Work Flow Policy and Within-Worker and Between-Workers Variability in Performance

Kenneth Howard Doerr

Tali Freed

Terence R. Mitchell

Chester A. Schriesheim and Xiaohua (Tracy) Zhou

Work flow policies are shown to induce a change in average between-workers variability (worker heterogeneity) and within-worker variability in performance times. In a laboratory experiment, the authors measured the levels of worker heterogeneity and within-worker variability under an individual performance condition, a work sharing condition, and a fixed assignment condition. The work sharing policy increased the levels of worker heterogeneity and worker variability, whereas the fixed assignment policy decreased them. These effects, along with work flow policy main effects on mean performance times and variability are examined. This article represents an initial step in understanding effects that may be important in the selection of an operating policy, the ignorance of which may lead to costly misestimates of performance.

A *flow line* is a production line in which all work follows the same sequence of operations. Because of the popularity and efficiency of the flow line, it has been the subject of considerable research (for a review, see Gagnon & Ghosh 1991; Ghosh & Gagnon, 1989). However, differences in worker ability and variability have been virtually ignored in operational models of flow lines, in spite of a wealth of psychological research evidence to suggest that significant differences exist, even for simple manual tasks (Doerr & Arreola-Risa, 2000; Hunter, Schmidt, & Judiesch, 1990; Rothe, 1978; Schmidt, Hunter, Outerbridge, & Goff, 1988). Recently, however, a class of work sharing systems (WSSs) has been proposed (Bartholdi & Eisenstein, 1996; Zavadlav, McClain, & Thomas, 1996) that not only acknowledges worker differences but (as we explain later in this article) relies on them as well. In this article we examine the importance of between-workers variability (differences in the mean performance of individual workers) and within-worker variability (differences in the performance of a single worker over time) on two kinds of flow lines.

The issue of within-worker and between-workers variability is important for a variety of reasons. First, for those doing work in the design of flow lines, in this article we question one of the most common assumptions of research on this topic: that within-line variability is random (unbiased) noise. Typically, when investigating a design change, one assumes that the only variability of interest is due to the manipulation (between-line variability). Our results show that at least some types of design changes affect within-line variability as well. Second, this article represents a preliminary investigation into the magnitude with which flow line efficiency is impacted by within-worker and between-workers variability and the degree to which this impact is moderated by work flow policies. The existence of such variability, as noted above, is not seriously in question (though few articles have appeared that recognize it). It may be that the flow line design literature has ignored within-worker and between-workers variability because researchers believe it does not make any practical difference. However, this question has not been studied, and this article represents an initial step in the investigation of that issue. Finally, with our comparison of two work flow policies, this article contributes to improved managerial decision making by examining new factors (heterogeneity and variability) that appear to affect group performance under those policies. The potential implications of research into these factors could be that management should not attempt to select the “most efficient” policy (regardless of the employees) or even the “fastest” employee (regardless of the work flow policy) but should instead think of selecting the best policy for a particular group or the best worker for an existing group and policy. Of course, such prescriptions cannot be derived from a single empirical study, but our work represents an initial investigation into these factors. Although it may seem overly mechanistic to evaluate employees in terms of their individual variability, we point out that such a procedure would be unquestioned if applied to any other input to the production (or service) process: One of the points of quality management is the control of variability in methods and materials. It is therefore reasonable to suggest that workers, as a major source of variability (Doerr & Arreola-Risa, 2000), might usefully be selected and managed with some attention to variability as well.

In the next section, we examine two work flow policies in more detail, to be able to hypothesize, in the Within-Worker and Between-Workers Variability on Flow Lines section, the different impact they will have on individual and group performance and variability.

Work Flow Policies

In the context of a flow line, *work flow* describes the way work moves between workers on the line. We use the term *work flow policy* (WFP) to describe all of the methods management has available to control work flow. These control methods all affect the interactions between workers. The impact of these interactions on the subsequent use of skills and motivational variables has been largely ignored in the applied psychology and operations management literature.

We are testing a behavioral model in this article, and our analysis was conducted by examining observable behaviors. However, to understand and explain our results, we draw on relevant literature from the field of industrial and organizational psychology. For example, we draw on findings of motivational effects from studies of interdependence (Kiggundu, 1981; Thomas, 1957; Thompson, 1967), social loafing (Comer, 1995; Latane, Williams, & Harkins, 1979), social compensation (Plaks & Higgins, 2000; K. D. Williams & Karau, 1991), autonomy (Klein, 1989; Langfred, 1999), feedback (Matsui, Kakuyama, & Onglatco, 1987), and equity (Harder, 1992) to predict and explain our results.

We are interested in two different types of work flow policies. With a fixed assignment system (FAS), the workload assigned to each worker is a sequence of contiguous operations, fixed from batch to batch. (A batch is a set of tasks to perform or a set of products to process.) The workload is performed in a limited physical zone, or workstation; the workstations typically do not overlap, work cannot be preempted, and the coordination required between workers at adjacent workstations is highly constrained. Of course, in some cases, more than one worker may be assigned to a workstation, but for simplicity we assume one worker per workstation.

With a WSS, the workload assignment is allowed to change from batch to batch (Zavadlav et al., 1996). The change of workload requires one worker to preempt, or interrupt, the work of another. An upstream worker must communicate the status of work on the batch to the downstream worker, and the two workers must coordinate the handoff of any required tooling or material needed to work on the batch. Thus, WSS involves a type of worker interaction and structural interdependence that does not exist with an FAS. It has been shown that such systems can, under certain conditions, outperform equally balanced lines and achieve a steady state assignment in which faster workers are more heavily loaded in proportion to their average individual performance (Bartholdi & Eisenstein, 1996).

The FASs and WSSs can be related to the typologies of group tasks given by Steiner (1972). The FAS is a kind of conjunctive task in which the "group performance is determined by the least able member" (Steiner, 1972, p. 17), whereas the WSS most closely resembles an additive task in which the group performance "depends upon the sum of the individual efforts" (Steiner, 1972, p. 17). The FAS is more complicated than the conjunctive tasks studied by Steiner because our tasks are variable. The WSS is more complicated than the additive tasks studied by Steiner because our tasks are variable and our assignments, or matching, is dynamic. Steiner (1972) claimed that for conjunctive tasks "the ideal arrangement in cases of this kind is one that involves as much homogeneity as possible" (p. 112), whereas for additive (but not dynamic) tasks, he claimed that heterogeneity was "irrelevant to potential productivity" (p. 117). We show later in this article that the existence of variability complicates the first claim, whereas the existence of variability and dynamism makes the WSS so different from a typical additive task that the second claim does not apply.

WSSs have a set of boundary rules that tell the workers, from batch to batch, how to interact with the upstream and downstream workers and where the assignments should begin and end in each cycle. There are different rules that can be used. In the bucket brigade rules most commonly used (and the ones we used in our experiment), workers proceed toward the end of the line with their current batch until they are preempted or, in the case of the last worker on the line, until they finish the batch. If they catch up to the worker ahead of them, they must wait (i.e., they are blocked). Once the worker at the end of the line finishes the batch he or she is working on, he or she walks back to the adjacent upstream worker, preempts that person's work and then proceeds forward again with the new batch. Each worker in turn preempts the adjacent upstream worker except the worker at the beginning of the line, who begins a new batch.

Within-Worker and Between-Workers Variability on Flow Lines

In spite of the well-established existence of substantial differences between workers in performance (Doerr & Arreola-Risa, 2000; Doerr, Mitchell, Schriesheim, Freed, & Zhou, 2002; Dudley, 1968; Hunter et al., 1990; Knott & Sury, 1987) the FAS literature has typically ignored the impact of differences in worker ability. In 1989, a comprehensive literature review of over 150 articles (Ghosh & Gagnon, 1989) found no research that incorporated worker differences in ability, possibly because the focus was on establishing line balances in spite of such differences so that workers could more easily be interchanged or replaced (Parker & Wall, 1998).

In contrast, the dynamics of the workload assignment depend on differences in average individual performance with WSS; faster workers are required to do more work. Under WSS, workers should be ordered from slowest to fastest for maximum efficiency (Bartholdi & Eisenstein, 1996; Bartholdi, Eisenstein, & Foley, 2001). Thus, WSS both assumes and depends on differences in average individual performance in ways that an FAS does not. FAS models assume that workers on a line are identical and that any variability is endemic to the tasks rather than to the workers (Doerr, Klastorin, & Magazine, 2000). WSS models, however, assume that workers on the line are significantly different from each other, but that within-worker variability is so insignificant that it can be ignored.

Although there is a considerable body of evidence that substantial differences between workers exist in terms of performance (Dunnette, 1983; Rothe, 1978; Schmidt & Hunter, 1983), the stability of these differences is still an open question. It seems that worker performance is dynamic: Workers' performance relative to each other on a given performance criterion may change over time (the dynamic criterion problem; Austin & Villanova, 1992; Ployhart & Hakel, 1998). Moreover, even at a given point in time, substantial within-worker variability in performance times has been shown to exist (Doerr et al., 2002; Knott & Sury, 1987).

It can be shown that a line operating with a WSS policy should be more efficient (in terms of the time required to complete work) than a line operating with an FAS policy and balanced workloads (Doerr et al., 2002). However, the dominance of the WSS policy relies on the existence of stable (constant) worker differences in performance. Without stable worker differences, the performance of a WSS may degrade and become chaotic because upstream workers will become blocked by downstream workers in unpredictable times and places. Although there is anecdotal evidence to suggest that a WSS can perform well even under certain kinds of variability (Bartholdi et al., 2001), to our knowledge there has been no systematic experimental work done to test the impact of within-worker variability on the performance of a WSS.

Given the sort of dynamic relative performance predicted by studies of the criterion problem (Austin & Villanova, 1992; Hofmann, Jacobs, & Baratta, 1993; Hofmann, Jacobs, & Gerras, 1992; Ployhart & Hakel, 1998), an ordering of workers from slowest to fastest is problematic over time unless the ordering is reassessed at regular intervals. Moreover, given substantial within-worker variability (Doerr & Arreola-Risa, 2000), the order may be established only for average case performance: The ordering may not hold at any particular point in time, because a "slow" worker may be faster than a "fast" worker on any particular cycle. Thus, under such conditions a WSS line may be less efficient than an FAS line.

The clarity of the task assignment under a WSS is less clear, because the assignment changes slightly from cycle to cycle. A lack of clarity in task assignments is likely to lead to a lack of visibility and accountability for performance and encourage social loafing (Latane et al., 1979). This assignment clarity is also related to the construct of dynamic task complexity defined by Wood (1986), in that assignment clarity involves changes over time in the acts required to accomplish a task. Wood, Locke, and Mento (1987) argued that this construct will interact with motivation to influence performance. This variability in assignment may create an increased cognitive load that detracts attention from the work rate or work quality (i.e., the workers have to think more about which tasks to do and consequently think less about how fast or how well they do each task). The coordination needed to accomplish each cycle's assignment is also less predictable under a WSS. Assignment clarity is thus an additional factor that should favor an FAS over a WSS.

Mathematical models of WSS performance ignore these behavioral impacts of assignment variability, and they assume constant worker differences and no within-worker variability (e.g., Bartholdi & Eisenstein, 1996). Thus, predictions that a WSS is more efficient than an FAS, drawn from these models, may not be realized when the policy is implemented with human workers (Doerr et al., 2002).

Hypothesis 1: The average group performance of WSS will be no better than performance of an FAS.

In Hypothesis 1, we predicted that the theoretical dominance of WSSs over FASs (Doerr et al., 2002) would not be observed empirically. At the same time, we proposed that these policies themselves would change the level of within-worker and between-workers variability observed on a flow line. We proposed that this change would be affected at least in part by a motivational response to the policies that differs depending on the relative performance of the individual.

The idea that the context of a task might produce a motivational response is not new (Mitchell, 1997; Mowday & Sutton, 1993). We have already mentioned the social loafing literature, which suggests that some workers will shirk in contexts in which their performance is less directly observed. Other contexts may encourage workers to slow down or speed up, depending on the workers' relative ability. For example, work on social compensation (Plaks & Higgins, 2000; K. D. Williams & Karau, 1991) suggests that when employees are engaged in meaningful work, faster employees will speed up if they are aware of their relative ability; the more important or meaningful a task, the greater the effect.

The idea that WFPs can produce a motivational response that depends on the relative performance of the employees is also not new. The Koehler effect is the tendency for heterogeneous groups to perform better than would be expected from their individual performances (Hertel, Kerr, & Messé, 2000). Hertel et al. (2000) found that the Koehler effect may occur on conjunctive tasks (similar to FASs) but not additive tasks (similar to WSSs).

In the sort of serial interdependence workers on a flow line experience (Thompson, 1967), each worker depends on the previous, upstream worker, but in the FAS policy we examined, all workers depend on the slowest, or bottleneck worker to set the pace—no one can work faster than the bottleneck pace in the long run. Whereas, in the WSS policy, the fastest worker sets the pace, and his or her pace determines when the line resets; thus, other workers are dependent on the fastest worker to determine their assignment boundaries on each cycle.

One of the specific motivational factors that is thought to come into play when one worker is depended upon by others is *felt responsibility*. Felt responsibility is a motivational force that grows out of expectations that one person should act to maximally facilitate and minimally hinder another (Thomas, 1957). To capture sources of felt responsibility, Kiggundu (1978, 1981, 1983) defined a variable called *initiated interdependence*, which measures the degree to which one employee feels that others rely upon him or her to accomplish their work. To the extent that initiated interdependence produces a sense of felt responsibility in an employee (because, for example, a downstream employee is waiting for him or her to pass along work), it should yield an improvement in performance.

Initiated interdependence describes only one half of a dyadic relationship. To describe the other half, Kiggundu (1978, 1981, 1983) defined a variable called *received interdependence*, which is felt by one employee when he or she depends upon another to accomplish his or her work. Kiggundu (1978, 1983) did not find the positive motivational impact for received interdependence that was found for initiated interdependence. In fact, to the extent that received interdependence is associated with reduced autonomy, it is likely to have a generally negative motivational impact (Klein, 1989). Depending on their relative performance and the WFP in place, employees on a flow line may experience either primarily initiated or primarily received interdependence.

Prior to our experiment, we believed the following would be true: On the static FAS line, the faster employees will experience the most interruption of work by a peer, through the *blocking* (waiting to pass work or move downstream) and *starving* (waiting for work from upstream) caused by adjacent employees. Conversely, on a dynamic WSS line, slower employees will experience the most interruptions relative to the amount of work accomplished. Thus, the fastest employee on the static FAS line and the slowest one on a WSS line are most likely to experience negative motivational states because of the control of their work pace by another employee. Because this is likely to be perceived as a loss of autonomy, it should be detrimental to performance (Klein, 1989; Langfred, 1999).

In comparison, the slowest employee on a static line will experience the most responsibility for others because he or she is the most frequent cause of starving or blocking another employee. This experience will be shared by the fastest employee on a WSS line because he or she controls the end of every cycle, and the whole line resets according to his or her pace. Consequently, these employees will be most likely to experience positive motivation because they have to provide work to others and maintain the work flow. Because this is likely to be perceived as increased pressure to perform, the effect will be a positive motivational impact (Kiggundu, 1983; Stewart & Barrick, 2000; Wong & Campion, 1991).

In particular, because the slowest worker on an FAS line will be the bottleneck, he or she should experience something like initiated interdependence and feel the greatest amount of pressure to increase his or her work rate. Schultz, Juran, and Boudreau (1998) observed that bottleneck workers increased their work rate when interstation buffers were reduced (thus increasing the likelihood that they would cause starvation and blocking of adjacent workers). However, the slowest worker is not a bottleneck on the WSS line; this worker should experience something like received interdependence. His or her work will merely be preempted by an adjacent worker as needed to keep that worker busy. Moreover, the constant preemption may be perceived as negative feedback by the slowest worker, who will not experience any autonomy or control over the work pace of the group.

> *Hypothesis 2A:* The slowest worker will perform more quickly on an FAS line than on a WSS line.

Because the fastest worker on the FAS line will experience the most idle time due to starvation and blocking, he or she is likely to experience less autonomy and control over work methods. Because his or her pace will be determined by the slowest worker, he or she should experience something like received interdependence. The attention of the fastest worker may be distracted more frequently because of increased idle time. Attention, and especially the focus and direction of attention, is considered to be a primary motivational process and has been shown to be related to the amount of effort exerted on a task (Mitchell, 1997). Thus, the fastest worker may be expected to slow down on an FAS. However, the fastest worker has much more autonomy and control on a WSS line. His or her work will never be preempted by an adjacent worker, and he or she will never be blocked or starved. The fact that other workers depend on the fastest worker to set their pace should cause them to experience something like initiated interdependence. The potential to control and maintain a steady work pace is likely to be a positive motivational force.

> *Hypothesis 2B:* The fastest worker on a flow line will perform faster on a WSS than on an FAS.

Hypotheses 1, 2A, and 2B deal with mean performance. Hypothesis 1 postulates an observable effect in relative mean line performance between WFPs based on considerations of within-worker and between-workers variability and assignment clarity. Hypotheses 2A and 2B postulate that work flow policies will induce an observable effect in mean individual performance, depending on the relative abilities of the employees.

Hypotheses 2A and 2B also imply something about the level of between-workers differences, or heterogeneity, one would expect to see on a line. Under an FAS, Hypotheses 2A and 2B predict that faster workers will slow down, whereas slower workers will speed up: a regression to the mean or a decrease in heterogeneity of performances. Under WSS, Hypotheses 2A and 2B predict that faster workers will speed up and slower workers will become even slower: an increase in heterogeneity. Thus, Hypothesis 2C is corollary to Hypotheses 2A and 2B.

> *Hypothesis 2C:* Between-workers variability in performance will be greater under a WSS than under an FAS.

Thus, Hypotheses 1 and 2 (A, B, and C) predict a multiphase relationship in which preexisting differences in worker ability and variability produce differences in the relative effectiveness of work flow polices, and work flow policies act to moderate this effect by reducing heterogeneity in FAS lines but increasing it on WSS lines.

We next turn to an examination of within-worker variability. The existence of substantial and systematic within-worker variability over time is well documented (Deadrick & Madigan, 1990; Hofmann et al., 1992; Rothe, 1978). Moreover, individuals appear to have systematic differences in within-worker variability that can be predicted in field settings (Deadrick & Gardner, 1997; Hofmann et al., 1993; Ployhart & Hakel, 1998). This within-worker variability has been measured not only in longitudinal studies of shifts of mean performance but also in examinations of systematic differences between workers in within-worker work-rate distributions over short time frames, for example, during the course of an experiment (Doerr & Arreola-Risa, 2000; Knott & Sury, 1987). Thus individual variability is important not only because it exists but rather because it is both systematic and different across individuals.

Although variability is recognized as an important topic in the design of production flow lines, within-worker variability has been virtually ignored in the operations management literature. The source of variability in task times is almost always assumed, at least implicitly, to be the tasks themselves and not the workers (Doerr & Arreola-Risa, 2000). Thus, models that deal with variability in the design of production flow lines (e.g., Bartholdi et al., 2001; Carraway, 1989) cannot be used to address questions about the impact that differences between workers in within-worker variability have on flow line performance, because they implicitly assume that workers are interchangeable and that particular workers do not affect the variability that (in those models) is inherent in the tasks themselves.

We proposed that within-worker variability will be affected by work flow policies, because different policies will produce more or less clarity and simplicity in the work flow. In discussing Hypothesis 1, we noted the ways that variability in work assignment may reduce group performance in WSS. Here, we note that variability in work assignment may also induce variability in individual performance.

Moreover, compared with an FAS, WSS involves more coordination between workers because they must preempt one another and communicate about the status of the work that they are passing along. WSS also potentially involves a greater range of activities and more physical movement along the line than an FAS. These factors will combine to affect the variability of individual work times because (apart from any difference in mean performance due, e.g., to the coordination, preemption, and movement time) they will create intermittent distractions in the work flow and require a dispersion of effort and attention. Hence,

> *Hypothesis 3:* Within-worker variability will be greater on a WSS line than on an FAS line.

In Hypothesis 2C we predicted changes in between-workers variability in performance across different policies. In Hypothesis 3 we predicted changes in within-worker variability in performance across different policies. In the next section, we describe an experiment conducted to investigate these three hypotheses.

Method

Participants consisted of 105 undergraduate students enrolled in an introductory operations management class at a private university in the southeastern United States. Participation in this experiment could be used to partly satisfy a research requirement for the course, but other activities could satisfy that requirement as well. No coercion was applied to garner student participation, and the investigators did not recruit from their own classes. Participation from classes in which students were recruited averaged 84%. Neither the age, gender, nor ethnicity of the participants was tracked, but participants were diverse in gender and ethnicity and represented a cross-section of the undergraduate population and were not noticeably different from the general student population.

The experiment consisted of a behavioral simulation of an order-picking operation. Totes (approximately 10 in. [25.4 cm] × 6 in. [15.24 cm] × 0.5 in. [1.27 cm]) of Halloween toys were arranged in five racks, with 9 totes to a rack (see Figure 1). Totes were placed three to a column on a table so that once a participant was standing in front of a column of 3 totes, every tote in the column was within easy arms reach for every participant. A rack thus consisted of three contiguous columns of 3 totes each. The face of each rack was 30 in. (76.20 cm) across. The first and fifth racks were separated from the second and fourth racks, respectively, by a distance of 20 in. (50.80 cm), whereas the second, third, and fourth racks were separated from each other by a distance of 10 in. (25.40 cm). Therefore, the total length of the line was 210 in. (533.40 cm) and an even division into three stations would have five columns (15 totes) distributed across 70 in. (177.80 cm).

Manipulations consisted of an individual-work condition (IWC), an FAS condition, and a WSS condition. During every condition, one investigator stood on the left side of the racks (the beginning of the line) and gave

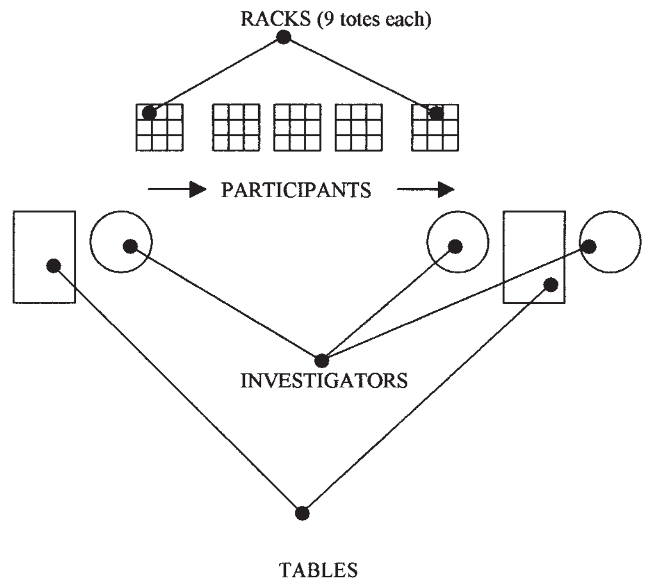

Figure 1. The order-picking racks.

participants an order form and a bag to be filled according to the order form. A second investigator stood on the right side of the racks and collected the order form and filled bag. A third investigator stood across a table from the second, and participants were told that this investigator was in charge of quality control and would be verifying their work. Participants were told to work as quickly as possible without making errors. During the experiment, participants filled a total (across all conditions) of 90 customer orders (30 in each condition), with each order consisting of requests for various items from each of the totes. The orders were balanced so that the same number of requests came from each column of 3 totes.

During the IWC, there were 10 trials in which participants were given a bag and a customer order listing 15 items to be placed in the bag. The items on the customer orders were printed according to their physical sequence on the line so that a participant could start filling the bag at one end of the line and proceed down the line to the end filling the bag along the way. When the participant reached the end of the line, he or she gave the filled bag to a second investigator and returned to the beginning of the line. To avoid a confound with learning effects, we used only the average times from the last three bags to compute individual work rates and variability. Although significant learning seemed to take place as the initial two to three bags were filled, a paired *t* test comparing the mean time each participant required to fill each of the last three bags ($M = 42.37$ s, $\sigma = 5.54$) with the mean time that same participant required to fill the previous bag ($M = 42.86$ s, $\sigma = 6.48$) revealed no significant difference, $t(104) = 1.12$, $p > .05$, indicating that any remaining learning was insignificant when participants filled the last three bags.

For the FAS and WSS conditions, participants were grouped randomly into teams of 3, and those teams were used for both conditions. In both conditions, participants were assigned to positions on the line in ascending order of the work rates calculated from the individual condition so that the slowest participant was at the beginning of the line and the fastest participant was at the end of the line. These relative work rates, assessed during the experiment with stopwatch observations, were later verified by the use of videotape observations. In no case was an inappropriate assignment made (e.g., in no case was a participant who performed more slowly during the individual condition inadvertently “promoted” to an advanced position on the line). Thirty orders were filled in both conditions, but only observations from 24 orders (5–28) were used to compute individual and group work rates. (The first 4 orders were discarded to control for possible learning effects, and the last 2 orders were discarded because the participants may have become aware that the condition was ending and thus worked at a different pace.) To test for order effects between the FAS and WSS conditions, we designed the experiment so that half of the participants went from an individual work to an FAS condition and then to a WSS condition, whereas the rest went from an IWC to a WSS condition and then an FAS condition. All dependent variables were tested for significant order effects using paired *t* tests, but no significant order effects were found, so we collapsed the cells for analysis.

During the FAS condition, participants were assigned columns of totes so that each participant had five columns of totes to pick from, distributed across an equal physical space of 70 in. (177.80 cm; see Figure 1). A customer order would be given to the first participant, who would pick five items and then hand the order along with the partially filled bag to the next participant. (For the first orders, the last 2 participants were given partially filled bags and told to continue the order as if the upstream worker had just handed the order to them.) Participants were not allowed to build interstation buffers but instead had to (a) wait (become blocked) if the downstream participant was not ready to receive an order and (b) wait (become starved) if an upstream participant was not ready to pass an order down.

During the WSS condition, 3 participants were again assigned to the line to fill customer orders. Participants were again arranged from slowest to fastest on the line. This time, however, participants followed the bucket brigade rules (outlined above) to fill the orders. Starting positions for the first orders were determined by computing the number of columns each participant should cover on average at a steady state, n_i, then positioning the participants from slowest to fastest so that n_i columns were between them. The number of columns was determined as follows:

$$n_i = \left[\frac{\alpha_i^{-1}}{\sum_i \alpha_i^{-1}} \times 15 \right]$$

where α_i is the velocity of the *i*th participant and [*a*] is the nearest integer to *a*. This procedure places participants in the starting position that is the closest to the steady state starting positions they would gravitate toward over a long period of time (Bartholdi & Eisenstein, 1996). These steady-state starting positions should be optimal, in the sense that they eliminate starving and blocking in a system without within-worker variability (Doerr et al., 2002). Proponents of a WSS would prescribe this as an initial starting position if it could be implemented in advance, because it would reduce the amount of inefficiency (relative to some other starting position) incurred in the transition state. In other words, we were starting the workers where they were supposed to wind up starting after a large number of cycles. Starting them at that point reduces potential inefficiencies that may occur in getting from an arbitrary starting point to this prescribed starting point. Thus, this prescribed point should be the starting position that most favors the WSS system. Because one of the points of our article is to question the efficiency of the WSS system, this is a conservative procedure.

Measures

Data were collected on videotape by placing a camera approximately 20 ft (about 6.09 m) in front of the simulated pick area. The camera was started before the first experimental condition began and allowed to run throughout data collection. We coded data coded from the videotape by using a stopwatch to measure performance times. Although these stopwatch measures of performance were subject to some potential coding error, test–retest reliability ($r = .90$) was deemed acceptable.

Group performance was measured by observing the time required for an order to travel from the beginning to the end of the line (flow time). Note that the group performance measure includes idle time due to starvation and blocking, but the individual performance measure does not. For the IWC, group performance was defined as the average flow time of the 3 participants in the group used for the WSS and FAS conditions. Group heterogeneity was measured by taking the standard deviation of the mean individual performance times for the group.

Individual performance was measured by observing the time an individual required to fill his or her portion of an order from the moment he or she received the order and bag (either from the investigator at the beginning of the line or from an upstream participant) until the moment the participant was ready to release the order and bag (either to the investigator at the end of the line or to a downstream participant). The performance time was converted to a time per pick by dividing the time used by the number of picks (15 in the individual condition, 5 in the FAS condition, and a variable number in the WSS condition). Thus, individual performance does not include idle time due to starvation, blocking, or coordination, nor is it directly impacted by the fact that a varying number of picks might be required under different policies.

Because we did not collect measures of ability apart from individual performance in the IWC, we wanted to further assess the reliability of this measure. Although it is one of the points of our hypotheses that individual performance will be affected by work flow policy, we also expected to see some stability in this measure across policies. To test this, we correlated individual performance means across policies. The correlation between individual performance means in the IWC and FAS individual performance means was moderate but significant ($r = .68$, $p < .01$). Likewise, the correlation between individual performance means in the IWC and WSS individual performance means was moderate but significant ($r = .54$, $p < .01$). Although not large, the difference between these two correlations was significant (Fisher's $z = 2.27$, $p = .01$)

A couple of observations are worth making about these correlation results. First, they establish the point that there is some stability in our individual performance measure but that individual performance is also determined by work flow policy. If one considers individual performance in the individual work condition as a surrogate for an ability measure, these correlations can also be seen as an estimate of the predictability of individual performance under a WSS or an FAS policy, from ability measures.

Second, because one of the justifications of the WSS policy is that it should allow workers to proceed at their own pace, it is somewhat counterintuitive that FAS individual performance means would correlate more strongly with individual performance means under the IWC. As explored further below, we believe this is mostly caused by a line position effect in the WSS policy. To check this, we recalculated the correlations controlling for line position effects, and the difference between the two policies disappeared (the IWC–FAS correlation dropped to .67 and the IWC–WSS correlation rose to .62). Thus, it appears that under both policies, individual performance is somewhat predictable from individual performance under the IWC, but the WSS policy has a line position effect that attenuates the relationship.

Finally, within-worker variability was measured with a coefficient of variation (CV) for each participant. This was computed by dividing the standard deviation of individual performance times by the mean individual performance time.

One might argue that faster workers would tend to be less variable ones and hence cause correlation between our two individual performance measures. The existence of such correlation would not be especially problematic, because we have no hypotheses that relate these two performance measures (i.e., we have no hypotheses about the relationship between heterogeneity and variability); however, this correlation is worth examining, both because it deals with the divergent validity of our individual performance measures and because the relationship is interesting in its own right. To test this, we correlated individual performance means with individual performance standard deviations under each of the three policies. As expected, faster workers were somewhat less variable under the IWC ($r = .43$, $p < .00$). Controlling for line position, this relationship is even stronger under the WSS ($r = .65$, $p < .00$) but weaker under the FAS policy ($r = .22$, $p = .02$), and the difference was significant (Fisher's $z = 5.57$, $p < .01$). This is not surprising, because (as we show below) the

WSS policy induces variability in slower workers, whereas the FAS policy induces variability in faster workers. The size of the difference in correlation between the FAS and WSS policy is somewhat surprising, however.

Analyses

To test for differences in group performance and heterogeneity by condition, we used a paired *t* test to compare the WSS and FAS conditions. We also conducted a one-way analysis of variance (ANOVA) to test for differences among the WSS condition, FAS condition, and IWC, with a Sheffé test to make post hoc pairwise comparisons between the three conditions. Note that our test of Hypothesis 1 requires that we accept that there is no difference between the means—that is, we need to accept the null hypothesis. The problems with doing this with the tools of traditional hypothesis testing are well known (Harcum, 1990; Malgady, 1998; Weitzman, 1984), and we will not add to this debate. Rather, we present our results, including an effect size and overlap (Cohen, 1988) and claim support for our hypothesis to the extent that the effect size is trivial and the distributions overlap.

To test for the interactive effects of line position and policy condition on individual performance times and variability, we used a general linear model (GLM; Cohen & Cohen, 1984). To test for differences between individual line positions and policies, we used a cell means test (Toothaker, 1993). This procedure involves a *t* test for significant difference between specific cell means of interest and is more informative than a simple test for interaction or a Sheffé test for differences within levels of a single factor. Critical values for the *t* tests were derived using a procedure developed by Cicchetti (1972).

Results

Results for group performance are shown in Table 1. As can be seen there, the flow times for the two policies are almost identical. A one-tailed *t* test, paired $t(34) = .68$, $p = .50$, shows a small, nonsignificant difference (effect size $d = .13$) and a large degree of overlap ($U = .098$, indicating that the two distributions overlap by over 90%; Cohen, 1988). This provides support for Hypothesis 1. The lack of a practical or significant difference is surprising in light of the normative models that predict that the WSS will be more efficient than the FAS (Doerr et al., 2002). We also conducted an ANOVA to test for differences among all three work conditions and found that the work flow policies did not improve flow time above the average flow time observed for the members of the group under the individual work condition, $F(2, 102) = .348$, $p = .707$, $\eta^2 = .063$ (Sheffé tests also indicated that there were no significant differences between any pair of the three policies.)

Results for individual performance are shown in Table 2. A GLM $F(8, 306) = 20.356$, $p < .00$, $\eta^2 = .418$, indicated significant policy effects, $F(2, 306) = 23.762$, $p < .00$, $\eta^2 = .280$, significant line position effects, $F(2, 306) = 44.917$, $p < .00$, $\eta^2 = .348$, and a significant interaction, $F(4, 306) = 6.372$, $p < .00$, $\eta^2 = .221$. As can be seen in Table 2, the slowest worker (at the beginning of the line) was faster under an FAS than under a WSS, and a cell means test indicated that the difference was significant at = .05, $t(3, 60) = 3.809 > 2.404$. This provides strong support for Hypothesis 2A. However, contrary to Hypothesis 2B, the faster worker (at the end of the line) was also faster under an FAS policy. Thus, Hypothesis 2B received no support. In fact, all workers were faster under the FAS policy. This may seem surprising given the results of the analysis of Hypothesis 1, but the explanation is quite simple: The WSS policy, as expected, reduced idle time associated with starvation and blocking. That is, the workers on the FAS line, while actually filling orders, were faster than the workers on the WSS line, but the groups' overall performance was no better because of the idle time on the FAS.

Results for group heterogeneity are shown in Table 3. Compared with the individual condition, worker performance was more homogenous under an FAS and more heterogeneous under a WSS. An ANOVA indicated that the differences among the three policies were significant, $F(2, 102) = 14.460$, $p < .00$, $\eta^2 = .344$, and a Sheffé test indicated that performance under the WSS policy was significantly more heterogeneous than performance under the FAS policy, $t = 5.275$, $p < .00$. Thus Hypothesis 2C received strong support.

Results for individual variability are shown in Table 4. Within-subject variability was higher under a WSS than under an FAS. It is interesting to note that the average CV under the individual policy (0.187) was higher than the average CV under an FAS (0.140) but lower than the average CV under WSS (0.296). A GLM, $F(8, 306) = 15.667$, $p < .00$, $\eta^2 = .388$) indicated that the differences between policies were significant, $F(2, 306) = 44.60$, $p < .00$, $\eta^2 = .349$, and a Sheffé test showed that the pairwise

Table 1
Flow Time

Policy	*M*	*SD*
WSS	44.09	4.81
FAS	44.86	6.58
IND	44.26	5.31

Note. Means and standard deviations are shown in seconds per order. Each policy group contained 35 participants. WSS = work sharing system; FAS = fixed assignment system; IND = individual work.

Table 2
Individual Task Times

	Line position					
	Beginning		Middle		End	
Policy	*M*	*SD*	*M*	*SD*	*M*	*SD*
WSS	3.31	0.55	3.52	0.71	2.71	0.38
FAS	2.91	0.38	2.83	0.41	2.57	0.32
IND	3.27	0.40	2.89	0.35	2.61	0.29

Note. Means and standard deviations are shown in seconds per pick. Each policy group contained 105 participants. WSS = work sharing system; FAS = fixed assignment system; IND = individual work.

Table 3
Standard Deviation of Each Group's Average Seconds per Pick

Policy	*SD*
WSS	0.552
FAS	0.257
IND	0.352

Note. Each policy group contained 35 participants. WSS = work sharing system; FAS = fixed assignment system; IND = individual work.

Table 4
Coefficients of Variation

Policy	Line position: Beginning	Middle	End
WSS	0.260	0.396	0.231
FAS	0.129	0.152	0.138
IND	0.196	0.190	0.175

Note. Coefficients of variation are shown in seconds per pick. Each policy group contained 35 participants. WSS = work sharing system; FAS = fixed assignment system; IND = individual work.

difference between the WSS and the FAS policies was also significant, $t = 9.21$, $p < .00$. These results provide limited support for Hypothesis 3. However, the GLM also showed that there was a significant main effect for line position, $F(2, 306) = 7.896$, $p < .00$, $\eta^2 = .184$, and a significant interaction between line position and policy, $F(4, 306) = 5.070$, $p < .00$, $\eta^2 = .203$. This significant interaction makes a simple interpretation of the main effect of work flow policy problematic, because it suggests that the relationship is not as straightforward as suggested by Hypothesis 3. Starting with the slowest worker, the CVs observed in the individual work condition were .196, .190, and .175. Thus, whereas there is a slight tendency for faster workers to also be less variable, the difference between participants was small. As already noted, the CVs observed for the line positions under the FAS policy were somewhat smaller (0.129, 0.152, and 0.138), but the middle-of-the-line position now shows the most variability, whereas the beginning-of-the-line position shows the least. Under the WSS policy, the CVs were all larger than those observed under the other two conditions (0.260, 0.396, and 0.233), but the end-of-the-line position exhibited the least variability. Thus, although it appears that the WSS policy does induce variability in individual work rates, the effect is much stronger for the middle-of-the-line position. Conversely, whereas it appears that the FAS policy reduces variability in individual work rates, the effect is somewhat weaker for the middle-of-the-line position.

Discussion

The primary implication of the analysis supporting Hypothesis 1 is that the relative efficacy of a WFP is significantly affected by behavioral factors. Theoretical and mathematical models show that WSS policy should dominate an FAS policy. However, these models assume stationary work rates from individuals that are unaffected by the WFP and continuous work flow unaffected by the need to coordinate the boundaries of work assignments. Our findings did not show that the WSS policy dominates the FAS policy, but the findings should not be interpreted as showing that the two are identical and certainly not that the FAS is better. Still, the failure to find the dominance predicted by the mathematical models is itself interesting as a counterexample to those models, and it demonstrates the need to include worker effects in models of WFP performance.

The result that the two policies may not be significantly different is practically important for at least two reasons. First, companies spend money to implement WSSs because they think it will improve productivity. Our findings suggest that this expense may be inappropriate. Second, WSSs will evidently increase between-workers differences in performance. The result of this will be either to increase differences in pay (if individual performance is a significant component of pay) or to increase inequity (if individual performance is not a significant component of pay). Either outcome is likely to have negative consequences (Cobb & Frey, 1996; Harder 1992; S. Williams, 1999).

The primary implication of the results supporting Hypothesis 2A is that line position effects are also significant determinants of individual and group work rates and that these effects are also dependent on the type of WFP in place. Although these line position effects are similar to the effects first investigated analytically by Hillier and Boling (1966), ours is the first article we know of that shows line position effects on individual performance.

Contrary to Hypothesis 2, we found that the effect of the WSS policy was not to increase the speed of the fastest worker and decrease the speed of the slowest but rather to decrease every worker's speed, most notably the middle worker. There are at least four potential explanations for the policy main effect that individual performance is decreased under a WSS but increased under an FAS. The first is an effect due to assignment clarity. The increased cognitive load associated with maintaining a dynamic boundary may reduce individual performance for all workers on the WSS line. Another explanation is rooted in the way that feedback on relative performance is received by the workers under the two policies. Feedback is generally motivational, and feedback on relative performance may be especially so (Matsui et al. 1987). Under the FAS it is obvious on every cycle who is fastest, because that person finishes before the adjacent worker and has to wait. This feedback is likely to create a motivational force that is not felt under the WSS, in which feedback on the relative speed is not as clear from cycle to cycle. Related to this is a social loafing (Latane et al., 1979) explanation: Workers under a WSS are slower because their individual contributions are less obvious. Still another possibility has to do with the perceived equity of the system. The rewards for participation were the same for all participants: satisfying a course requirement. But it may have become clear to the participants that faster work on the WSS line would translate into a heavier workload. Such perceived inequity is likely to be demotivational (Harder, 1992).

A possible explanation for the significant decrease in the performance of the middle worker under the WSS may be that the boundaries between workers exert a greater cognitive load under the WSS. For the workers at the beginning and end of a WSS line, there is only one boundary point between workers. However, for the worker in the middle, there are two boundary points in every cycle, the locations of which are beyond his or her control.

The analysis supporting the corollary Hypothesis 2C shows the impact these line position effects can have on group heterogeneity. Whereas the effects on the heterogeneity of work rates on the WSS and FAS lines were opposite (WSS increased heterogeneity, although FAS decreased it), the net effect on mean group performance is unclear. Whereas an increase in heterogeneity should be associated with less blocking on a WSS (because a slower worker is less likely to catch up to a faster worker when the difference between them is greater), and a decrease in heterogeneity should be associated with less starvation on an FAS (because a faster workers will wait less time for a slower worker when the difference between them is less), our analysis of flow times (see Table 1) did not demonstrate any significant reduction in flow time compared with the individual work policy. Further work is needed to examine the significance of any reduction in starvation and blocking due to the effect of WFP on heterogeneity and whether that reduction is more significant for one policy than for another.

The analysis supporting Hypothesis 3 that demonstrates that WFPs have an impact on the variability of individual work rates is significant for researchers who study flow lines, because it suggests that within-cell variance cannot be ignored when making between-cell comparisons of WFP performance. Although it was not hypothesized, our data also showed that the middle-of-the-line positions tend to exhibit more variability than the other line positions under the WSS, perhaps because of the dispersion of attention required to monitor both upstream and downstream assignment boundaries. Although some research has examined the allocation of given levels of individual variability to different line positions as a part of a WFP (e.g., Lau, 1992), to our knowledge, our research is the first to suggest that line position itself changes the magnitude of that variability.

The managerial implication of these changes in variability will likely depend on the magnitude of the underlying variability. There exist industrial settings in which CVs are considerably larger than those found in this laboratory study (Doerr & Arreola-Risa, 2000; Knott & Sury, 1987). The magnitude of the increase in individual variability associated with the WSS policy on this data set (a 33%–108% increase in CV, depending on the line position) suggests that further investigation should be undertaken to determine a threshold level of variability at which the statistically significant individual level effects described in this analysis begin to make a practical difference in group performance.

In spite of the dearth of operations management models incorporating within-worker and between-workers variability, firms seem to be aware of it. Automation, between-workers buffer inventories, and methods training can all be seen, in part, as (expensive) ways to ameliorate the "problem" of within-worker and between-workers variability, whereas a WSS can be seen as an attempt to exploit between-workers variability rather than to avoid it. Our results suggest that WSS may not be able to successfully exploit between-workers differences, because the policy itself induces within-worker variability.

Limitations, Extensions, and Summary

This article has examined within-worker and between-workers differences in performance, factors typically ignored in operations management models of production lines. There are at least three reasons why this examination is significant. First, the existence of within-worker variability that is affected by work flow policy suggests that statistical analysis of work flow policy interventions must account for within-cell variability. Second, we have shown that the effect of within-worker and between-workers variability on flow line efficiency is moderated by work flow policies. Although the existence of such variability was not seriously in question, few articles have appeared that recognize it, and very little empirical work demonstrates it. Third, this article represents an initial step in the investigation of the issue of the practical significance of the effect of within-worker and between-workers variability and demonstrates that the significance of the effect may depend on the work flow policy. Although we believe that the contribution of this article in each of these three areas is significant, the article has a number of limitations and needs to be extended in a number of ways.

A primary limitation is the difficultly in generalizing from student subjects. However, it is important to note that this limitation applies equally to both the WSS and the FAS systems we examined. Thus, because our main goal was to compare these two systems, we see no reason why the use of student subjects should have biased this comparison. In other words, we fully recognize that the motivation and behavior of full-time workers employed in production jobs will not be the same as that of students engaged in a behavioral simulation (however elaborate the simulation). However, this limitation does not directly impact the generalizability of our results, because our results deal with differences between the impact of two operating policies on motivation and performance. Although it may also be true that the differences between the impact of the policies would not be the same for students as for workers employed in production jobs, we see no strong argument as to why this should be so (or how it would bias the results). Nonetheless, it is clear that laboratory research of the type reported in this article needs to be cross-validated in the field, on different sorts of tasks, with longer flow lines, and with other workflow policies. We point to the need for such cross-validation as an obvious extension of the current work.

In this article, we have compared balanced (equal workload) FAS lines with WSS lines that naturally unbalance the workload so that faster workers receive more work. This is appropriate because, as we noted, the FAS literature has typically ignored worker differences in ability in the past. However, differences in average individual performance have recently been addressed in the FAS literature. Additional line-balancing algorithms have appeared that account for differences in average individual performance and variability (Doerr et al., 2000). These algorithms load faster workers more heavily than slower workers, in proportion to their average performance and variability. Such "imbalanced" lines outperform balanced lines when worker differences exist (Doerr et al. 2000). Clearly a comparison between WSS lines and these "imbalanced" FAS lines is another obvious extension of the present work.

We also reported line position effects, but because we did not manipulate line position by assigning one participant to multiple positions or by randomly assigning positions (so that relative speed and position interactions could be checked), our ability to make statements about these effects is limited. Moreover, we examined only three position lines, which limits our ability to make statements, for example, about generalized middle-position effects. We did not manipulate line position, because we were trying to avoid a confound with sequence and policy effects—as slow to fast is the prescribed sequence for WSS, our manipulation provides a clear comparison on that basis and allows us to examine only dynamic versus static work assignments. We feel that this was appropriate, because the focus of our study was not line position effects. However, interactions between line position and relative speed need to be examined in future work.

Whereas this article reports on observable behaviors (performance times), the underlying logic supporting the effects we observed rests, in part, on psychological variables that we did not measure (e.g., interdependence, assignment clarity variables).

Thus, although many of our hypotheses were supported, and those hypotheses are consonant with the psychological explanations we have discussed, it remains a limitation of this article that some underlying psychological mechanisms have not been directly investigated. In part, this is due to the lack of an appropriate measurement scale. Although scales have been developed for received and initiated interdependence (Kiggundu, 1978, 1981, 1983), and modifications of those scales have been used successfully in some research (e.g., Stewart & Barrick, 2000), the scales may not be appropriate for studies of sequential interdependence or in lab settings. Scale development is also needed for assignment clarity, which to our knowledge is a new construct, and for received and initiated interdependence. Such scale development is a clear direction for future work.

Although our article draws on some of the individual-difference literature for theoretical support (e.g., Ployhart & Hakel, 1998; Schmidt & Hunter, 1983), we did not measure any individual-difference psychological constructs (e.g., preference for group work; Shaw, Duffy, & Stark, 2000), nor did we measure differences in cognitive ability or attempt to correlate our individual work condition performance measure with any other individual-difference variable. We believe that the relationship between individual performance and measures of, for example, individual differences in cognitive ability are beyond the scope of the current article, in which we attempt to assess the interaction between individual performance and work flow policy on group performance. However, more investigation is clearly needed into the relationship between individual-difference variables and the factors examined in this article.

In addition, we held constant obvious external factors, such as pay, that may differ between policies and have motivational implications. Future work is needed to embed the phenomenon we study in a broader social–psychological context. In doing this, it is important to note that a cross-level structural model is needed. In the current research, some of the variables were measured at the individual level (individual performance and variability) and some at the group level (group performance and heterogeneity). Likewise, some of the psychologically related variables will need to be measured at the group level (e.g., cohesion, the Koehler effect) and some at the individual level (e.g., social compensation, assignment clarity, social loafing, and motivation), whereas some should be measured at multiple levels (e.g., equity and interdependence).

It is difficult to draw managerial implications from a single experiment, and this is a limitation of any empirical work. However, our results do suggest that firms should be cautious about expecting dramatic productivity improvements from a WSS implementation. The proposition that managers need to consider within-worker and between-workers variability when selecting an operating policy such as a WSS warrants further examination.

In summary, we found that contrary to mathematical models of human performance, work flow policies induce changes in individual work rates, relative work rates (performance heterogeneity), and individual variability. Although the impact of work flow policy on individual work rates has been shown before (Doerr, Mitchell, Klastorin, & Brown, 1996; Schultz et al., 1998), this has been in regards to the impact of changes in buffers. To our knowledge, ours is the first study to investigate the impact of changes in boundary (static or dynamic) rules and the first to establish the impact on heterogeneity and individual variability.

Our findings about the impact of WFP on individual differences in average performance and variability suggest that future research on flow line policies needs to consider who is on the line (average performance and variability), where they are on the line (line position effects), and the impact of WFP on the psychology and behavior of the group and the individual.

References

Austin, J. T., & Villanova, P. (1992). The criterion problem: 1917–1992. *Journal of Applied Psychology, 77,* 836–874.

Bartholdi, J. J., & Eisenstein, D. D. (1996). A production line that balances itself. *Operations Research, 44,* 21–34.

Bartholdi, J. J., Eisenstein, D. D., & Foley, R. D. (2001). Performance of bucket brigades when work is stochastic. *Operations Research, 49,* 710–719.

Carraway, R. (1989). A dynamic programming approach to stochastic assembly line balancing. *Management Science, 35,* 459–471.

Cicchetti, D. V. (1972). Extension of multiple-range tests to interaction tables in the analysis of variance: A rapid approximate solution. *Psychological Bulletin, 77,* 405–408.

Cobb, A. T., & Frey, F. M. (1996). The effects of leader fairness and pay outcomes on superior/subordinate relations. *Journal of Applied Social Psychology, 26,* 1401–1426.

Cohen, J. (1988). S*tatistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cohen, J., & Cohen, P. (1984). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Comer, D. R. (1995). A model of social loafing in real work groups. *Human Relations, 48,* 647–667.

Deadrick, D. L., & Gardner, D. G. (1997). Distributional ratings of performance levels and variability. *Group & Organization Management, 22,* 317–342.

Deadrick, D. L., & Madigan, R. M. (1990). Dynamic criteria revisited: A longitudinal study of performance stability and predictive validity. *Personnel Psychology, 43,* 717–744.

Doerr, K. H., & Arreola-Risa, A. (2000). A worker-based approach for modeling variability in task completion time. *IIE Transactions, 32,* 625–636.

Doerr, K. H., Klastorin, T. D., & Magazine, M. J. (2000). Synchronous unpaced flow lines with worker differences and overtime cost. *Management Science, 46,* 421–435.

Doerr, K. H., Mitchell, T. R., Klastorin, T. D., & Brown, K. A. (1996). Impact of material flow policies and goals on job outcomes. *Journal of Applied Psychology, 81,* 142–152.

Doerr, K. H., Mitchell, T. R., Schriesheim, C. A., Freed. T., & Zhou, T. (2002). Heterogeneity and variability in the context of flow lines. *Academy of Management Review, 27,* 594–607.

Dudley, N. A. (1968). *Work measurement: Some research studies*. London: Macmillan.

Dunnette, M. D. (1983). Aptitudes, abilities and skills. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 473–520). NY: Wiley-Interscience.

Gagnon, R., & Ghosh, S. (1991). Assembly line research: Historical roots, research life cycles and future directions. *OMEGA International Journal of Management Science, 19,* 381–399.

Ghosh, S., & Gagnon, J. R. (1989). A comprehensive literature review and analysis of the design, balancing and scheduling of assembly systems. *International Journal of Production Research, 27,* 637–670.

Harcum, E. R. (1990). Guidance from the literature for accepting a null hypothesis when its truth is expected. *The Journal of General Psychology, 117,* 325–329.

Harder, J. W. (1992). Play for play: Effects of inequity in a pay-for-performance context. *Administrative Science Quarterly, 37,* 321–335.

Hertel, G., Kerr, N. L., & L. A. Messé. (2000). Motivation gains in performance groups: Paradigmatic and theoretical developments on the Köhler effect. *Journal of Personality and Social Psychology, 79,* 580–601.

Hillier, F., & Boling, R. (1966). The effect of some design factors on the efficiency of production lines with variable operation times. *Journal of Industrial Engineering, 17,* 651–658.

Hofmann, D. A., Jacobs, R., & Baratta, J. E. (1993). Dynamic criteria and the measurement of change. *Journal of Applied Psychology, 78,* 194–204.

Hofmann, D. A., Jacobs, R., & Gerras, S. J. (1992). Mapping individual performance over time. *Journal of Applied Psychology, 77,* 185–195.

Hunter, J. E., Schmidt, F. L., & Judiesch, M. K. (1990). Individual differences in output variability as a function of job complexity. *Journal of Applied Psychology, 75,* 28–42.

Kiggundu, M. N. (1978). *The integration of task interdependence in the job characteristics theory of employee responses*. Unpublished doctoral dissertation, University of Toronto, Ontario, Canada.

Kiggundu, M. N. (1981). Task interdependence and the theory of job design. *Academy of Management Review, 6,* 499–508.

Kiggundu, M. N. (1983). Task interdependence and job design: Test of a theory. *Organizational Behavior and Human Performance, 31,* 145–172.

Klein, J. A. (1989, March–April). The human cost of manufacturing reform. *Harvard Business Review, 67,* 60–66.

Knott, K., & Sury, R. J. (1987). A study of work-time distributions on unpaced tasks. *IIE Transactions, 19,* 50–55.

Langfred, C. (1999). *Work group design and autonomy: A field study of the interaction between task interdependence and group autonomy*. Paper presented at the annual meeting of the Academy of Management national conference, Chicago.

Latane, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology, 37,* 822–832.

Lau, H. S. (1992). On balancing variances of station processing times in unpaced lines. *European Journal of Operational Research, 56,* 345–356.

Malgady, R. G. (1998). In praise of value judgments in null hypothesis testing . . . and of "accepting" the null hypothesis. *American Psychologist, 53,* 797–798.

Matsui, T., Kakuyama, T., & Onglatco, M. L. U. (1987). Effects of goals and feedback on performance in groups. *Journal of Applied Psychology, 72,* 407–415.

Mitchell, T. R. (1997). Matching motivational strategies with organizational contexts. *Research in Organizational Behavior, 19,* 57–149.

Mowday, R. T., & Sutton, R. I. (1993). Organizational behaviors linking individuals and groups to organizational contexts. *Annual Review of Psychology, 44,* 195–229.

Parker, S., & Wall, T. (1998). *Job and work design*. Thousand Oaks, CA: Sage.

Plaks, J. E., & Higgins, E. T. (2000). Pragmatic use of stereotyping in teamwork: Social loafing and social compensation as a function of inferred partner–situation fit. *Journal of Personality and Social Psychology, 79,* 962–974.

Ployhart, R. E., & Hakel, M. D. (1998). The substantive nature of performance variability: Predicting individual differences in intraindividual performance. *Personnel Psychology, 51,* 859–901.

Rothe, H. F. (1978). Output rates among industrial employees. *Journal of Applied Psychology, 63,* 40–46.

Schmidt, F. L., & Hunter, J. E. (1983). Individual differences in productivity: An empirical test of estimates derived from studies of selection procedure utility. *Journal of Applied Psychology, 68,* 407–414.

Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., & Goff, S. (1988). Joint relation of experience and ability with job performance: Test of three hypotheses. *Journal of Applied Psychology, 73,* 46–57.

Schultz, K. L., Juran, D. C., & Boudreau, J. W. (1998). Modeling and worker motivation in JIT production systems. *Management Science, 44,* 1595–1607.

Shaw, J. D., Duffy, M. K., & Stark, E. M. (2000). Interdependence and preference for group work: Main and congruence effects on the satisfaction and performance of group members. *Journal of Management, 26,* 259–279.

Steiner, I. D. (1972). *Group process and productivity*. New York: Academic Press.

Stewart, G. L., & Barrick, M. R. (2000). Team structure and performance: Assessing the mediating role of intrateam process and the moderating role of task type. *Academy of Management Journal, 43,* 135–148.

Thomas, E. (1957). Effects of facilitative role interdependence on group functioning. *Human Relations, 10,* 347–366.

Thompson, J. D. (1967). *Organizations in action*. NY: McGraw-Hill.

Toothaker, L. E. (1993). *Multiple comparison procedures*. Thousand Oaks, CA: Sage.

Weitzman, R. A. (1984). Seven treacherous pitfalls of statistics, illustrated. *Psychological Reports, 54,* 355–363.

Williams, K. D., & Karau, S. J. (1991). Social loafing and social compensation: The effects of expectations of co-worker performance. *Journal of Personality and Social Psychology, 61,* 570–581.

Williams, S. (1999). The effects of distributive and procedural justice on performance. *The Journal of Psychology, 133,* 183–193.

Wong, C.-S., & Campion, M. A. (1991). Development and test of a task level model of motivational job design. *Journal of Applied Psychology, 76,* 825–837.

Wood, R. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes, 37,* 60–82.

Wood, R. E., Locke, E. A., & Mento, A. J. (1987). Task complexity as a moderator of goal effects: A meta-analysis. *Journal of Applied Psychology, 72,* 416–440.

Zavadlav, E., McClain, J. O., & Thomas, J. L. (1996). Self-buffering, self-balancing, self-flushing production lines. *Management Science, 42,* 1151–1164.