1988). Hence a formal proposition regarding the non-monotonic nature of the impact of workload on performance, based on this combined effect, would deviate from the view of workload as traditional constraint indicator. Such a proposition is worthy of testing and could help explain some of the inconsistencies observed across studies in this area.

Proposition 1. *Specific factors traditionally interpreted as monotonic constraints, such as "workload", can provide greater predictive strength when their impact on performance is modeled as non-monotonic.*

3.2. Considerations for model design

As argued by authors such as Bobko (1985), more attention needs to be directed at modeling such non-monotonic relationships to distinguish them from relationships thought to exhibit simpler dynamics (e.g. performance impacts of equipment and other physical constraints studied by Peters et al., 1980, etc.). Unfortunately such considerations still remain under-represented in modern research. Again, risks to model validity due to such oversight may be particularly prevalent in "constraint" studies, which have relied

explicitly on indicators (such as workload) that are strongly theoretically associated with non-monotonical performance effects (e.g. O'Connor et al., 1984; Adkins and Naumann, 2001).

Moreover, given the present arguments, the anticipated effects of indicators relating to workload neither consistently fall into the category of "situational constraints" nor that of "situational facilitators", as conceptually distinguished by Kane (1993). Rather, depending on their relative levels, these situational variables transition between constraint and facilitator characterizations. In order to distinguish this dynamics, an alternate categorization that explicitly alludes to the transient non-monotonic nature of their effects would be helpful. Though there may be less cumbersome terms available to delineate this category, for the present we will suggest the term "bipolar factors" or "bipoles". Fig. 4 depicts an extension of Kane's categorization that takes the existence of these special variables into account.

Furthermore, the previous set of theoretical arguments would suggest that the performance effect-model of indicators categorized as bipolar factors, such as workload, would follow a significantly different conceptual structure from that of situational constraints. This structure is presented in Fig. 5 and implies special attention should be applied to non-linear and non-monotonic relationships if the analysis of such models is to yield interpretable and managerially practical results.

Proposition 2. *Across a range of contexts under study, the impacts of situational variables are more appropriately assessed when the theoretically or anecdotally supported potential for bipolar (i.e. u-curve) dynamics is formally considered for integration in model analysis.*

The importance of such a proposition is primarily in its ability to drive understanding and discussion of anticipated dynamics in specific operational settings. The proposition itself should be testable through basic examinations of the equality of standardized slopes for specific ranges covered by variables used in performance regressions. Ultimately, while the recasting of performance antecedent models posed in Fig. 5 suggests a clear deviation from traditional linear-effect assumptions, the assessment of models that include non-monotonic elements need not be complex. Regression

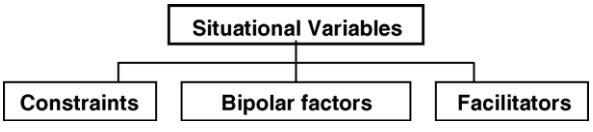

Fig. 4. Extended situational categorization to include bipolar factors.

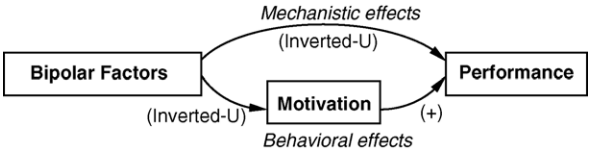

Fig. 5. Effect relationships associated with bipolar factors in contrast to those of situational constraints (Fig. 2).

and path analysis can still be readily applied by researchers, provided additional non-monotonic transforms of specific variables (e.g. a parabolic workload term, namely, workload^2) are included.

Yet model development and assessment is not the only analytical implication of the non-monotonic nature of bipolar factors. Another is the necessary caution that needs to be administered when constructing individual factors for use in model analysis. The need to distinguish between indicators, categorized as monotonic constraints and indicators that lead to fundamentally different effect-forms, is further emphasized by Spector and Jex (1998). In particular this distinction has lead them to propose a separate scale designed to measure situational constraints (e.g. their Organizational Constraints Scale) and alternate designed to measure workload (e.g. their Quantitative Workload Inventory). While not formally discussing the analytical issue of consistent monotonicity, they recognize that the two issues should remain separate as they fundamentally capture distinct effects on performance, which should be expected to take on alternate dynamics and should be viewed as subject to alternate moderating and mediating effects. In part, such specification emphasizes the critical nature of the scale development phase in empirical research.

3.3. Empirical check on case dynamics

To provide a quick check on whether an inverted-U relationship actually better describes the relationship between 'workload' and 'productivity' as described by the case, a week's worth of work records for each of the case phases (pre-staffing change, post-first wave, post-second wave) were extracted from the hospitals ERP system. Simple GLS regressions of productivity as a linear function of workload (both measured as described in the case) were conducted based on the combined data set of these three periods. K–S tests for normality showed no anomalies in either of these variable distributions at the $p < 0.20$ level. Dummy variables for each period were included in these regressions to account for any externalities that might not have been apparent in the case examinations. While

these dummy variables failed to provide additional accounts of variance in productivity, the hierarchical addition of a quadratic form of workload (i.e. workload^2) provided a significant increase in R^2 (from 0.172 to 0.244). As anticipated, the coefficient on this quadratic term was negative and highly significant, while the linear term remained positive and significant. This provided evidence that a bipolar depiction of workload effects (i.e. taking on the role as both a positive and negative driver depending on situational context) seemed to be strongly quantifiable in this general setting.

To assess whether motivation in fact played any key role in the relationship between workload and productivity, an additional hierarchical step was conducted by adding in the motivational measures gathered by the audit teams. These terms provided an additional significant increase in the model's R^2 (to 0.316) but eliminated the significance of the workload terms. A subsequent regression of motivation on workload showed that the linear and quadratic forms of workload accounted for approximately 35% of the variance observed in the staff motivation scores and described another inverted-U relationship in the signs of their coefficients. Taken together the evidence drawn from objective measures of workload and productivity, and the subjective accounts of motivation described a strong mediating role of motivation in this context akin to the lower half of the model presented in Fig. 5.

4. Conclusions

Melnyk and Handfield prefaced the 1998 *Journal of Operations Management* special issue on theory-driven empirical research by pointing out that while the desire for specific solutions to well-defined management problems still exists, there is a further need to gain a more thorough understanding of the dynamic environments in which these problems are set. This call for greater detail paralleled points made by earlier authors such as Powell and Johnson (1980) in their criticism of the lack of behavioral considerations in OM research. Empirical research provides one avenue by which such added knowledge can be developed, but only if sufficient care is taken to ensure interpretability of data collected, analysis conducted and results derived. With regards to gaining greater insights into the role of constraints in work environments, this care can be facilitated through the simultaneous consideration of alternate experiences and insights provided by operations and organizational behavior researchers over the last three decades. The present work has touched upon an area of concern to researchers in both disciplines and represents a starting point for future interdisciplinary work on the topic.

One conclusion derived from the literature of both disciplines is the recognition that the performance effect of certain factors, such as workload, should not generally be categorized as unidirectional in nature. Under varying conditions, such ''bipolar'' factors may exhibit either performance-limiting or performance-enhancing effects. This variety of effects can exist across a set of contexts or set of observations under study of even a single context. Unfortunately such a characterization is far from the norm perceived by most academics and practitioners. Traditional assumptions of strictly monotonic relationships between factors viewed as ''constraints'' and performance therefore continue to risk unanticipated gaps between the expected impacts of policy changes and those results actually realized, as exemplified by our case discussion in the prior section. Specifically, the perceptions of materials managers in this case, with regard to assumed strictly negative relationships between workload levels and individual value-added productivity, failed in predicting the performance impact of staff changes. Given appropriate considerations of motivating work challenges and subsequent impacts on individual performance, their enthusiasm regarding staff additions may have been curbed. Fortunately, in this single case the bottom line dollar impact of their subsequent staffing decision was not viewed as overwhelming in the grand scheme. Nevertheless, one wonders how many similar cases of insufficient behavioral considerations have lead to misplaced views of constraints, inappropriate workplace changes and subsequent muted or even counterproductive results that pose much greater strategic and financial difficulties for firms.

Ultimately, both theory and case experience suggest that reference to factors as operational ''constraints'' must be used with caution to prevent misguided assumptions in important management decision schemes. Moreover, in approaching policy developments and operational changes, factors that are thought to have non-monotonic effects on performance should be clearly distinguished from other factors whose effects on performance can be theoretically argued as unidirectional. The result of such a separate treatment can become apparent through the distinction of such factors-items (e.g. through factor aggregations separate from unidirectional items) or in the special treatment of bipolar factors in overall model structures (e.g. the inclusion of a parabolic term in path analysis).

Lastly, given the various arguments cited in the present research note, more attention to capturing and distinguishing direct mechanistic and indirect behavioral effects is called for in future comprehensive studies of performance constraints. The ability to account for both types of effects in work environments should draw from both operation management and behavioral theory and has the potential for greatly augmenting variance accountability in empirical studies. Only a handful of recent studies to date have attempted to account for both types of phenomenon and their respective dynamics through empirical or theoretical analysis (e.g. Schultz et al., 2003; Bertrand and van Ooijen, 2002; Nie, 2000; Villanova, 1996). Future empirical studies that capture objective measures and use such measures to predict performance results through established operational models in tandem with behavioral modeling efforts are encouraged.

References

Adkins, C.L., Naumann, S.E., 2001. Situational constraints on the achievement–performance relationship: a service sector study. Journal of Organizational Behavior 5 (22), 453–465.

Ahire, S.L., Golhar, D.Y., Waller, M.A., 1996. Development and validation of TQM implementation constructs. Decision Sciences 27 (1), 23–49.

Amundson, S.D., 1998. Relationships between theory-driven empirical research in operations management and other disciplines. Journal of Operations Management 16, 341–359.

Benson, H., Allen, R.L., 1980. How much stress is too much? Harvard Business Review September–October 86–92.

Bertrand, J.W.M., 1983. The use of workload information to control job lateness in controlled and uncontrolled release production systems. Journal of Operations Management 3 (2), 79–92.

Bertrand, J.W.M., van Ooijen, H.P.G., 2002. Workload based order release control and productivity: a missing link. Production Planning and Control 13 (7), 665–678.

Bertrand, J.W.M., Wortmann, J.C., 1981. Production Control and Information Systems for Component Manufacturing Shops. Elsevier, Amsterdam.

Bobko, P., 1985. Removing assumptions of bipolarity: towards variation and circularity. Academy of Management Review 10, 99–108.

Bollen, K.A., Lennox, R., 1991. Conventional wisdom on measurement: a structural equation perspective. Psychological Bulletin 110 (2), 305–314.

Campbell, J.P., Dunnette, M.D., Lawler, E.E., Weick, K.E., 1970. Managerial Behavior, Performance and Effectiveness. McGraw Hill, New York.

Campbell, J.P., Pritchard, R.D., 1976. Motivation theory in industrial and organizational psychology. In: Dunnette, M.D. (Ed.), Handbook of Industrial and Organizational Psychology. second ed. Consulting Psychologists Press, Palo Alto, CA, pp. 687–732.

Cummings, L.L., 1977. Needed research in production/operations management: a behavioral perspective. Academy of Management Review 2 (3), 500–504.

Dachler, H.P., Mobley, W.H., 1973. Construct validity of an instrumentality–expectancy–task–goal model of work motivation: some theoretical boundary conditions. Journal of Applied Psychology 58, 397–418.

Deci, E.L., Connell, J.P., Ryan, R.M., 1989. Self-determination in a work organization. Journal of Applied Psychology 74, 580–590.

De Treville, S., Antonakis, J., 2006. Could lean production job design be intrinsically motivating? Contextual, configurational, and levels-of-analysis issues. Journal of Operations Management, forthcoming.

De Vries, G., 1987. Nursing workload measurement as management information. European Journal of Operational Research 29, 199–208.

Elsass, P., 2000. Stress and the workload of professional women in Sweden. Academy of Management Executive 14 (2), 123–124.

Flynn, B.B., Schroeder, R.G., Sakakibara, S., 1994. A framework for quality management research and an associated instrument. Journal of Operations Management 11, 339–366.

Fry, T.D., Fox, J.F., Blackstone, J.H., 1992. An analysis and discussion of the optimized production technology software and its use. Production and Operations Management 1 (2), 229–243.

Goldratt, E., Cox, J., 1986. The Goal: An Ongoing Improvement Process. North River Press, New York.

Guzzo, R.A., Gannett, B.A., 1988. The nature of facilitators and inhibitors of effective task performance. In: Schoorman, D., Schneider, B. (Eds.), Facilitating Work Effectiveness. Lexington Books, Lexington, MA.

Hayes, J.M., Hill, A.V., 2001. A preliminary investigation of the relationships between employee motivation/vision, service learning, and perceived service quality. Journal of Operations Management 19 (3), 335–349.

Hensley, R.L., 1999. A review of operations management studies using scale development techniques. Journal of Operations Management 17, 343–358.

Herroelen, W., Leus, R., 2001. On the merits and pitfalls of critical chain scheduling. Journal of Operations Management 19, 559–577.

Hinkin, T.R., 1995. A review of scale development practices in the study of organizations. Journal of Management 21 (5), 967–988.

Hogan, J., Hogan, R., 1996. Motives, Values and Preferences Manual. Hogan Assessment Systems, Tulsa, OK.

Hollenbeck, J.R., Klein, H.J., 1987. Goal commitment and the goal setting process: problems, prospects and proposals for future research. Journal of Applied Psychology 72, 212–220.

Icmeli, O., Erenguc, S.S., 1996. The resource constrained time/cost tradeoff project scheduling problem with discounted cash flows. Journal of Operations Management 14 (3), 255–275.

Kane, K., 1993. Situational factors and performance: an overview. Human Resource Management Review 3 (2), 83–103.

Karasek, R.A., 1979. Job demands, job decision latitude and mental strain: implications for job redesign. Administrative Science Quarterly 24, 285–311.

Kendar, S.N., Wu, S.D., Sorer, R.H., Doshi, B.J., 2001. Melt scheduling to trade off material waste and shipping performance. Operations Research 49 (5), 629–645.

Kerlinger, F.N., 1986. Foundations of Behavioral Research, third ed. Holt, Rinehart and Winston, New York.

Klein, H.J., Kim, J.S., 1998. A field study of the influence of situational constraints, leader-member exchange, and goal com-

mitment on performance. Academy of Management Journal 41 (1), 88–95.
Lawler, E.E., 1973. Motivation in Work Organizations. Brooks/Cole, Monterey, CA.
Lawler, E.E., Suttle, J.L., 1973. Expectancy theory and job behavior. Organizational Behavior and Human Performance 9, 482–503.
Lieberman, M.B., Demeester, L., 1999. Inventory reduction and productivity growth: linkages in the Japanese automotive industry. Management Science 45 (4), 466–485.
Linderman, K., Schroeder, R.G., Zaheer, S., Choo, A.S., 2003. Six sigma: a goal-theoretic perspective. Journal of Operations Management 21 (2), 193–203.
Locke, E.A., 1968. Towards a theory of task motivation and incentives. Organizational Behavior and Human Performance 3, 157–189.
Locke, E.A., Latham, G.P., 1990. A Theory of Goal Setting and Task Performance. Prentice Hall, Englewood Cliffs, NJ.
Mabert, V.A., 1985. Short interval forecasting of emergency phone calls (911) work loads. Journal of Operations Management 5 (3), 259–271.
Malhotra, M.K., Grover, V., 1998. An assessment of survey research in POM: from constructs to theory. Journal of Operations Management 16, 407–425.
Mangiameli, P.M., Krajewski, L.J., 1983. The effects of workforce strategies on manufacturing operations. Journal of Operations Management 3 (4), 183–196.
Mathieu, J.E., Tannenbaum, S.I., Salas, E., 1992. Influences of individual and situational characteristics on measures of training effectiveness. Academy of Management Journal 35 (4), 828–847.
Melnyk, S.A., Handfield, R.B., 1998. May you live in interesting times . . . the emergence of theory-driven empirical research. Journal of Operations Management 16, 311–319.
Nie, W., 2000. Waiting: integrating social and psychological perspectives in operations management. Omega 28, 611–629.
Nunnally, J.C., 1978. Psychometric Theory, second ed. McGraw Hill, New York.
O'Connor, E.J., Peters, L.H., Pooyan, A., Weekley, J., Frank, B., Erenkrantz, B., 1984. Situational constraint effects on performance, affective reactions and turnover: a field replication and extension. Journal of Applied Psychology 69, 663–672.
O'Connor, E.J., Peters, L.H., Rudolf, C.J., Pooyan, A., 1982. Situational constraints and employee affective reactions: a partial field replication. Group and Organization Studies 7 (4), 418–428.
Parker, S., Sprigg, C.A., 1999. Minimizing strain and maximizing learning: the role of job demands, job control and proactive personality. Journal of Applied Psychology 84, 925–939.
Parkinson, C.N., 1958. Parkinson's Law: The Pursuit of Progress. John Murray, London.
Peters, L.H., O'Connor, E.J., 1980. Situational constraints and work outcomes: the influence of a frequently overlooked construct. Academy of Management Review 5, 391–397.
Peters, L.H., O'Connor, E.J., Eulberg, J.R., 1985. Situational constraints: Sources, consequences, and future considerations. In Ferris, G.R., Rowlands, K.M. (Eds.), Research in Personnel and Human Resources Management. JAI Press, Greenwich, CT, pp. 79–114.
Peters, L.H., O'Connor, E.J., Rudolf, C.J., 1980. The behavioral and affective consequences of performance-relevant situational variables. Organizational Behavior and Human Performance 25, 79–96.
Phillips, J.S., Freedman, S.M., 1984. Situational performance constraints and task characteristics: their relationship to motivation and satisfaction. Journal of Management 10 (3), 321–331.
Powell, G.N., Johnson, G.A., 1980. An expectancy–equity model of productive system performance. Journal of Operations Management 1 (1), 47–56.
Ravlin, E.C., Meglino, B.M., 1987. Effects of values on perception and decision-making: a study of alternative work value measures. Journal of Applied Psychology 72, 666–673.
Rungtusanatham, M., 2001. Beyond improved quality: the motivational effects of statistical process control. Journal of Operations Management 19 (6), 653–673.
Sakakibara, S., Flynn, B.B., Schroeder, R.G., Morris, W.T., 1997. The impact of just-in-time manufacturing and its infrastructure on manufacturing performance. Management Science 43 (9), 1246–1257.
Schmenner, R.W., 1988. The merit of making things fast. Sloan Management Review 30 (1), 11–17.
Schmenner, R.W., Swink, M.L., 1998. On theory in operations management. Journal of Operations Management 17, 97–113.
Schneider, B., 1978. Person-situation selection: a review of some ability–situation interaction research. Personnel Psychology 31, 281–297.
Schriesheim, C.A., Powers, K.J., Scandura, T.A., Gardiner, C.C., Lankau, M.J., 1993. Improving construct measurement in management research: comments and a quantitative approach for assessing the theoretical content adequacy of paper-and-pencil survey-type instruments. Journal of Management 19 (2), 385–417.
Schultz, K.L., McClain, J.O., Thomas, L.J., 2003. Overcoming the dark side of worker flexibility. Journal of Operations Management 21, 81–92.
Shaw, J.B., Weekley, J.A., 1985. The effects of objective work-load variations of psychological strain and post-work-load performance. Journal of Management 11 (1), 87–98.
Spector, P.E., Jex, S.M., 1998. Development of four self-report measures of job stressors and strain: interpersonal conflict at work scale, organizational constraints scale, quantitative workload inventory and physical symptoms inventory. Journal of Occupational Health Psychology 3 (4), 356–367.
Steel, R.P., Mento, A.J., 1986. Impact of situational constraints on subjective and objective criteria of managerial job performance. Organizational Behavior and Human Decision Processes 37, 254–265.
Tesluk, P.E., Mathieu, J.E., 1999. Overcoming roadblocks to effectiveness: incorporating management of performance barriers into models of work group effectiveness. Journal of Applied Psychology 84 (2), 200–217.
Theorell, T., Karasek, R.A., 1996. Current issues relating to psychological job strain and cardiovascular disease research. Journal of Occupational Health Psychology 1, 9–26.
van Ooijen, H.P.G., Bertrand, J.W.M., 2003. The effects of a simple arrival rate control policy on throughput and work-in-process in production systems with workload dependent processing rates. International Journal of Production Economics 85, 61–68.
van Yperen, N.W., Hagedoorn, M., 2003. Do high job demands increase intrinsic motivation or fatigue or both? The role of job control and job social support. Academy of Management Journal 46 (3), 339–348.

Villanova, P., 1996. Predictive validity of situational constraints in general versus specific performance domains. Journal of Applied Psychology 81 (5), 532–547.

Villanova, P., Roman, M.A., 1993. A meta-analytic review of situational constraints and work-related outcomes: alternative approaches to conceptualization. Human Resource Management Review 3 (2), 147–175.

Vroom, V., 1964. Work and Motivation. Wiley, New York.

Yano, C.A., Rachamadugu, R., 1991. Sequencing to minimize work overload in assembly lines with product options. Management Science 37 (5), 572–586.

Yerkes, R.M., Dodson, J.D., 1908. The relation of strength of stimulus to rapidity of habit-formation. Journal of Comparative Neurology and Psychology 18, 459–482.