SERVERS' EFFECTIVENESS IN SERVICE QUEUING SYSTEMS: A BEHAVIORAL OPERATIONS MANAGEMENT APPROACH

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Ahmad M. Ashkanani

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2017

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Benjamin Dunford, Chair

Krannert School of Management

Dr. David Schoorman

Krannert School of Management

Dr. Suresh Chand

Krannert School of Management

Dr. Kenneth Musselman

Regenstrief Center for Healthcare Engineering

Approved by:

Dr. Mark Bagnoli

Head of the Departmental Graduate Program

To my beloved.

ACKNOWLEDGMENTS

I would like to thank my academic advisor, Benjamin Dunford for his guidance over the past few years. I owe a great deal for his unconditional support, motivation, and belief in my ability to excel in academia. I would also like to give special thanks to my committee members, David Schoorman, Suresh Chand, and Kenneth Musselman for their valuable guidance and advice. Finally, I thank my family and friends for their support and encouragement. To my parents and grandmothers, thank you for keeping me in your prayers. To my wife and daughter, thank you for bringing joy to my life. To my brothers and sisters, thank you for believing in me. To my uncles, aunts, and cousins, thank you for your kindness. To my friends, thank you for making me laugh.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Ashkanani, Ahmad M. Ph.D., Purdue University, August 2017. Servers' Effectiveness in Service Queuing Systems: A Behavioral Operations Management Approach. Major Professor: Benjamin B. Dunford.

Previous research in the service queuing literature has treated servers' psychological factors as fixed effects, which limits our understanding of the behavioral drivers of service effectiveness. This dissertation presents a model of service worker productivity that examines the joint effects of system-level workload and individual-level motivation on productivity. Using multilevel models, I tested my hypotheses in a call center setting with a pooled queue structure and limited financial incentives. I found that workload and intrinsic motivation jointly influenced servers' productivity. In particular, intrinsically motivated servers were more productive and less prone to workload effects. In contrast, the productivity of less intrinsically motivated servers was lower and exhibited a U-shaped response to workload levels. Furthermore, I found that the intrinsic motivation effect on servers' productivity was more favorable than the extrinsic motivation effect in this setting. I discuss the implications of these findings and present three recommendations for building theories of service effectiveness that are more valid and useful to practitioners.

1. INTRODUCTION

Service work is one of the largest and fastest-growing sectors of the US economy (Hecker, 2004). Service organizations face the challenge of becoming more productive while employing fewer workers to reduce labor costs, which is a major component of operational expenses. In many service settings, a great emphasis is placed on adaptive task behaviors that involve workers' responses to unpredictable work demands (LePine, Colquitt, & Erez, 2000; Pulakos, Arad, Donovan, & Plamondon, 2000). Recent empirical evidence in the behavioral queuing literature suggests that servers adapt to changes in system workload by adjusting their service time (e.g., Kc & Terwiesch, 2009; Schultz, Juran, Boudreau, McClain, & Thomas, 1998; Tan & Netessine, 2014).

However, there has been little consensus about the shape or direction of the relationship between system workload and service time where evidence from the literature shows a negative linear (e.g., Kc & Terwiesch, 2009), inverted U-shaped (e.g., Tan & Netessine, 2014), or even N-shaped (e.g., Berry Jaeker & Tucker, 2017) relationships (Delasay, Ingolfsson, Kolfal, & Schultz, 2015). Moreover, while a growing body of evidence suggests that productivity varies among service workers (e.g., Grant, 2008; McCarthy et al., 2012), many of the studies in the behavioral queuing literature control for these differences as fixed effects (i.e., by including a dummy variable that captures average productivity differences among service providers), which limits our understanding of the underlying psychological factors that drive individual-level differences in performance. Another limitation in the extant behavioral queuing literature is the common assumption that servers react to system workload in a uniform fashion, which might not be the case as individual differences among servers might lead to behavioral differences in the way they respond to uncertain and dynamic work demands (Lazarus & Folkman, 1984).

In this dissertation, I integrate findings from the behavioral queuing literature with motivation theories to examine how system-level workload and individual-level motivation jointly influence servers' productivity in a pooled queuing setting with limited financial incentives. First, I hypothesize that workload influences within-server variations in productivity level where, on average, servers slowdown in response to higher workload levels, but only up to a certain workload threshold. Next, using multilevel models (Singer & Willett, 2003), I test whether differences exist between servers in terms of their productivity level (as measured by their service time as an inverse proxy of productivity), productivity dispersion (as measured by the within-day fluctuations in service time), and reaction to system workload. I then examine the role of intrinsic motivation as an influencer of between-server variation in both productivity level and dispersion. Next, I examine the joint effects of workload and intrinsic motivation and propose that (a) intrinsic motivation mitigates the "within-server" workload effect on servers' productivity level and (b) workload magnifies the "between-server" intrinsic motivation effect on servers' productivity level. Finally, I propose that the intrinsic motivation effect on servers' productivity is more favorable than the extrinsic motivation effect in a pooled queuing setting with an incentive structure that does not reward busy period performance. I test these hypotheses using longitudinal operational data from a large US call center.

This research design allows me to contribute to three distinct literatures: behavioral queuing theory, motivation, and the interface between operations management (OM) and organizational behavior and human resources management (OBHR). First, I show that servers' motivation influences interindividual variation in servers' productivity level, productivity dispersion, and reaction to system workload, challenging the common assumption that servers react to workload in a uniform fashion. Second, I show that system workload constitutes a boundary condition for the motivation effect where the positive intrinsic motivation effect is more evident under high levels of system workload. Third, I answer calls to integrate OM and OBHR theories in specific operational contexts to offer better insights into workers' behavior (Boudreau, Hopp, McClain, & Thomas, 2003). From a practical standpoint, this research offers insights to decision makers in making better staffing, scheduling, and work design decisions. Finally, I present three recommendations for building theories of service effectiveness that incorporate individual differences, team-level factors and mechanisms, and organizational withdrawal behavior in service effectiveness models.

2. THEORETICAL BACKGROUND AND HYPOTHESES DEVELOPMENT

My theoretical framework is based on the relationships between system workload, agents' motivational attitudes, and agents' productivity levels and fluctuations. Following recent trends in the behavioral queuing literature, I examine the role of system-level workload as a predictor of the intraindividual (i.e., within-server) variance in servers' productivity. Next, building on organizational behavior theories, I examine the interindividual (i.e., between-server) variation in servers' productivity and explore the role of intrinsic motivation as a predictor of such variation. I then examine the joint effects of intrinsic motivation and system workload in influencing servers' productivity at both the within- and between-server levels. Finally, I contrast the effects of intrinsic and extrinsic motivation on servers' productivity in the context of my research. Figures 2.1 and 2.2 summarize my hypotheses.

Note that it is important to specify the context in which the predictions are made. Thus, the following hypotheses predict servers' performance in a call center setting with a pooled queue structure, "next available agent" routing rule, high information visibility (i.e., servers are able to observe queue length, waiting time of first caller in the queue, and number of busy servers), low task complexity, focal task independence (i.e., each call is processed by a single server), and no performance-based incentives for busy period performance. I explain the research setting in more detail in later sections of the dissertation. I will use the terms server, worker, and agent interchangeably throughout the dissertation.

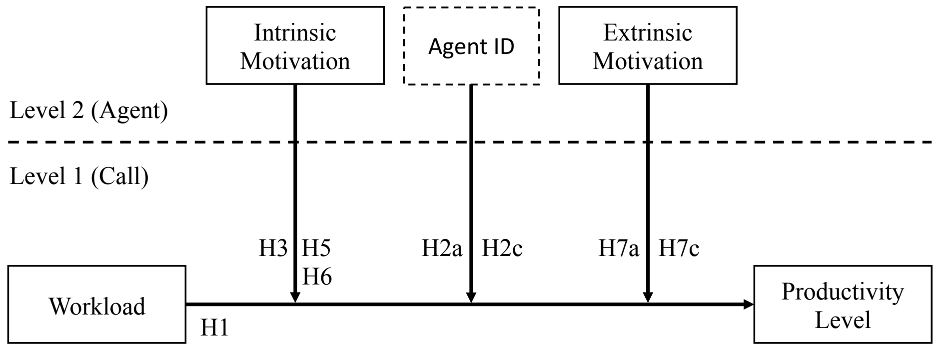

Figure 2.1. Productivity Level Model

2.1 Within-Server Variation in Productivity: The Workload Effect

A growing line of research in behavioral operations literature has challenged the classical assumption in queuing theory, which suggests that servers' productivity is exogenous to system workload levels (e.g., Batt & Terwiesch, 2016; Berry Jaeker & Tucker, 2017; KC, 2014; Kc & Terwiesch, 2009; KC & Terwiesch, 2012; Schultz et al., 1998; Staats & Gino, 2012; Tan & Netessine, 2014). This line of research suggests that servers strategically adapt their service rates in response to system workload levels. However, the results are mixed regarding the shape and direction of the workload-service time relationship (Delasay et al., 2015). For example, in two separate studies of healthcare workers, Kc & Terwiesch (2009) found that workers accelerated their work pace in response to higher workload levels. The results of this study, among others, suggest a negative linear relationship between workload and service time (e.g., Kc & Terwiesch, 2009; KC & Terwiesch, 2012; Staats & Gino, 2012). In contrast, in an empirical study of restaurant servers, Tan & Netessine (2014) found an inverted U-shaped relationship between workload and service time where at low workload levels, service time increased with the increase in workload, but after reaching a certain workload threshold, servers accelerated their pace with further increases in workload. The authors attribute this behavior to servers' strategic management of the speed-quality tradeoff where servers are encouraged to increase the service quality (as indicated by sales amount) at the expense of service time when the cost of waiting is lower, and vice versa. Another study by Berry Jaeker & Tucker (2017) suggests that in an inpatient hospital setting, the relationship between workload and service time is N-shaped. The authors found that service time increased with the increase in workload up to a first tipping point, after which patients with less severe ailments or injuries were discharged earlier to reduce congestion, resulting in lower service times. However, the authors also found a second tipping point, after which service time increased with the increase in workload due to the higher proportion of remaining patients with more urgent care needs.

While these results might seem contradictory at first, it is important to note that we need to consider the interplay between multiple factors that jointly affect workers' service time. These factors include, but are not limited to, the incentive structure (e.g., fixed pay vs. performance-based pay), queue structure (e.g., pooled vs. parallel queues), routing rule (round-robin vs. next available agent), information visibility (e.g., visible vs. blocked queue length), task complexity (simple vs. complex tasks), and task interdependence (e.g., independent vs. interdependent tasks). Indeed, recent studies suggest that queue structure and information visibility influence servers' productivity (e.g., Shunko, Niederhoff, & Rosokha, 2017; Song, Tucker, & Murrell, 2015). For example, in a field study examining physicians' productivity in a healthcare setting, Song et al. (2015) found that service time was 17% shorter in a parallel queuing system, compared to a pooled queuing system with a fixed-pay structure. The authors attributed this result to a higher sense of "ownership over patients and resources [in a parallel] queuing system" (Song et al., 2015, p. 3032). Similarly, an experimental study by Shunko et al. (2017) found that human servers were less productive in pooled compared to parallel queuing systems regardless of the incentive structure (i.e., fixed pay vs. performance-based) when information visibility was high. For the fixed-pay experimental condition, the authors attributed the slow-down behavior to servers' dispensability of effort, which was driven by higher levels of task interdependence as "servers work[ed] collectively to clear the customer queue"

(Shunko et al., 2017, p. 3). The authors also noted that the queue structure effect was more evident when system workload was higher. The results of these studies suggest that servers are more likely to slow down in a pooled than a parallel queuing system with full information visibility and a fixed-pay incentive structure. Moreover, this slowdown effect is more likely to be observed under high (rather than low) levels of workload (Shunko et al., 2017). This notion is consistent with the "opportunity" aspect of the COMU framework (Boudreau et al., 2003), which suggests that humans are influenced by the presence of situational constraints and opportunities that facilitate or hinder the achievement of their desired goals. For example, if servers wish to service fewer customers in a queuing system, then the presence of a pooled queuing structure (where workload is shared) and higher levels of workload (where idle time is minimal) increases the opportunity for those servers to engage is their desired behavior.

Thus, in my research setting, I expect the average servers to slowdown in response to higher levels of workload. Moreover, I expect the workload effect to be curvilinear such that service time increases with the increase in workload up to a tipping point, after which service time either decreases or saturates. The rationale of the curvilinear effect is as follows: First, the waiting costs become too high when the system is highly congested (Tan & Netessine, 2014), which creates additional incentives for servers to process calls faster to reduce customer waiting time. This notion is supported by analytical studies that suggest the costs of servers' extra effort can be offset by reductions in customers' waiting times (George & Harrison, 2001; Stidham & Weber, 1989). Indeed, interviews with call agents in my research setting suggest that customers become angry as they wait for extended periods of time, which in turn creates stress for the call agents (e.g., "[sometimes] I feel like we get verbally attacked because we are not available"). Second, although agents have the discretion to extend their service times without being noticed if those extensions are short, computer systems at modern call centers track all call details, and supervisors are thus likely to be alarmed if service time extensions are too long, which in turn could lead to unfavorable outcomes for the agents. Taken altogether, I expect the following:

Hypothesis 1 *The relationship between system-level workload and agents' individual-level productivity is curvilinear with a U-shape.*

In the above discussion, I examined the roles of system-level factors in influencing servers' reaction to workload. However, other factors that might impact servers' reaction to workload involve individual differences among service workers as I discuss below.

2.2 Between-Server Variation in Productivity: The Role of Individual Differences

OM scholars realize the existence of individual differences in productivity among service workers, leading many of them to control for these differences as fixed effects (e.g., Batt & Terwiesch, 2016; KC, 2014; KC, Staats, & Gino, 2013; Pendem, Staats, Green, & Gino, 2016; Song et al., 2015; Tan & Netessine, 2015). This approach allows scholars "to control for unobservable individual [server] effects that do not vary over time, such as level of motivation, innate ability, and practice routines. These are important to account for because they may significantly influence a [server's] productivity level in ways that cannot be measured" (Song et al., 2015, p. 3040). One limitation of the fixed effects approach is that it does not examine the underlying psychological forces that drive differences in servers' performance (i.e., why do we observe differences in productivity between servers?). In addition, a common assumption in the behavioral queuing literature stipulates that workers react to system workload in a uniform fashion, which might not always be the case, as I discuss below. Moreover, many of these studies focus on the productivity level of service workers, but little attention is paid to differences in productivity dispersion (i.e., within-day variation in productivity) among service workers. To address these concerns, I refer to organizational behavior theories that examine the role of individual differences in predicting worker performance.

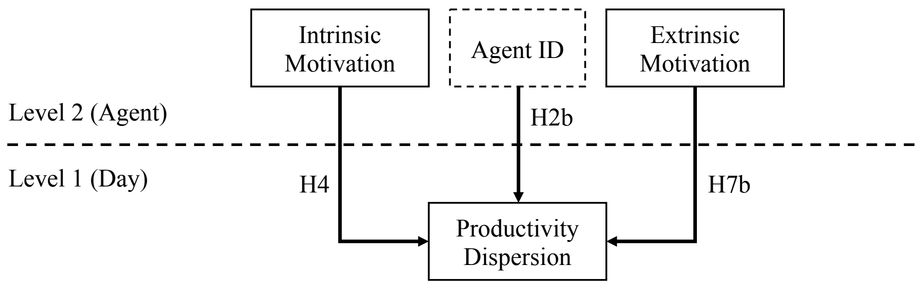

Figure 2.2. Productivity Dispersion Model

Organizational scholars suggest that there are two major categories of individual differences that affect workers' performance: ability-related (can do) and motivation-related (will do) factors (Cortina & Luchman, 2012). These individual-level factors lead to between-worker differences in task performance, contextual performance, and adaptive performance. Task performance involves behaviors needed to execute the technical core of products and services (Borman & Motowidlo, 1993; Campbell, 1990). In a call center setting, these behaviors include answering incoming calls, providing information to callers about offered services, operating communication systems, and relaying messages between customers and providers as necessary (O*NET, 2017). Individual differences in knowledge (e.g., information about offered products and/or services), skills (e.g., active listening), or abilities (e.g., oral expression) might lead to differences in task-related service time among call agents. Contextual performance involves voluntary behaviors that support or undermine the work environment, including citizenship or counterproductive work behaviors (Borman & Motowidlo, 1993). Examples of counterproductive work behaviors, which are defined as "employee behaviors that intentionally hinder organizational goal accomplishment" (Colquitt, LePine, & Wesson, 2013, p. 41), include social loafing and organizational withdrawal (e.g., working too slowly to shift workload to coworkers, taking too many breaks, etc.). Individual differences in motivation and/or organizational commitment, among other differences, might lead to variation in contextual performance (Colquitt et al., 2013). Finally, adaptive performance involves behaviors that characterize workers' responses to changes in dynamic work environments (Pulakos et al., 2000). In a call center setting, these behaviors involve handling work stress (e.g., dealing with angry callers) and dealing with uncertain work demands (e.g., sudden spikes in call volume). Individual differences in workers' attitudes might lead to differences in how they react to work stressors (Lazarus & Folkman, 1984). Indeed, some call agents in my research setting were more likely to skip work during busy days (e.g., Monday and Tuesday) and they attributed their withdrawal behavior to stress-related reasons (e.g., "Mondays are stressful"). In contrast, other call agents were less likely to call off during busy days and they attributed their behavior to motivation-related reasons (e.g., "The day goes by faster. [I] Like to stay busy.").

Taken altogether, I expect to observe interindividual variation in the productivity level (as indicated by the average productivity of servers), productivity dispersion (as indicated by the within-day fluctuations in servers' productivity), and reaction to system workload among service workers.

Hypothesis 2 *There exist statistically significant differences between agents in terms of their (a) productivity level, (b) productivity dispersion, and (c) reaction to system workload.*

Next, I examine the role of individual-level motivation as a predictor of the between-server differences in productivity.

Why Are Some Servers More Motivated Than Others?

Motivation is defined as "a set of energetic forces that originates both within and outside an employee, initiates work-related effort, and determines its direction, intensity, and persistence" (Colquitt et al., 2013, p. 164). The direction of effort refers to what employees do at a given time (e.g., speed-up vs. slowdown), the intensity of effort refers to how hard they try (e.g., increasing productivity by 5% vs.

20%), and the persistence of effort refers to how long they maintain their effort levels (e.g., increasing productivity for the next 5 vs. 30 minutes). Expectancy theory (Vroom, 1964), one of the earliest motivation theories, conceptualizes motivation as a multiplicative function of three constructs: expectancy, instrumentality, and valence. Expectancy is "a probability assessment reflecting an individual's belief that a given level of effort will lead to a given level of performance" (Mitchell & Daniels, 2003, p. 228). Instrumentality refers to the belief that a given performance level leads to obtaining certain outcomes (e.g., would poor (high) performance be punished (rewarded)?). Valence refers to the value a worker assigns to those outcomes (e.g., are the rewards valuable?). This view suggests that workers' motivational force is given by: $Motivation = Expectancy \times Instrumentality \times Valence$ (Vroom, 1964). If any of these forces is zero, then motivational force is zero. Applying expectancy theory to a call center setting suggests that workers are motivated when their effort leads to performance levels that help them acquire attractive (or avoid unattractive) outcomes.

However, which outcomes are of interest to workers in such setting? In the following discussion, I explore the answer to this question by considering two categories of outcome: intrinsic and extrinsic outcomes.

2.3 The Intrinsic Motivation Effect

Employees' intrinsic motives (e.g., task enjoyment, accomplishment, skill development, knowledge gain, lack of boredom; Colquitt et al., 2013) can be considered as the outcomes of interest in Vroom's (1964) expectancy theory. Individual differences in intrinsic motivation, which is defined as the desire to exert effort due to enjoying the task itself, have been linked to differences in workers' performances (e.g., Amabile, 1988; George, 2007; Grant & Berry, 2011; Grant, 2008). For example, Grant (2008) found a positive link between intrinsic motivation and the number of fundraising requests made by agents in a fundraising center, suggesting that intrinsic motivation positively influences both the direction and intensity of effort. These results, among others, suggest that in my research setting, agents with high compared to low intrinsic motivation levels are more motivated to exert effort that leads to higher productivity and thus a higher volume of serviced calls, an outcome that they value.

Hypothesis 3 *Intrinsic motivation is positively associated with the productivity level of call agents.*

Grant (2008) also found a positive link between intrinsic motivation and the persistence of firefighters (as measured by the number of overtime hours). This result is consistent with the principles of Conservation of Resources (COR; Hobfoll, 1989) theory. The COR principles suggest that workers are motivated to conserve their current resources, that they expend these resources to gain additional valuable resources, and that a lack of initial resources lead to defensive mechanisms to conserve the remaining resources (Halbesleben, 2010; Halbesleben & Bowler, 2007; Halbesleben, Harvey, & Bolino, 2009; Halbesleben & Wheeler, 2008, 2011; Ng & Feldman, 2012; Vinokur & Schul, 2002). If we consider workers' motivation to be a resource, then expectancy theory suggests that workers with lower compared to higher levels of intrinsic motivation have lower initial levels of resources (i.e., lower overall motivation levels). Hence, these individuals are more likely to defend their current resources when faced with the "opportunity" to do so (Boudreau et al., 2003). In contrast, workers with higher levels of motivation are more likely to expend their resources and persist in their effort as long as investments in effort lead to achieving the outcomes they value (e.g., processing more calls), which in turn leads to lower variation in their daily productivity level. Thus, I hypothesize the following:

Hypothesis 4 *Intrinsic motivation is negatively associated with the productivity dispersion of call agents.*

However, one question of interest relates to whether the intrinsic motivation-driven, between-server differences in productivity are evident at all times. Note that in the discussion above, I mentioned that servers with low intrinsic motivation levels are more likely to conserve their resources when faced with the "opportunity" to do so. This "opportunity" might be contingent on the system workload levels, as I discuss further below. Another question of interest relates to the prediction of Hypothesis 2c, which suggests that servers might react to their respective system workload levels in a non-uniform fashion. If so, then what explains the differences in servers' reaction to system workload? To answer these questions, I examine the interaction between intrinsic motivation and system workload in the section below.

2.4 The Joint Effects of Intrinsic Motivation and Workload

In the discussion above, I propose that the relationship between servers' productivity and system workload is curvilinear with a U-shape (Hypothesis 1). Moreover, I propose that servers react to system workload in a non-uniform fashion (Hypothesis 2c), which suggests that the workload effect in Hypothesis 1 might not apply to all servers. The differences in servers' reaction to workload might be influenced by the individual-level differences in intrinsic motivation. In Hypothesis 4, I propose that intrinsic motivation reduces variation in servers' productivity, which suggests that intrinsically motivated servers might be less influenced by the workload effect, since their productivity level remains stable throughout the day. This prediction is supported by the principles of the COR theory, in which workload triggers the defensive mechanism of servers with low intrinsic motivation levels, while the acquisition of new resources (i.e., processing more calls) encourages intrinsically motivated servers to stay productive throughout the day regardless of workload levels. Thus, I hypothesize the following:

Hypothesis 5 *Intrinsic motivation attenuates the within-agent workload effect on agents' productivity; the workload-driven, within-agent differences in productivity level are smaller for agents with higher intrinsic motivation levels.*

Hypothesis 3 predicts that intrinsic motivation is positively associated with servers' productivity levels, suggesting that differences in productivity levels exist between servers. However, Hypotheses 4 and 5 suggest that the productivity of workers with lower levels of intrinsic motivation fluctuates throughout the day, partly due to servers' responses to system workload levels. In contrast, the productivity of workers with higher intrinsic motivation levels is more stable throughout the day, regardless of workload levels. Taken together, I expect the between-server differences in servers' productivity level to be larger when the system is congested and vice versa. To understand the logic of this hypothesis, let us consider the case of a call center that employs 10 call agents. If one agent was busy servicing a call while the other agents were idle (i.e., no customers waiting in the queue), then extending the service time of the focal call is less likely to influence the total count of calls processed by the busy agent in the short run, since any new incoming calls will be routed to other idle agents. In contrast, if all call agents were busy servicing calls (i.e., customers are waiting in the shared queue), then extending the service time of a focal call is more likely to reduce the total count of service calls, thus providing more opportunity for individuals with low-intrinsic motivation levels to engage in social loafing behavior. Thus I hypothesize the following:

Hypothesis 6 *Workload magnifies the between-agent intrinsic motivation effect on agents' productivity; the intrinsic motivation-driven, between-agent differences in productivity level increase as system workload increases.*

Next, I compare the effects of extrinsic motivation on server productivity to the effects of intrinsic motivation in my research setting.

2.5 Intrinsic vs. Extrinsic Motivation

Employees' extrinsic motives (e.g., pay, promotion, praise, lack of disciplinary action; Colquitt et al., 2013) can be considered as another set of outcomes of interest in Vroom's (1964) expectancy theory. Individual-level differences in extrinsic motivation, which is defined as the desire to exert effort to attain external outcomes, have been linked to differences in worker performance (Amabile, 1988; George, 2007). Applying the principles of expectancy theory, I note that the instrumentality link between performance (e.g., servers' productivity) and extrinsic outcomes of interest (e.g., financial rewards) is weak in my research setting since the incentive structure does not reward higher performance levels during busy periods. This suggests that the motivational force of workers might not be driven by extrinsic motivation levels in this research setting. In contrast, as discussed earlier, intrinsic motivation has a positive influence on the motivation levels of workers. Thus, I expect the intrinsic motivation effect on servers' productivity to be more favorable than the extrinsic motivation effect in my research setting.

Hypothesis 7 *The effects of intrinsic motivation on call agents' (a) productivity level, (b) productivity dispersion, and (c) reaction to system workload are more favorable than the effects of extrinsic motivation in a pooled queuing setting with limited financial incentives.*

3. EMPIRICAL SETTING

My data comes from a US call center that is part of a large health system. This call center processes over 60,000 calls per month and employs around 80 call agents. The call center handles patient scheduling requests for different service lines within the healthcare system (e.g., pediatrics, neurology, etc.). The center operates Monday through Friday from 8am to 4:30pm. The center utilizes a single shift per day, so all agents are expected to work from 8am to 4:30pm. In this center, the call volume fluctuated both within and across days, as illustrated in Figures 3.1 and 3.2, respectively.

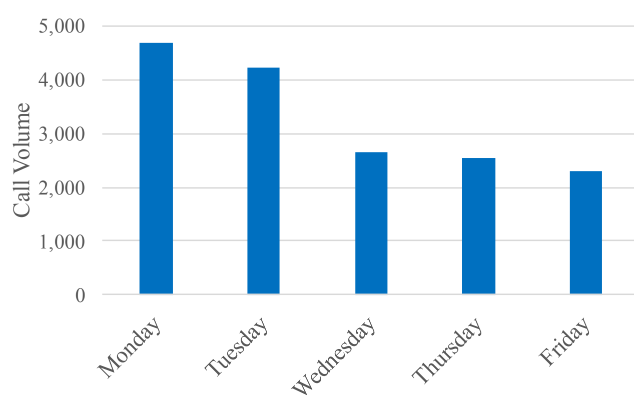

Figure 3.1. Example of Daily Call Volume Over a 1-Week Period

This call center has a standardized call process flow (Figure 3.3). Upon receiving the call, an automated computer system checks whether an agent assigned to assist with that specific service line is available. If all agents are busy, then the call is placed in a pooled first-come-first-served (FCFS) virtual queue with a "next available agent" routing rule. Once an agent becomes available, the computer system assigns the call to a specific agent who is responsible for processing the patient's request. The call agent has access to an electronic knowledge base that contains medical information

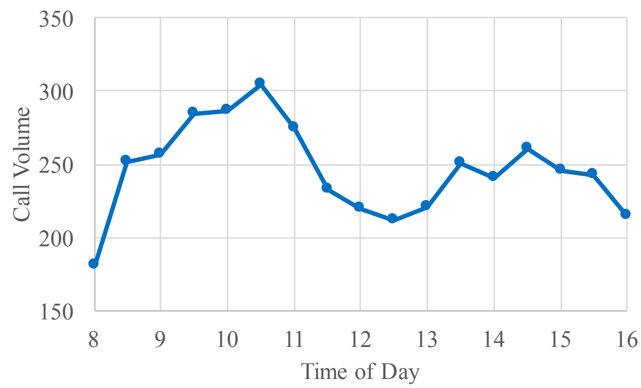

Figure 3.2. Example of Call Volume Fluctuations Over a 1-Day Period

pertaining to each service line. The call agent could also consult with other agents via an electronic chat system. However, this did not transfer the responsibility for processing the patient's request to other agents. Typical agent tasks involve taking patients' histories, checking the availability of healthcare providers, scheduling patient appointments with healthcare providers, and/or sending patient messages to healthcare providers and provider messages to patients.

Based on this process flow, the information system divides the patient request sojourn time into five non-overlapping time slices: queue time, ring time, talk time, hold time, and wrap-up time. Queue time indicates the amount of time a customer had to wait before being assigned to a call agent. Ring time denotes the time it took an agent to answer a call[1]. Talk time indicates the time an agent spent talking to a patient, while hold time indicates the amount of time a patient was put on hold. Finally, wrap-up time denotes the amount of post-call time spent by an agent to complete processing a service request. Note that an agent becomes available to receive new calls only after the request is fully processed.

To test my hypotheses, I collected data over 3 phases: archival call logs from November 2016, agent survey data from February 2017, and archival call logs from

[1]Call agents receive prompts on their screen indicating call assignments and need to click a button to answer the calls.

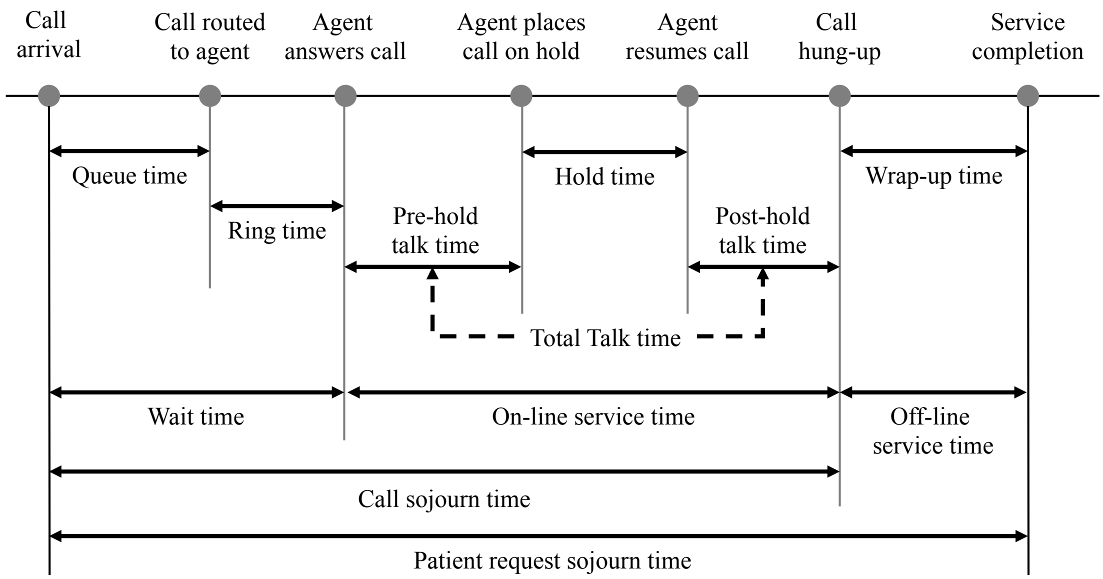

Figure 3.3. Standard Patient Request Flow in the Call Center

March 2017. In the first study, I tested Hypotheses 1 and 2 using the November 2016 data. In the second study, I combined the survey data with the archival call logs from March 2017 to replicate the findings of the first study and test the remaining hypotheses.

4. STUDY 1

In this study, I test Hypotheses 1 and 2 using data obtained from archival call logs.

4.1 Data

My sample includes all calls processed in the call center during November 2016, which came to a total of 67,505 calls processed by 82 agents. The information system at the call center tracks detailed information at each stage of the call. The data set includes the date of call, time of call, call service line, agent who handled the call, queue time, ring time, talk time, hold time, and wrap-up time. I used this information to construct my study variables as outlined in the section below. I dropped calls that had zero talk time (e.g., customer hung up before talking to an agent), were received outside normal work hours, or had an extreme call wrap-up time[1], leaving 66,221 calls and 82 agents. At the time of study 1, agents were paid using a fixed pay structure with no performance-based incentives. In the discussion that follows, i denotes the focal call, j denotes the agent, d denotes the day, s denotes the service line, and time period refers to the 30-minute time interval during which the focal call was received. For example, if a call was received at 9:32am, then the time period for this call is 9:30am-10:00am.

4.2 Measures

Tables 4.1 and 4.2 include summary statistics of Study 1 variables that were used in the productivity level and productivity dispersion models, respectively. Operationalization of these measures are discussed below.

[1] I considered wrap-up times that exceeded the sample 99th percentile of call wrap-up time to be outliers.

Table 4.1.
Means, Standard Deviations, and Correlations (Study 1: Productivity Level Model)

Variable	Mean	SD	1	2	3	4	5	6
1. Call Wrap-Up Time	61.23	109.55	—					
2. Workload	3.15	1.24	**0.07**	—				
3. Number of Agents	13.06	4.35	0.00	**-0.06**	—			
4. On-Line Service Time	245.16	212.51	**0.13**	**-0.02**	**0.04**	—		
5. Time	4.18	2.40	0.00	**-0.08**	**0.01**	**0.02**	—	
6. Overwork$_{K=4}$	0.17	1.15	**0.07**	**0.60**	**-0.11**	0.00	**0.27**	—

Note: Bold denotes significance at the 5% level.

Table 4.2.

Means, Standard Deviations, and Correlations (Study 1: Productivity Dispersion Model)

Variable	Mean	SD	1	2	3	4	5	6
1. Wrap-Up Dispersion	98.33	83.54	—					
2. Workload Level	2.80	1.14	**0.09**	—				
3. Workload Dispersion	0.61	0.23	**0.11**	**0.78**	—			
4. Number of Calls	46.43	21.72	**-0.20**	**0.63**	**0.45**	—		
5. Number of SLs	4.34	2.41	**0.07**	**0.72**	**0.59**	**0.45**	—	
6. Number of Servers	21.43	9.10	**0.07**	**0.41**	**0.27**	**0.30**	**0.45**	—

Note: Bold denotes significance at the 5% level. M= Mean. SD= Standard Deviation. SLs= Service Lines.

4.2.1 Dependent Variables

Call wrap-up time (WT_{ij}). Service time is a common "inverse" operational measure of servers' productivity that was used in multiple studies across different service settings (e.g., Kc & Terwiesch, 2009; Song et al., 2015; Tan & Netessine, 2014). I used call wrap-up time, which is defined as the post-call time spent by an agent to complete a patient's service request (see Figure 3.3), as a proxy for service time. I measured this dependent variable by calculating the number of seconds spent by an agent to complete a focal service request after the caller hung up the call. I used call wrap-up time rather than talk or hold time, since servers have more "opportunity" (Boudreau et al., 2003) to engage in social loafing behavior during the wrap-up stage of the service request. Unlike the on-line stage of service time, the agent could extend wrap-up time without interacting with the focal customer during or after the off-line stage, suggesting the agent is less likely to exert additional effort (e.g., talking more with the caller to extend the talk time portion of the call) or face mistreatment from the customer (e.g., customers might get angry if they were placed on hold for extended periods of time) if he/she decides to extend the service time. Another benefit of using the wrap-up time (rather than talk time) is that it reduces concerns about speed-quality tradeoffs (Kc & Terwiesch, 2009; Tan & Netessine, 2014) during the call wrap-up stage. Unlike the on-line stage, where spending less time talking to the caller might lead to lower service quality (e.g., missing medical history information), servers in this call center are expected to spend less time in the wrap-up stage.

Dispersion in daily wrap-up time (WD_{dj}). This measure estimates the variance in agents' daily wrap-up times. Higher dispersion values indicate agents' productivity fluctuates more within a day. I operationalized this measure as the standard deviation of daily wrap-up times for a given agent, which resulted in 1,452 agent-day observations of this dependent variable.

4.2.2 Independent Variables

Workload (WL_{ij}). Let S_{ij} denote a set of service lines assigned to agent j during the day call i was received, NC_{si} denote number of calls routed to service line s during the time period call i was received, and NA_{si} denote the number of agents assigned to service line s on the day call i was received. Then, I operationalize the workload seen by agent j during the time period call i was received as $WL_{ij} = \sum_{s\in S_{ij}} \frac{NC_{si}}{NA_{si}}$. Note that this workload measure adjusts for the number of coworkers by dividing the number of calls routed to an agent's virtual queues by the number of coworkers assigned to service those queues at a given day and time[2]. As an illustration of this measure, consider an agent servicing two service lines (i.e., two virtual queues). Assume the number of calls routed to the first (second) virtual queue during a given 30-minute time period was 10 (21) calls and the number of agents assigned to service that virtual queue was 5 (7) agents, respectively. Then, the workload seen by that agent in the given time period is 5 calls/agent.½-hour $\left(= \frac{10\,calls}{5\,agents.1/2-hour} + \frac{21\,calls}{7\,agents.1/2-hour} \right)$.

4.2.3 Controls

For the productivity level model, I controlled for variables expected to influence call wrap-up time: service line indicator (to account for different service line requirements that might influence wrap-up time; see Figure 4.1), $overwork_K$, which was measured by calculating the average workload level seen by an agent over the past K periods[3] (to account for extended workload and fatigue effects; Delasay et al., 2015; Kc & Terwiesch, 2009; Staats & Gino, 2012), day of week indicator (to account for day-specific effects such as start of week effect), on-line service time (to account for amount of time spent by an agent to process a focal call while the patient was on-line), and time of day (to account for time-related effects).

[2]These numbers were based on the full (rather than reduced) sample to measure workload levels more accurately.

[3]Following Kc & Terwiesch's (2009) operationalization, I chose $K = 4$ hours.

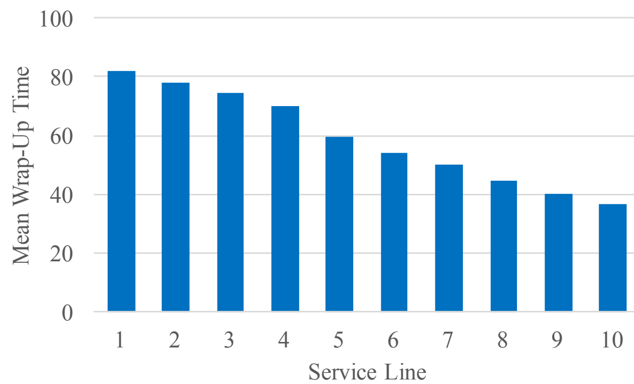

Figure 4.1. Example of Uncontrolled Mean Wrap-Up Times by Service Line (in Descending Order)

For the productivity dispersion model, I controlled for variables expected to relate to fluctuations in daily wrap-up time: day of week indicator (to account for day-specific fixed effects), total number of calls serviced by an agent in a specific day, total number of service lines serviced by a call agent in a given day, number of servers (i.e., coworkers) servicing the lines assigned to a focal agent on a given day, workload level, which was operationalized as the mean of system workload seen by an agent on a given day (to account for busier days), and workload dispersion, which was operationalized as the standard deviation of workload seen by an agent over a given day (to account for fluctuations in workload within a day).

4.3 Econometric Specification

I used multilevel modeling (Singer & Willett, 2003) to account for the nesting effects in my data: calls within agents (productivity level model) and days within agents (productivity dispersion model). Let WT_{ij} denote the wrap-up time of call i that was processed by agent j, X_{ij} denote a vector of control variables, and WL_{ij} denote the workload level seen by agent j during the time period call i was received. Then, the final productivity level model is given by Model 1 below:

Model 1: Productivity Level Model (Study 1)

Level 1 Model	$WT_{ij} = \beta_{0j} + X_{ij}B_{1j} + \beta_{2j}WL_{ij} + \beta_{3j}WL_{ij}^2 + r_{ij}$
Level 2 Model	$\beta_{0j} = \gamma_{00} + u_{0j}$ $B_{1j} = \Gamma_{10}$ $\beta_{2j} = \gamma_{20} + u_{2j}$ $\beta_{3j} = \gamma_{30}$
Assumptions	$r_{ij} \sim N(0, \sigma_r^2)$ and $\begin{bmatrix} u_{0j} \\ u_{2j} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right)$ for call i and agent j

Note that r_{ij} denotes the within-agent residual, u_{0j} denotes the between-agent random effect on the intercept of wrap-up time, and u_{2j} denotes the between-agent random effect on the linear slope of the workload effect[4]. A positive linear and negative quadratic coefficients of workload indicate support for Hypothesis 1. Furthermore, significant variance component estimates of u_{0j} and u_{2j} indicate support for Hypotheses 2a and 2c, respectively.

Next, let WD_{dj} denote the dispersion in wrap-time for calls received in day d and processed by agent j, while Ψ_{dj} denote a vector of control variables. Then, the final productivity dispersion model is given by Model 2 below:

[4]I did not add a random effect to the quadratic term of workload since preliminary tests of the data showed that this term did not vary at the between-agent level.

Model 2: Productivity Dispersion Model (Study 1)

Level 1 Model	$WD_{dj} = \pi_{0j} + \Psi_{dj}\Pi_{1j} + \epsilon_{dj}$
Level 2 Model	$\pi_{0j} = \alpha_{00} + \delta_{0j}$
	$\Pi_{1j} = A_{10}$
Assumptions	$\epsilon_{dj} \sim N(0, \sigma_{\epsilon}^2)$ and $\delta_{0j} \sim N(0, \widehat{\sigma}_0^2)$ for day d and agent j

Note that ϵ_{dj} denotes the within-agent residual while δ_{0j} denotes the between-agent random effect on the intercept of dispersion in daily wrap-up time. A significant variance component estimate of δ_{0j} indicates support for Hypothesis 2b.

I ran these models using SAS 9.4 PROC MIXED procedure with Maximum Likelihood (ML) estimation method. The results of these models are discussed below.

4.4 Results

Hypothesis 2 predicted that there are operationally meaningful differences in (a) productivity level, (b) productivity dispersion, and (c) reaction to system workload among call agents. To test Hypothesis 2a, I examined the existence of an agent's (level 2) random effect on the (level 1) intercept of call wrap-up time. As shown in column 1 of Table 4.3, the random effects are statistically significant at both the within- (9060.64, $p < .01$) and between-agent (3542.15, $p < .01$) levels, suggesting that call agents differ significantly in their productivity levels. The intraclass correlation coefficient (ICC) of the specification in column 1 was 0.28, suggesting that 28% of the variance in call wrap-up time lay between agents (Singer & Willett, 2003). To test Hypothesis 2b, I examined the existence of an agent's effect on the intercept of daily wrap-up time dispersion. The results in column 1 of Table 4.4 confirm the existence of an agent random effect, suggesting that the agent effect explains around 58% of the total variation in productivity dispersion. Finally, to test Hypothesis 2c, I examined the existence of an agent's random effect on the slope of workload. As shown in column 3 of Table 4.3, the slope of workload varied significantly across agents (66.39, $p < .01$). This suggests that agents reacted to system workload in a non-uniform fashion. Thus, Hypothesis 2 was fully supported.

Hypothesis 1 predicted that the relationship between system-level workload and agents' productivity is curvilinear with a U-shape. Note that since my dependent variable is call wrap-up time, a positive linear and a negative quadratic workload coefficients indicate an inverted U-shaped relationship between workload and service time, and therefore a U-shaped relationship with individual-level productivity. As shown in column 3 of Table 4.3, workload has an inverted U-shaped relationship with agents' wrap-up time as indicated by the positive linear (11.03, $p < 0.01$) and negative quadratic (-0.94, $p < 0.01$) coefficients of workload. However, it is important to note that around 98% of the observed workload levels in the sample lay below the inflection point of 5.87 calls/agent.½-hour (see Figure 4.2). This suggests that on average, agents slowdown in response to higher levels of workload, but only up to a certain threshold, which supports Hypothesis 1.

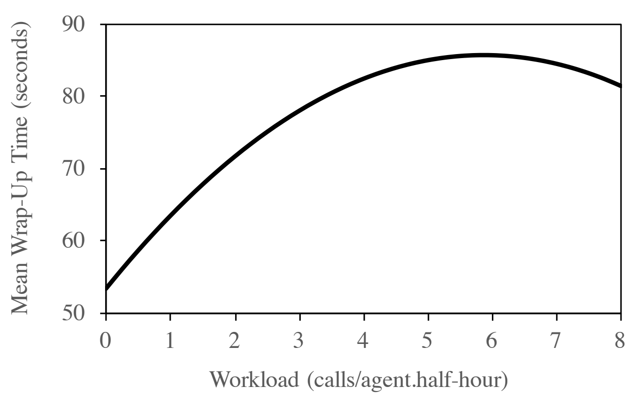

Figure 4.2. Call Wrap-Up Time as a Function of Workload (Study 1)

Collectively, these findings confirm that there were operationally meaningful differences in servers' productivity and reactions to system workload. Given that these

Table 4.3.
Workload Effect on Productivity Level (Study 1)

	Dependent variable: Call Wrap-Up Time		
	(1)	(2)	(3)
Fixed Effects			
Intercept	70.37***	40.76***	13.34
	(6.59)	(10.20)	(9.79)
Level-1 Main Effects (within-agent)			
Workload	—	—	11.03***
			(1.86)
Workload Square	—	—	-0.94***
			(0.24)
Random Effects (Variance Components)			
Residual (within-agent)	9060.64***	8739.71***	8671.52***
	(49.82)	(48.06)	(47.71)
Intercept (between-agent)	3542.15***	3597.58***	2357.56***
	(554.86)	(565.34)	(392.28)
Workload (between-agent)	—	—	66.39***
			(13.23)
Controls			
Service Line F.E.	—	Yes	Yes
Number of Agents	—	0.77***	0.76***
		(0.20)	(0.21)
On-Line Service Time	—	0.08***	0.08***
		(0.002)	(0.002)
Time	—	3.35***	4.78***
		(0.75)	(0.77)
Time Square	—	-0.41***	-0.53***
		(0.08)	(0.08)
Day of Week F.E.	—	Yes	Yes
$Overwork_{K=4}$	—	2.29***	-0.53
		(0.48)	(0.55)
-2 Log Likelihood (Deviance)	791775.7	789389.5	788989.3
Akaike Information Criterion	791781.7	789489.5	789095.3
ΔD		2386.2***	400.2***

Notes: N=66,221 calls/82 agents. Standard errors are in parenthesis. F.E.= fixed effects. ΔD= delta deviance. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 4.4.
Random Effects on Productivity Dispersion (Study 1)

	Dependent variable: Dispersion in Daily Wrap-Up Time	
	(1)	(2)
Fixed Effects		
Intercept	99.31***	50.62***
	(7.24)	(9.39)
Random Effects (Variance Components)		
Residual (within-agent)	2940.39***	2716.40***
	(112.33)	(103.83)
Intercept (between-agent)	4096.56 ***	3765.70***
	(669.64)	(621.13)
Controls		
Day of Week F.E.	—	Yes
Number of Calls	—	-0.73***
		(0.12)
Number of Service Lines	—	-4.46***
		(1.67)
Number of Servers	—	1.00***
		(0.27)
Workload Level	—	22.17***
		(3.63)
Workload Dispersion	—	14.6
		(10.32)
-2 Log Likelihood (Deviance)	15979.3	15863.8
Akaike Information Criterion	15985.3	15887.8
ΔD		115.5***

Notes: N=1,452 agent-day observations/82 agents. Standard errors are in parenthesis. F.E.= fixed effects. M= mean. SD= standard deviation. ΔD= delta deviance. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

servers were subject to a shared work environment, further investigation is warranted to understand the drivers of these behavioral differences. To address this issue, I conducted a second study in which I examined agents' motivational attitudes as potential individual-level factors that influence the between-server variability in servers' productivity level, productivity dispersion, and reaction to system workload.

5. STUDY 2

In this study, I test Hypotheses 1-7 by combining survey data with archival call logs, as outlined below.

5.1 Survey Procedure

An electronic survey was distributed to the call center team members by e-mail during the last week of February 2017. To encourage participation, participants were given a $10 gift certificate as compensation for completing the survey. In addition, team members had an opportunity to enter a raffle for a chance to win one of five gift certificates, each worth $100. Participants were ensured that their responses would be kept confidential, and that only aggregate-level reports will be shared with the leadership team. The number of call agents who filled in Study 2's survey measures was 57 agents, representing a 78% participation rate.

5.2 Data

My sample includes all calls processed in the call center during March 2017 for a total of 67,194 calls processed by 73 agents. Similar to Study 1, the archival logs included detailed information on each stage of the call. I followed a procedure similar to that of Study 1 to construct my Study 2 variables, as outlined in the section below. I dropped calls that had zero talk time (e.g., customer hung-up before talking to an agent), were received outside normal work hours, had an extreme call wrap-up time, or were processed by a survey non-participant, leaving 52,574 calls and 57 agents. Finally, at the time of Study 2, agents were paid using an hourly pay structure with non-guaranteed performance-based incentives that were not directly linked to busy-period performance[1].

5.3 Measures

Tables 5.1 and 5.2 include summary statistics of Study 2 variables that were used in the productivity level and productivity dispersion models, respectively. Operationalization of these measures are discussed below.

5.3.1 Dependent Variables

Call wrap-up time (WT_{ij}). Measured in a similar fashion to Study 1.

Dispersion in daily wrap-up time (WD_{ij}). Measured in a similar fashion to Study 1.

5.3.2 Independent Variables

Workload (WL_{ij}). Measured in a similar fashion to Study 1[2].

Intrinsic Motivation (IM_j). Intrinsic motivation was measured with a four-item scale used in Grant (2008). Agents were asked "Why are you motivated to do your job?" and were presented with the following statements: "Because I enjoy the work itself", "Because it's fun", "Because I find the work engaging", and "Because I enjoy it". The agents were asked to indicate the degree to which they agreed with the above statements using a 7-point Likert-type scale with anchors of 1 (strongly disagree) to 7 (strongly agree). The mean intrinsic motivation score was 5.50 ($\alpha = 0.94$).

Extrinsic Motivation (EM_j). Extrinsic motivation was measured with a four-item scale used in Grant (2008). Agents were asked "Why are you motivated to do

[1]Agents were evaluated on weekly/monthly mean performance metrics that were not directly tied to busy period performance. Higher performance resulted in higher chances of winning a raffle prize, but rewards were not guaranteed.

[2]Again, system workload was measured based on the full (rather than reduced) sample to reflect more accurate workload levels.

Table 5.1.
Means, Standard Deviations, and Correlations (Study 2: Productivity Level Model)

Variable	Mean	SD	1	2	3	4	5	6	7	8
1. Wrap-Up Time	53.99	93.54	—							
2. Workload	3.00	1.14	**0.09**	—						
3. Intrinsic Motivation	5.50	1.29	**-0.16**	**0.01**	(0.94)					
4. Extrinsic Motivation	5.61	1.49	**0.02**	-0.01	-0.05	(0.90)				
5. Number of Agents	13.62	4.60	**-0.04**	**-0.15**	**0.09**	**0.05**	—			
6. On-Line Service Time	240.64	204.99	**0.09**	**-0.03**	**0.02**	**-0.03**	**0.01**	—		
7. Time	4.23	2.38	0.01	**-0.04**	-0.00	**-0.01**	-0.01	**0.03**	—	
8. Overwork$_{K=4}$	0.13	1.10	**0.10**	**0.71**	**0.01**	0.00	**-0.12**	0.00	**0.34**	—

Note: Bold denotes significance at the 5% level. Coefficient alpha estimates of reliability are in parentheses on the diagonal.

Table 5.2.

Means, Standard Deviations, and Correlations (Study 2: Productivity Dispersion Model)

Variable	Mean	SD	1	2	3	4	5	6	7	8
1. Wrap-Up Dispersion	47.42	17.45	—							
2. Workload Level	2.90	0.97	**0.07**	—						
3. Workload Dispersion	0.60	0.2	**0.14**	**0.72**	—					
4. Intrinsic Motivation	5.50	1.29	**-0.28**	0.02	-0.01	(0.94)				
5. Extrinsic Motivation	5.61	1.49	**0.06**	0.03	0.03	-0.05	(0.90)			
6. Number of Calls	47.42	17.45	**-0.17**	**0.50**	**0.38**	**0.09**	0.03	—		
7. Number of SLs	3.50	1.34	**0.08**	**0.73**	**0.58**	0.01	-0.05	**0.42**	—	
8. Number of Servers	22.49	6.53	**-0.10**	**0.52**	**0.27**	**0.21**	0.05	**0.33**	**0.42**	—

Note: Bold denotes significance at the 5% level. M= Mean. SD= Standard Deviation. SLs= Service Lines. Coefficient alpha estimates of reliability are in parentheses on the diagonal.

your job?" and were presented with the following statements: "Because I need to pay my bills", "Because I need to earn money", "Because I have to", and "Because I need the income". The agents were asked to indicate the degree to which they agreed with the above statements using a 7-point Likert-type scale with anchors of 1 (strongly disagree) to 7 (strongly agree). The mean extrinsic motivation score was 5.61 ($\alpha = 0.90$).

5.3.3 Controls

I used the same set of controls as in Study 1. Other additional control variables were considered but were omitted from the study to achieve model parsimony[3].

5.4 Econometric Specification

Following the empirical strategy used in Study 1, I used multilevel modeling to account for the nesting effects in my data. Next, I discuss the main differences between the Study 1 and Study 2 models.

Let $\overline{WL}_{D(i)}$ denote the mean workload level during the day call i was received, $\overline{IM}$ denote the grand sample mean of intrinsic motivation, and $\overline{EM}$ denote the grand sample mean of extrinsic motivation. Then, the main difference between the productivity level models in Studies 1 and 2 is the mean-centering of the workload variables and the inclusion of mean-centered intrinsic and extrinsic motivation effects (both as main and cross-level interaction effects). I used mean-centering to facilitate the interpretation of the model results given the presence of interaction effects. The final productivity level model is given by Model 3 below:

[3]The additional potential control variables were statistically insignificant and did not improve model fit (as measured by a χ^2 test on the difference in deviance statistics). These variables included agents tenure, position (i.e., generalist vs. specialist), and work location (i.e., main office, clinics, or home).

Model 3: Productivity Level Model (Study 2)

Level 1 Model	$WT_{ij} = \beta_{0j} + X_{ij}B_{1j} + \beta_{2j}(WL_{ij} - \overline{WL}_{D(i)}) + \beta_{3j}(WL_{ij} - \overline{WL}_{D(i)})^2 + r_{ij}$
Level 2 Model	$\beta_{0j} = \gamma_{00} + \gamma_{01}(IM_j - \overline{IM}) + \gamma_{02}(EM_j - \overline{EM}) + u_{0j}$ $B_{1j} = \Gamma_{10}$ $\beta_{2j} = \gamma_{20} + \gamma_{21}(IM_j - \overline{IM}) + \gamma_{22}(EM_j - \overline{EM}) + u_{2j}$ $\beta_{3j} = \gamma_{30}$
Assumptions	$r_{ij} \sim N(0, \sigma_r^2)$ and $\begin{bmatrix} u_{0j} \\ u_{2j} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right)$ for call i and agent j

Similar to Study 1, positive linear and negative quadratic coefficients of workload indicate support for Hypothesis 1. Also, significant variance component estimates of u_{0j} and u_{2j} indicate support for Hypotheses 2a and 2c, respectively. Furthermore, a negative coefficient of the intrinsic motivation main effect indicates support for Hypothesis 3.

Let $\overline{WL}$ denote the grand mean of system workload, $\theta_{WL_{High} \to WT\,|\,IM}$ ($= E[WT\,|\,IM = m, WL = w_{High}, X = X_o]$) denote the conditional effect of high workload on wrap-up time as a function of intrinsic motivation (Hayes, 2013), $\theta_{WL_{Low} \to WT\,|\,IM}$ ($= E[WT\,|\,IM = m, WL = w_{Low}, X = X_o]$) denote the conditional effect of low workload on wrap-up time as a function of intrinsic motivation, and $\Theta_{WL \to WT\,|\,IM}$ ($= \theta_{WL_{High} \to WT\,|\,IM} - \theta_{WL_{Low} \to WT\,|\,IM}$) denote the difference between the two conditional effects as a function of intrinsic motivation. Then, to test Hypothesis 5, I examined the 3-D surface plot of the joint conditional effects of intrinsic motivation and workload on wrap-up time (controlling for other variables), probed $\Theta_{WL \to WT\,|\,IM}$ at 25 arbitrary and equally spaced values of intrinsic motivation that covered the whole range of the IM scale to visualize the difference between the workload conditional effects as a function of intrinsic motivation (Aiken & West, 1991; Hayes, 2013), and examined the rate of change in $\Theta_{WL \to WT\,|\,IM}$ for a unit of change in intrinsic motivation (i.e., $\frac{\partial \Theta_{WL \to WT \mid IM}}{\partial IM}$). I followed Aiken & West's (1991) recommendation to determine high and low values of workload where a workload level that is one standard deviation above (below) $\overline{WL}$ reflects a high (low) system workload level, respectively.

Next, let $\mu_{IM_{High} \to WT \mid WL}$ $(= E[WT \mid IM = m_{High}, WL = w, X = X_o])$ denote the conditional effect of high intrinsic motivation on wrap-up time as a function of workload, $\mu_{IM_{Low} \to WT \mid WL}$ $(= E[WT \mid IM = m_{Low}, WL = w, X = X_o])$ denote the conditional effect of low intrinsic motivation on wrap-up time as a function of workload, and $M_{IM \to WT \mid WL}$ $(= \mu_{IM_{High} \to WT \mid WL} - \mu_{IM_{Low} \to WT \mid WL})$ denote the difference between the two conditional effects as a function of workload. Then, to test Hypothesis 6, I examined the 3-D surface plot of the joint conditional effects of intrinsic motivation and workload on wrap-up time (controlling for other variables), probed $M_{IM \to WT \mid WL}$ at 25 arbitrary and equally spaced values of system workload that covered the range of the scale to visualize the difference between the intrinsic motivation conditional effects as a function of workload (Aiken & West, 1991; Hayes, 2013), and examined the rate of change in $M_{IM \to WT \mid WL}$ for a unit of change in workload (i.e, $\frac{\partial M_{IM \to WT \mid WL}}{\partial WL}$). Again, I followed Aiken & West's (1991) recommendation to determine high and low values of intrinsic motivation where an intrinsic motivation score that is one standard deviation above (below) $\overline{IM}$ reflects a high (low) intrinsic motivation level, respectively.

As for the productivity dispersion model, the main difference between the Study 1 and Study 2 models is the inclusion of mean-centered intrinsic and extrinsic motivation main effects in the level-2 intercept, as indicated in Model 4 below:

Model 4: Productivity Dispersion Model (Study 2)

Level 1 Model	$WD_{dj} = \pi_{0j} + \Psi_{dj}\Pi_{1j} + \epsilon_{dj}$
Level 2 Model	$\pi_{0j} = \alpha_{00} + \alpha_{01}(IM_j - \overline{IM}) + \alpha_{02}(EM_j - \overline{EM}) + \delta_{0j}$
	$\Pi_{1j} = A_{10}$
Assumptions	$\epsilon_{dj} \sim N(0, \sigma_\epsilon^2)$ and $\delta_{0j} \sim N(0, \widehat{\sigma}_0^2)$ for day d and agent j

A significant variance component estimate of δ_{0j} indicates support for Hypothesis 2b while a negative coefficient of intrinsic motivation indicates support for Hypothesis 4.

Finally, to test Hypothesis 7, I used t-tests to examine the difference in the coefficients of the main and cross level effects of intrinsic and extrinsic motivation in the productivity models above where applicable. T-tests supporting the following relations indicate support for Hypotheses 7a, 7b, and 7c, respectively: $\gamma_{01} < \gamma_{02}$, $\alpha_{01} < \alpha_{02}$, and $\gamma_{21} < \gamma_{22}$.

I ran these models using SAS 9.4 PROC MIXED procedure with Maximum Likelihood (ML) estimation method. The results of these models are discussed below.

5.5 Results

To test Hypothesis 2, I followed an empirical strategy similar to the one outlined in Study 1. Supporting Hypothesis 2a, there was significant variation in call wrap-up time at both the within- and between-agent levels (column 1 of Table 5.3). I found around 29% of the total variance in call wrap-up time was attributed to the agent effect. Supporting Hypothesis 2b, I found significant variation in daily wrap-up time dispersion at both levels of analysis where 53% of the total variance in daily wrap-up time dispersion was attributed to the agent effect (column 1 of Table 5.4).

Finally, supporting Hypothesis 2c, the results in column 3 of Table 5.3 show significant between-agent variation in agent's reaction to system workload (37.53, $p < 0.01$). Therefore, Hypothesis 2 was fully supported.

To test Hypothesis 1, I also followed an empirical strategy similar to that of Study 1. As shown in Table 5.3 (column 3), the results suggest an inverted U-relationship between workload and agent's wrap-up time. Figure 5.1 illustrates this relationship where at lower levels of workload, agents slowed down with the increase in workload, suggesting a social loafing behavior. However, agents' wrap-up time saturated at higher workload levels. For example, an increase in workload from the minimum to the 25th percentile resulted in a significant increase of 10.17 seconds in call wrap-up time ($p < 0.01$). In contrast, an increase in workload from the 50th to the 75th percentile resulted in a non-significant increase of 0.74 seconds in call wrap-up time. Thus, Hypothesis 1 was supported.

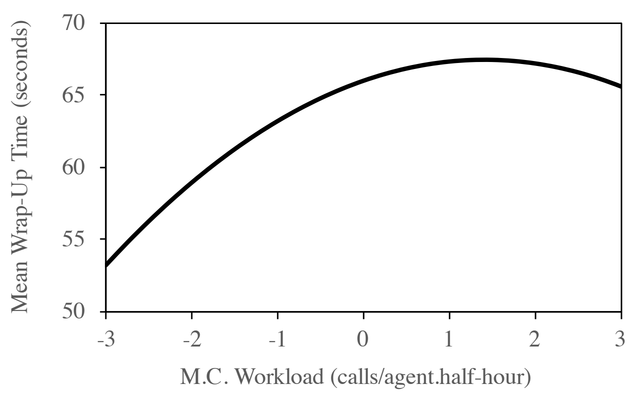

Figure 5.1. Call Wrap-Up Time as a Function of Mean-Centered (M.C.) Workload (Study 2)

Next, I examined Hypothesis 3, which predicted that intrinsic motivation is positively related to call agents' productivity levels. As shown in Table 5.3 (column 4), the coefficient of intrinsic motivation is negative and statistically significant (-11.21, $p < 0.05$). This suggests that on average, an agent with a high intrinsic motivation score (i.e., 1 standard deviation above the mean) wrapped up calls 29 seconds faster than an agent with a low intrinsic motivation score (i.e., 1 standard deviation below the mean). This result supports Hypothesis 3.

Hypothesis 4 predicted that intrinsic motivation is negatively related to fluctuations in agents' daily wrap-up time. Supporting this hypothesis, the coefficient of intrinsic motivation in column 3 of Table 5.4 is negative and statistically significant (-13.36, $p < 0.01$). This suggests that variance in wrap-up time is lower for agents with higher intrinsic motivation levels.

The joint effects of intrinsic motivation and workload on call wrap-up time

Next, to examine the joint effects of workload and intrinsic motivation on call wrap-up time, I plotted wrap-up time (z-axis) as a function of the mean-centered intrinsic motivation (y-axis) and the mean-centered system workload (x-axis), as illustrated in Figure 5.2, using the regression coefficients from column 5 of Table 5.3 (controlling for day of week at Monday, service line at the reference line, and remaining variables at their mean).

Hypothesis 5 predicted that intrinsic motivation attenuates the workload effect on agents' productivity. Looking at Figure 5.2, it is clear that the 3-D surface plot becomes flatter as we move along the y-axis toward higher values of intrinsic motivation, suggesting that intrinsically motivated agents are less affected by the workload effect. This insight is evident in the YZ view of the 3-D surface plot (Figure 5.3) where I note that the workload-driven variability in call wrap-up time (as indicated by the width of the band) decreases as intrinsic motivation increases. These observations provide initial support for Hypothesis 5. Next, I plot the conditional effects of high and low workload levels on wrap-up time as a function of intrinsic motivation (Figure 5.4). The plot shows that differences in wrap-up time that are driven by differences in workload levels (i.e., the vertical distance between the green and red lines) are smaller for agents with higher intrinsic motivation levels. To test this difference rigorously, I probe $\Theta_{WL \to WT \mid IM}$ at different values of intrinsic motivation that cover the IM scale. Figure 5.5 illustrates the point estimates and the confidence interval of $\Theta_{WL \to WT \mid IM}$ at different values of intrinsic motivation (based on t-tests). The plot shows that the workload-driven, within-agent differences in wrap-up time are significant for agents with lower IM scores. For example, an agent with a minimum intrinsic motivation level spends 18.76 seconds ($p < 0.05$) more on call wrap-up during periods of high compared to low workload. However, this workload-driven, within-agent difference in wrap-up time decreases as agents' intrinsic motivation increases to the point where it becomes statistically insignificant for agents with high intrinsic motivation levels. Furthermore, I found the rate of change in $\Theta_{WL \to WT \mid IM}$ for a unit increase in intrinsic motivation level to be negative (-3.49, $p < 0.05$), which means that a unit increase in intrinsic motivation reduces the difference in wrap-up time between high and low congestion periods by 3.49 seconds. Taken together, these results suggest that intrinsically motivated agents are more resilient to the workload effect, supporting Hypothesis 5.

Hypothesis 6 predicted that workload magnifies the intrinsic motivation effect on agents' productivity. Looking at Figure 5.2, I notice that differences in mean wrap-up times between agents with low and high intrinsic motivation levels become larger as we move along the x-axis toward higher levels of workload. Looking at the XZ view of the 3-D surface plot (Figure 5.6), I notice that the motivation-driven variability in call wrap-up time (as indicated by the width of the band) increases with the increase in system workload. These observations provide initial support for Hypothesis 6. Next, I plot the conditional effects of high and low intrinsic motivation levels on wrap-up time as a function of workload (Figure 5.7). The plot shows that differences in wrap-up time that are driven by differences in intrinsic motivation levels (i.e., vertical distance between the green and red curves) are larger for higher workload levels. To test these differences rigorously, I probe $M_{IM \to WT \mid WL}$ at different values of workload that cover the scale. Figure 5.8 illustrates the point estimates and the confidence interval (based on t-tests) of $M_{IM \to WT \mid WL}$ at different values of workload. The plot shows that the intrinsic motivation-driven, between-agent differences in wrap-up time

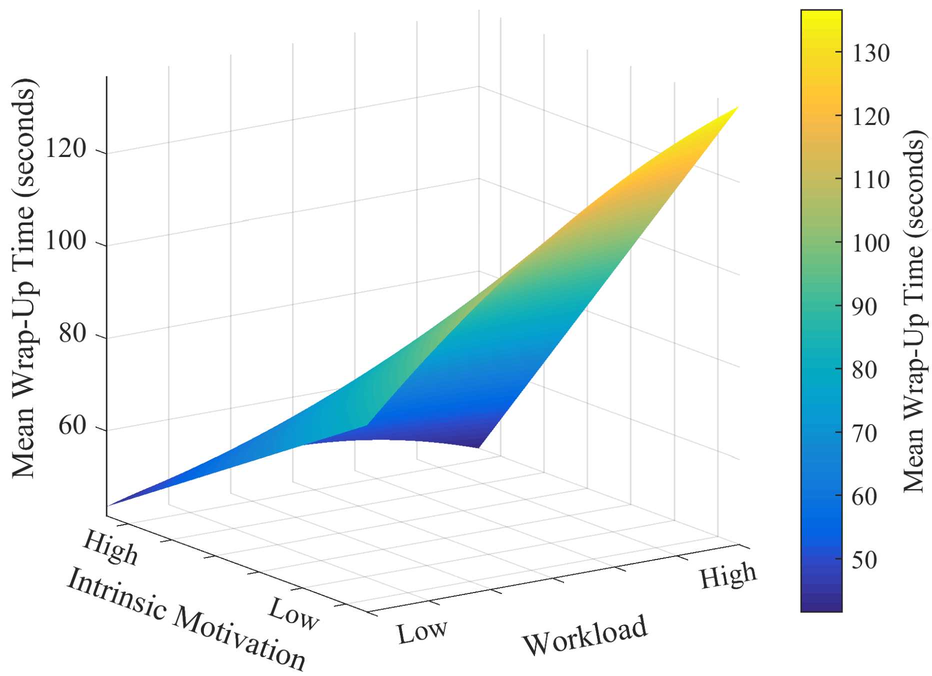

Figure 5.2. Mean Wrap-Up Time (z-axis) as a Function of Workload (x-axis) and Intrinsic Motivation (y-axis)

are statistically insignificant when workload levels are low. However, these differences increase and become significant when workload levels are high. For example, an agent with high intrinsic motivation level spends 40 fewer seconds on call wrap-up than an agent with low intrinsic motivation level when workload levels are 3 units above the sample mean ($p < 0.05$). Furthermore, I found the rate of change in $M_{IM \rightarrow WT|WL}$ for a unit increase in workload to be negative (-4.04, $p < 0.05$), suggesting that a unit increase in workload increases the difference in wrap-up time between agents with high and low intrinsic motivation levels by 4.04 seconds. Taken together, supporting Hypothesis 6, these results suggest that the between-agent differences in productivity level are more evident when the system is congested.

Finally, Hypothesis 7 predicted that the effects of intrinsic motivation on call agents' (a) productivity level, (b) productivity dispersion, and (c) reaction to system

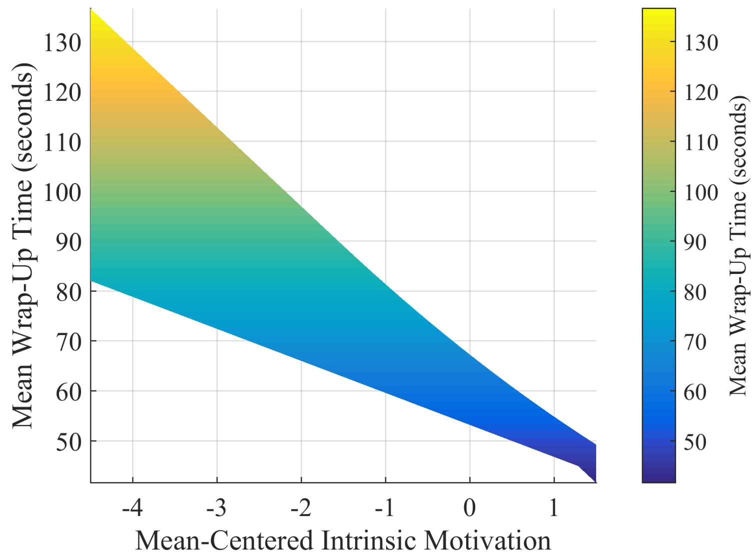

Figure 5.3. Mean Wrap-Up Time (z-axis) as a Function of Workload (x-axis) and Intrinsic Motivation (y-axis): YZ View

workload are more favorable than the effects of extrinsic motivation. I used t-tests to examine differences in the intrinsic and extrinsic motivation coefficients (including the interaction terms) in column 5 of Table 5.3 and column 3 of Table 5.4. I found the main effect of intrinsic motivation on agents' wrap-up time to be more favorable than the main effect of extrinsic motivation, although at a 10% level of statistical significance (-12.03, $p < 0.10$). Similarly, I found the intrinsic motivation effect on agents' wrap-up time dispersion to be more favorable than the extrinsic motivation effect (-15.20, $p < 0.05$). Finally, the intrinsic motivation effect on agents' reaction to workload was more favorable than the extrinsic motivation effect (-1.95, $p < 0.05$). Thus, Hypothesis 7 was supported.

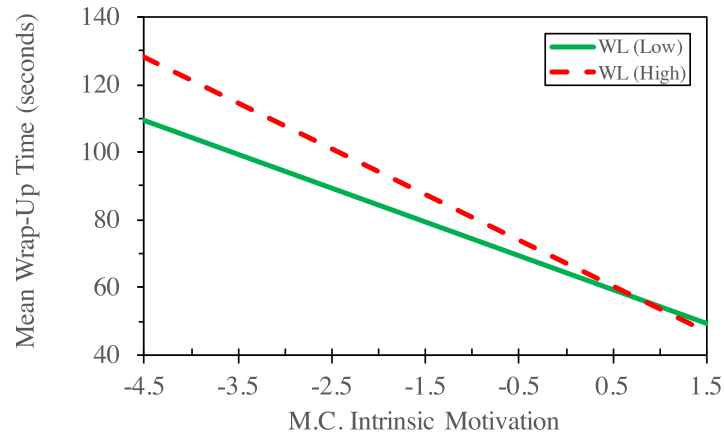

Figure 5.4. Conditional Effects of High and Low Workload on Wrap-Up Time as a Function of Intrinsic Motivation

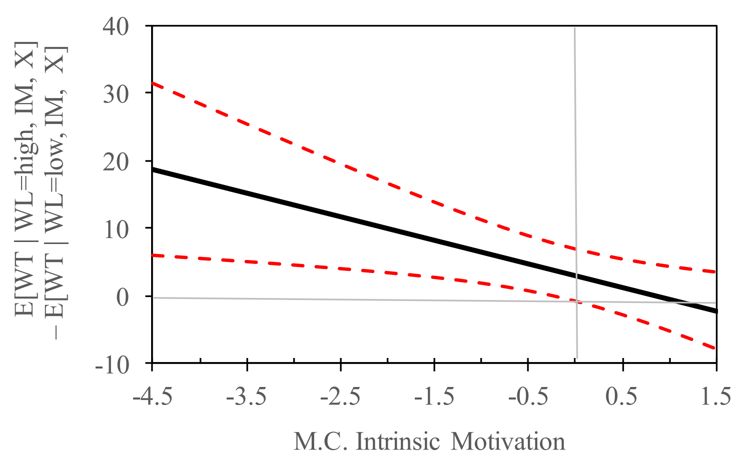

Figure 5.5. Difference in the Conditional Effects of High and Low Workload on Wrap-Up Time as a Function of IM (Notes: Black line indicates point estimates. Red dashed lines indicate 95% confidence interval)

Table 5.3.
Joint Effects of Workload and Motivation on Productivity Level (Study 2)

	Dependent variable: Call Wrap-Up Time				
	(1)	(2)	(3)	(4)	(5)
Fixed Effects					
Intercept	58.30***	52.32***	51.22***	51.19***	51.05***
	(6.71)	(7.96)	(7.94)	(7.71)	(7.70)
Level-1 Main Effects (within-agent)					
M.C. Workload	—	—	2.06**	2.07**	2.02**
			(0.99)	(0.99)	(0.94)
M.C. Workload Square	—	—	-0.73**	-0.73**	-0.74**
			(0.33)	(0.32)	(0.32)
Level-2 Main Effects (between-agent)					
M.C. Intrinsic Motivation	—	—	—	-11.21**	-11.12**
				(5.09)	(5.09)
M.C. Extrinsic Motivation	—	—	—	0.96	0.91
				(4.41)	(4.40)
Cross-Level Interaction Effects					
M.C. Intrinsic Motivation × *M.C. Workload*	—	—	—	—	-1.57**
					(0.62)
M.C. Extrinsic Motivation × *M.C. Workload*	—	—	—	—	0.38
					(0.55)
Random Effects (Variance Components)					
Residual (within-agent)	6363.51***	6154.83***	6141.67***	6141.70***	6141.76***
	(39.27)	(37.98)	(37.93)	(37.93)	(37.93)
Intercept (between-agent)	2554.91***	2682.20***	2584.79***	2377.99***	2372.46***
	(479.75)	(507.42)	(490.80)	(451.55)	(450.37)
M.C. Workload (between-agent)	—	—	27.53***	27.31***	22.39***
			(8.77)	(8.73)	(7.69)
Controls					
Service Line F.E.	—	Yes	Yes	Yes	Yes
Number of Agents	—	-0.72***	-0.65***	-0.65***	-0.65***
		(0.22)	(0.23)	(0.23)	(0.23)
Call Duration (On-Line)	—	0.07***	0.07***	0.07***	0.07***
		(0.002)	(0.002)	(0.002)	(0.002)
Time	—	3.31***	3.56***	3.56***	3.58***
		(0.75)	(0.77)	(0.77)	(0.77)
Time Square	—	-0.44***	-0.46***	-0.46***	-0.47***
		(0.08)	(0.08)	(0.08)	(0.08)
Day of Week F.E.	—	Yes	Yes	Yes	Yes
$Overwork_{K=4}$	—	2.32***	1.77**	1.77**	1.75**
		(0.56)	(0.67)	(0.67)	(0.67)
-2 Log Likelihood (Deviance)	609992.1	608243.8	608193.4	608188.7	608182.1
Akaike Information Criterion	609998.1	608317.8	608273.4	608272.7	608270.1
ΔD		1748.3***	50.4***	4.7*	6.6**

Notes: N=52,574 calls/57 agents. Standard errors are in parenthesis. F.E.= Fixed Effects. M.C.= Mean Centered. ΔD= Delta Deviance. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 5.4.
Motivation Effect on Productivity Dispersion (Study 2)

	Dependent variable: Dispersion in Daily Wrap-Up Time		
	(1)	(2)	(3)
Fixed Effects			
Intercept	85.35***	96.98***	95.37***
	(6.79)	(12.93)	(12.67)
Level-2 Main Effects (within-agent)			
M.C. Intrinsic Motivation	—	—	-13.36***
			(4.85)
M.C. Extrinsic Motivation	—	—	1.84
			(4.19)
Random Effects (Variance Components)			
Residual (within-agent)	2190.87***	2140.07***	2140.13***
	(94.22)	(92.11)	(92.11)
Intercept (between-agent)	2510.81***	2353.64***	2051.31***
	(490.16)	(467.2)	(410.01)
Controls			
Day of Week F.E.	—	Yes	Yes
Number of Calls	—	-0.40***	-0.40***
		(0.14)	(0.13)
Number of Service Lines	—	-4.56*	-4.44*
		(2.46)	(2.45)
Number of Servers	—	-0.24	-0.18
		(0.41)	(0.41)
Workload (M)	—	2.74	2.64
		(3.78)	(3.77)
Workload (SD)	—	33.40***	33.52***
		(11.44)	(11.43)
-2 Log Likelihood (Deviance)	12163.0	12134.1	12126.6
Akaike Information Criterion	12169	12158.1	12154.6
ΔD		28.9***	7.5**

Notes: N=1,138 agent-day observations/57 agents. Standard errors are in parenthesis. F.E.= fixed effects. M.C.= mean centered. M= mean. SD= standard deviation. ΔD= delta deviance.
*, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

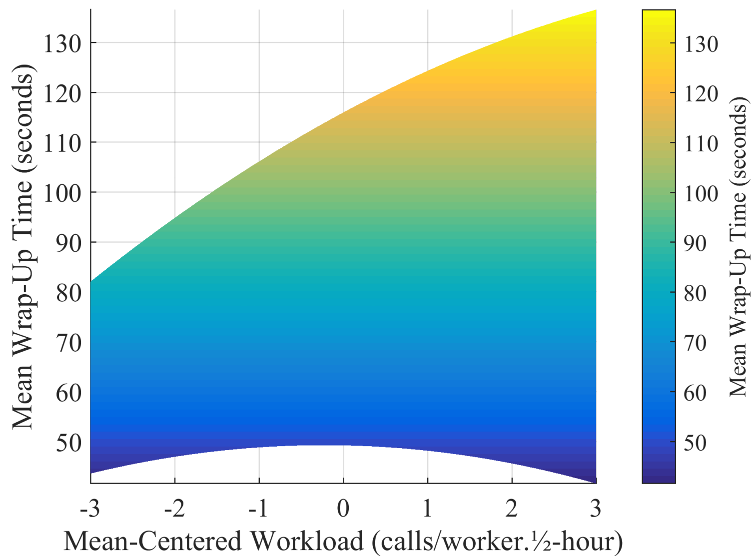

Figure 5.6. Mean Wrap-Up Time (z-axis) as a Function of Workload (x-axis) and Intrinsic Motivation (y-axis): XZ View

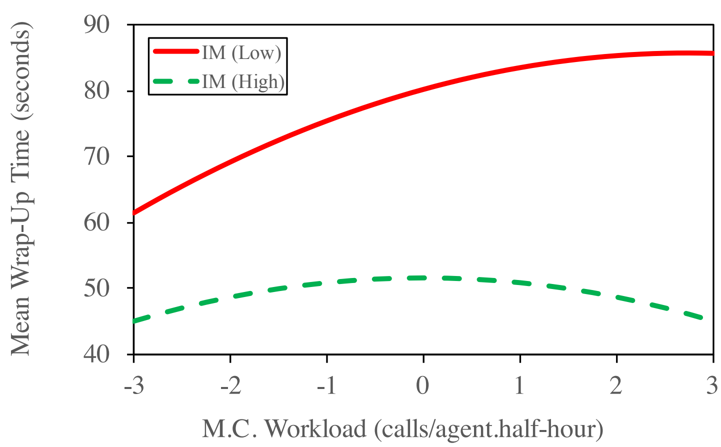

Figure 5.7. Conditional Effects of High and Low Intrinsic Motivation on Wrap-Up Time as a Function of Workload

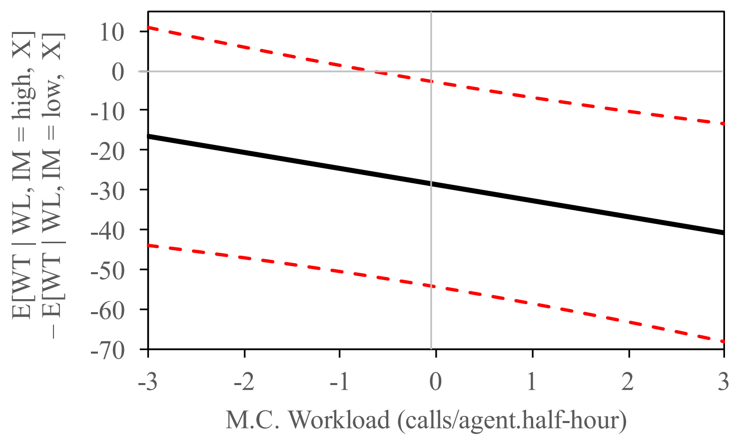

Figure 5.8. Difference in the Conditional Effects of High and Low Intrinsic Motivation on Wrap-Up Time as a Function of Workload
(Notes: Black line indicates point estimates. Red dashed lines indicate 95% confidence interval)

6. ROBUSTNESS CHECK AND SENSITIVITY ANALYSIS

6.1 Non-Response Bias

The final sample of Study 2 consisted of calls serviced only by agents who responded to the survey; agents who did not respond to the survey were omitted from the study. This raises concerns about the generalizability of the results if one assumes that there were systematic differences between survey respondents and non-respondents (e.g., if survey respondents were more productive than non-respondents). To address this concern, I note the following. First, around 78% of agents responded to the survey, which means that only a smaller portion of agents were omitted from the final sample. Second, the results of Study 2 replicated the findings of Study 1, which used the full sample, suggesting that the workload effect is consistent across the two samples. Third, as an additional robustness check, I used multilevel random intercept models (aka null models; Singer & Willett, 2003) to examine whether there were statistical differences between survey respondents and non-respondents in key productivity-related measures: talk time, hold time, and wrap-up time. Controlling for agents random effects, I found no evidence of differences between survey participants and non-participants for the three service time measures, thus reducing concerns about non-response bias.

6.2 Workload Operationalization

In the previous productivity level studies, I operationalized workload as the total count of call arrivals to an agent's virtual queues within a 30-minute time period, divided by the number of agents assigned to service those queues in a given day. While this measure adjusts for the number of agents for a given set of virtual queues, there is a concern with this operationalization. For a given set of virtual queues, a call agent observes the number of calls in the queues and the number of agents assigned to service those queues. However, this server does not directly observe the value of my workload measure. In other words, my workload measure assumes that the server makes a mental calculation to adjust the call arrival rate by the number of agents assigned to service those calls. While this assumption seems plausible, I ran the full productivity level model of Study 2 using an alternative workload measure that does not adjust for the number of servers assigned to a given set of virtual queues. In particular, I operationalized workload as the total count of calls routed to a given set of virtual queues in a given 30-minute time period, and I controlled for the number of agents assigned to service those queues as an independent variable. The mean (standard deviation) of the new workload measure was 40.60 calls/half-hour (15.40), respectively. The results of this model were consistent with the findings of Study 2 (see Table A.1).

6.3 Specification of High and Low Workload

To test Hypothesis 5, I examined differences in the conditional effects of high and low workload on wrap-up time as a function of intrinsic motivation. I followed Aiken & West's (1991) recommendation to determine high and low values of workload where high (low) workload level is conceptualized as a one standard deviation above (below) the mean workload level, respectively. To examine the sensitivity of the results to different specifications of low and high workload, I ran the analysis using two alternative pairs of low and high workload: first vs. third quartiles and 10th vs. 90th percentiles of system workload. The results of these analyses were consistent with the original findings.

6.4 Specification of High and Low Intrinsic Motivation

To test Hypothesis 6, I examined differences in the conditional effects of high and low intrinsic motivation on wrap-up time as a function of system workload. Again, I followed Aiken & West's (1991) recommendation to determine high and low values of intrinsic motivation where a high (low) intrinsic motivation is conceptualized as a one standard deviation above (below) mean intrinsic motivation level, respectively. To examine the sensitivity of the results to different specifications of low and high intrinsic motivation, I ran the analysis using two alternative pairs of low and high intrinsic motivation: first vs. third quartiles and 10th vs. 90th percentiles of intrinsic motivation. The results of these analyses were consistent with the original findings.

7. DISCUSSION

Prior research has assumed that servers react to system workload in a uniform fashion, independent of any individual-level factors. I present a model of service worker productivity that examines the joint effects of workload and intrinsic motivation on service time. This approach allows me to consider the interaction between system-level workload and individual-level motivation. Traditionally, these two lines of research have been studied in separate bodies of literature. The joint examination of the workload and motivation effects contributes to multiple literatures. First, contributing to the behavioral queuing literature, I show that servers' motivation influences the interindividual variation in servers' reaction to system workload. In addition, I show that servers' motivation influences the between-server variation in both the level and dispersion of servers' productivity. Taken together, these results suggest that intrinsically motivated servers are more productive and more resilient to system workload effects. Second, contributing to the motivation literature, I show that workload is a boundary condition for the intrinsic motivation effect. In particular, the intrinsic motivation-driven, between-server differences in productivity are more evident when system workload is higher. This finding addresses calls in the motivation literature for more research that examines "critical points when resources are needed" (Halbesleben, Neveu, Paustian-Underdahl, & Westman, 2014, p. 1353), acknowledges the "role of volition on human action when formulating [motivation] theories" (Locke & Latham, 2004, p. 399), and creates a boundaryless science of work motivation by considering "concepts developed in fields outside OB and I/O psychology" (Locke & Latham, 2004, p. 392). Finally, this research contributes to the OM-OBHR interface by proposing an integrative model of service worker productivity in a specific research context, thus providing insights into "how humans behave in specific operating environments" (Boudreau et al., 2003, p. 197). Indeed, I show that the intrinsic motivation effect on servers' productivity is more favorable than the extrinsic motivation effect in a pooled queuing setting with limited performance-based incentives, suggesting that careful examination of the interplay between operational context and individual-level factors is warranted.

The research design in this study has a number of advantages. First, the use of multilevel models allows me to control for nesting effects that might lead to violation of the independence assumption of residuals in classical regression models (e.g., OLS; Garson, 2012). Second, using a longitudinal research design allows me to ensure temporal order (e.g., individual-level factors are measured before the measurement of servers' subsequent behaviors) and track variation in servers' behaviors over time (Singer & Willett, 2003). Third, the granularity of the operational data allows me to accurately track dynamic changes in the environment (e.g., how workload levels fluctuate over time), isolate outcomes that are under the direct control of the server (e.g., ring, talk, hold, and wrap-up times) from outcomes that are influenced by external factors (e.g., queue time is influenced by staffing levels and the performance of other servers, both of which are not under the direct control of the server), and focus on key components of service time (e.g., off-line vs. on-line service time).

It is also important to take into account the limitations of these studies. First, I examined the impact of intrinsic and extrinsic motivation in a pooled queuing setting that does not reward agents for busy period performance. The results suggest that in such setting, the effect of intrinsic motivation on servers' productivity is more favorable than the effect of extrinsic motivation. However, the results might be different when examining the motivation-based effects in a parallel queuing setting where servers take more responsibility for their own queues (Shunko et al., 2017; Song et al., 2015). Moreover, the effects of extrinsic motivation on servers' productivity might also differ under a performance-based incentive structure that reward busy period performance. Under such an incentive scheme, the instrumentality of the performance-outcomes link is more favorable since servers who desire monetary outcomes are able to attain such rewards through higher performance. This suggests that the motivational force for extrinsically motivated servers might be higher in such settings (Vroom, 1964). Second, I focused my analysis only on the impact of workload and motivation on level and dispersion of service time, which warrants further investigation of the impact of those factors on service quality outcomes. Recent studies in the literature suggest that increases in service speed might come at the expense of service quality (e.g., Kc & Terwiesch, 2009; Tan & Netessine, 2014). However, based on my observations and interviews with call agents in my research site, the quality-speed tradeoff is less (more) of a concern during the off-line (on-line) service stage, respectively. Thus, by focusing on the call wrap-up time as an inverse proxy for agents' productivity, I limit concerns about the speed-quality tradeoff.

Next, I discuss other avenues for future research on service effectiveness, which include examining the relationship between team-level factors and outcomes. Based on my interactions with the call agents and the leadership team at my research site, I also encountered a great interest in exploring the roles of other individual-level factors (e.g., fatigue and burnout) in influencing servers' performance.

8. WHAT DIFFERENCES MAKE A DIFFERENCE? RECOMMENDATIONS FOR FUTURE RESEARCH ON SERVICE EFFECTIVENESS IN QUEUING SYSTEMS

In the previous sections, I examined the joint effects of intrinsic motivation and system workload on servers' productivity within a specific research context. Yet, there are other factors, mechanisms, and outcomes that influence service effectiveness in queuing systems. In this section, I present an integrative multilevel framework that examines the factor-mechanism-outcome (FMO) links in service queuing systems. I also present recommendations for building theories of service effectiveness that are more valid, complete, and useful to practitioners. Figure 8.2 summarizes these relationships and illustrates the multilevel nature of the problem.

8.1 Recommendation 1: Consider the Role of Individual Differences in Influencing Service Effectiveness

OM scholars and practitioners are interested in identifying factors that influence servers' effectiveness in queuing systems. The majority of extant behavioral operations studies examine the role of system-level factors in influencing servers' outcomes. These factors include, but are not limited to, system workload (e.g., Kc & Terwiesch, 2009), queue structure (e.g., Song et al., 2015), information visibility (e.g., Schultz, McClain, & Thomas, 2003), incentive structure (e.g., Shunko et al., 2017), task design strategies (e.g., Staats & Gino, 2012), and work breaks (e.g., Pendem et al., 2016). These factors have been linked to individual-level service times (e.g., patient transport speed, order processing time, meal duration, etc.) and service quality (e.g., sales) outcomes via multiple intermediary mechanisms (e.g., rushing, social pressure, learning, motivation, etc.).

While some of these studies consider the role of individual differences in influencing server behavior to some extent (e.g., by controlling for servers' fixed effects), the majority of the studies in the literature pay little attention to the underlying psychological forces that lie behind these differences. Organizational behavior theories provide insights into the role of individual differences in influencing workers' behaviors as discussed earlier (Cortina & Luchman, 2012). In the discussion that follows, I focus on two major facets that drive individual differences in service settings: motivation and stress.

8.1.1 Theoretical Facet 1: Motivation

Motivation theories provide insights into the drivers of human effort under different situational factors. These theories could be leveraged in behavioral queuing models to understand how different human servers react to different system-level factors. Figure 8.1 illustrates an example of an application of the following motivation theories in a service queuing context: expectancy theory, self-efficacy theory, COR theory, and motive theories. This example is based on a hypothetical server who exerts effort that leads to a given level of individual-level performance (e.g., service time), which in turn leads to a given level of queue-level performance (e.g., average waiting time in the queue, queue abandonment rate), in order to attain certain outcomes (e.g., financial rewards, praise from others, helping others, etc.). Note that in this model of server's behavior, system workload influences queue-level performance where additional units of workload lead to lower queue-level performance if the service rate is not adjusted properly. Also, information visibility influences servers' perceptions of the queue-level performance where higher levels of information visibility lead to more accurate assessments of the actual queue-level performance. I explain below how each of the motivation theories play a role in explaining servers' behaviors in a service queuing setting.

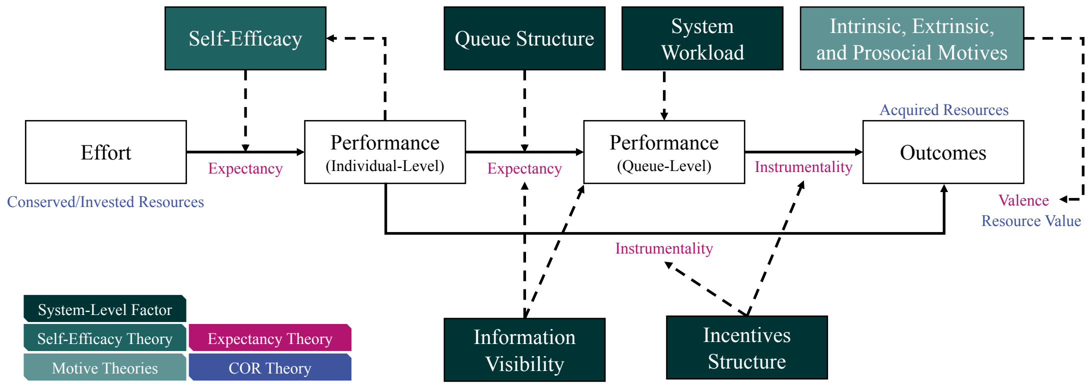

Figure 8.1. An Application of Motivation Theories in a Service Queuing Context

As mentioned in the previous chapters, Expectancy Theory (Vroom, 1964) models workers' motivation as a multiplicative function of three forces: expectancy, instrumentality, and valence. If one of these forces is zero, then the motivational force of the server is zero. In a queuing setting, expectancy refers to servers' beliefs that exerting more effort leads to higher individual-level performance (e.g., higher service rate), which in turn leads to higher queue-level performance (e.g., lower waiting time). Note that expectancy could be influenced by the queue structure since servers' might perceive more control over the queue performance (i.e., via their individual-level performance) under a dedicated (rather than pooled) queuing structure (Shunko et al., 2017; Song et al., 2015). Information visibility might also influence expectancy by providing servers with feedback that could be used to assess more accurately the influence of servers' performance on queue performance. Instrumentality in a service queuing context might be influenced by the incentive structure. Under a flat incentive structure (e.g., fixed pay), instrumentality between performance and financial rewards is weaker. In contrast, a performance-based structure might strengthen the instrumentality between monetary outcomes (e.g., bonuses) and individual-level performance[1], queue-level performance[2], or both. Finally, valence refers to the value of the acquired outcomes to the server.

Self-Efficacy theory (Bandura, 1977) posits another factor that influences servers' expectancy. Self-efficacy is defined as "the belief that a person has the capabilities needed to execute the behaviors required for task success" (Colquitt et al., 2013, p. 167). Studies show that individuals with higher self-efficacy levels exert more effort, persist longer, and perform better on assigned tasks (Bandura, 1997; Stajkovic & Luthans, 1998). Moreover, studies show that prior task accomplishments influence posterior self-efficacy levels (Bandura, 1982). In other words, servers' past success or failure with similar tasks influences their future self-efficacy levels, which in turn influence the amount of effort they exert and, subsequently, their individual-level performances.

Motive-based theories address the question of which outcomes motivate a worker. In the previous sections, I explored the role of intrinsic and extrinsic motivation in influencing servers' behavior in a pooled queuing environment with limited performance-based incentives. In such a setting, the intrinsic motivation effect was more favorable than the extrinsic motivation effect, which might be explained by the stronger (weaker) instrumentality between servers' performance and intrinsic (extrinsic) outcomes, respectively. However, the results might change under a performance-based incentive structure by which servers are rewarded with extrinsic outcomes as a compensation for their effort. Thus, it is important to consider both the context and motives when examining servers' effectiveness in queuing systems. Finally, recent evidence in the literature suggests that prosocial motivation, which is defined as the desire to exert effort to help others, is positively related to workers' productivity, persistence, and performance (Grant & Berry, 2011; Grant, 2008). Thus, prosocial motives might constitute a third category of server outcomes.

[1] If servers were rewarded based on their individual-level performance

[2] if servers were rewarded based on the queue-level performance

Conservation of Resources theory proposes that individuals are motivated to "conserve" their valuable resources and "acquire" new resources (Hobfoll, 1989). Principles of COR theory suggest the primacy of resource loss (i.e., harms that result from losing current resources overweigh benefits of acquiring equivalent resources), the need to invest current resources to acquire new resources, the lack of current resources leads to defensive mechanisms (i.e., to conserve remaining resources), and the surplus in initial resources leads to more resource investment (Halbesleben, 2010; Halbesleben & Bowler, 2007; Halbesleben et al., 2009, 2014; Halbesleben & Wheeler, 2008, 2011; Hobfoll, 1989; Ng & Feldman, 2012; Vinokur & Schul, 2002). Taken together, the COR principles suggest that individuals are motivated to expend the minimum amount of resources needed to acquire new valuable resources. In a service setting, server's effort denotes "invested resources", attained outcomes denote "acquired resources", and value attached to attained outcomes denotes "resource value". Note that individuals may assign different values to different outcomes, as discussed earlier (e.g., financial rewards, performing a task, helping others, etc.). From an OM perspective, the COR theory suggests that servers are strategic decision makers who aim to minimize their exerted effort and maximize their acquired outcomes.

Another key motivation theory is the Goal-Setting theory, the most dominant motivation theory in the extant literature, which suggests that setting specific challenging goals rather than vague or "do your best" goals leads to higher task performance (Locke & Latham, 1990; Mitchell & Daniels, 2003). This theory suggests that servers who set specific challenging goals (e.g., process x calls per hour, transport patients within x units of time, etc.) might outperform servers who do not set specific goals. However, it is important to note that servers need a minimum level of ability in order for goal-setting to influence their task performance (Locke & Latham, 1990). Moreover, goal-setting might be less effective if goal commitment was low (Tubbs, 1994), or if task complexity was extremely high (Wood, Mento, & Locke, 1987). In summary, motivation is a major predictor of servers' performance. Thus, behavioral queuing models could benefit from incorporating individual-level motivation theories to better understand servers' behaviors under different system designs and states.

8.1.2 Theoretical Facet 2: Stress, Fatigue, and Burnout

Stress, which is defined as "a psychological response to demands that possess certain stakes for the person and that tax or exceed the person's capacity or resources" (Colquitt et al., 2013, p. 130), is a central topic in the study of work behaviors. Stress might be caused by different work-related stressors, including task-related (e.g., workload, time pressure, interruptions), work schedule-related (working time arrangements, long working hours), and social (e.g., poor social interactions with coworkers) stressors (Sonnentag & Frese, 2012). Stress has been linked to physiological (e.g., back pain, high blood pressure, headache, illness), psychological (e.g., forgetfulness, clouded thinking, burnout), and behavioral (e.g., compulsive behaviors) strains (Burke, 2005), all of which have negative implications for workers' well-being and performance (Sonnentag & Frese, 2012). However, it is important to note that "not all individuals react in a uniform manner to the same stressor" (Sonnentag & Frese, 2012, p. 561). Hence, differences in the way individuals cope with work stressors are expected to lead to between-server differences in stress-related strains and performances (Lazarus & Folkman, 1984).

Recent studies in the behavioral operations literature examined the role of extended workload (a task-related stressor) on service time and service quality (Delasay et al., 2015; Kc & Terwiesch, 2009; Staats & Gino, 2012). Many of these studies examine the effects of short-term, within-day strains and assume that servers react to task-related stressors in a uniform fashion. However, future research could benefit from the examination of long-term strains (e.g., burnout) that might influence servers' reaction to workload-related strains. Burnout, which is defined as "the emotional, mental, and physical exhaustion that results from having to cope with stressful demands on an ongoing basis" (Colquitt et al., 2013, p. 141), is a long-term stress reaction that has negative consequences on workers' well-being and performance (Maslach, Schaufeli, & Leiter, 2001). From a conservation of resources perspective, burned-out servers start with lower levels of resources and are thus more likely to engage in defensive mechanisms (e.g., slowdown) to conserve their remaining resources when they face workload stressors (Halbesleben et al., 2014; Sonnentag & Frese, 2012). In summary, future behavioral OM research should consider differences between servers that might lead to differences in their reactions to work stressors.

8.2 Recommendation 2: Consider the Roles of Team-Level Factors and Mechanisms in Influencing Team-Level Outcomes

In order to become more competitive in the current global and dynamic economic environment, organizations are increasingly relying on cross-functional work teams as building structures that facilitate efficient responses to emerging organizational problems (Byrne, 1993; Donnellon, 1996; Ilgen, 1999; Mathieu, Maynard, Rapp, & Gilson, 2008). Many OM studies examined service time and service quality outcomes at the team level of analysis (e.g., Batt & Terwiesch, 2016; Berry Jaeker & Tucker, 2017; Huckman & Staats, 2011; KC, 2014; KC & Terwiesch, 2011, 2012; KC et al., 2013; Kuntz, Mennicken, & Scholtes, 2015; Song et al., 2015). While some of these studies examined the role of team-level factors (e.g., team composition) in influencing team-level outcomes (e.g., Huckman & Staats, 2011; Tan & Netessine, 2015), the majority of the studies in the behavioral operations literature paid little attention to team-level factors or mechanisms.

Drawing from the team effectiveness literature, I explore the influence of the following team factors and mechanisms on the effectiveness of team outcomes: team diversity, team psychological safety, behavioral team processes, affective/motivational team processes, and cognitive team processes. I selected these factors and mechanisms due to their relevance in service queuing settings; however, these are not the only factors that influence team-level outcomes. Hence, I refer readers interested in learning more about other potential team-level factors to Kozlowski & Bell (2012).

8.2.1 Theoretical Facet 1: Team Diversity (Team-Level Factor)

Research on the relationship between team diversity and team outcomes has yielded mixed results (Joshi & Roh, 2009; Kozlowski & Bell, 2012). On the one hand, the optimistic view holds that team diversity leads to higher team performance, since team members have access to a wider pool of knowledge and skills that facilitate knowledge-sharing and team creativity (Cox, Lobel, & McLeod, 1991; Janis, 1982; Nemeth, 1986; Van Der Vegt & Bunderson, 2005). On the other hand, the pessimistic view holds that team diversity undermines team performance since diverse teams suffer from social divisions that hinder team effectiveness and performance (Chatman, Polzer, Barsade, & Neale, 1998; Jehn, 1995; Rosenbaum, 1986; W. G. Wagner, Pfeffer, & O'Reilly, 1984). This suggests that different types of team diversity might have different influences on team performance.

Harrison, Price, Gavin, & Florey (2002) differentiate between surface-level diversity (based on demographic differences) and deep-level diversity (based on differences in attitudes and beliefs). Similarly, Joshi & Roh (2009) differentiate between task-oriented diversity (based on differences among team members in terms of skills and information) and relation-oriented diversity (based on demographic and cultural differences between team members). However, other scholars have argued for the need of multifactor approaches to team diversity, since the two-factor approaches "depend on the measurement of a limited set of variables, often operationalized as only one focal characteristic" (Mannix & Neale, 2005, p. 36). For example, Jehn, Northcraft, & Neale (1999) categorize team diversity into three types: social category diversity, value diversity, and informational diversity. Moreover, Mannix & Neale (2005) categorize team diversity into six types: social-category differences, differences in knowledge and skills, differences in values and beliefs, personality differences, organizational- or community-status differences, and differences in social and network ties.

The similarity-attraction theory suggests that individuals are attracted to others who share similar attributes. For example, Berscheid (1985) found that surface-level similarity (i.e., demographic similarity) is a predictor of attraction between individuals. In contrast, scholars found individuals are more likely to avoid interaction with other individuals whom they dislike (e.g., Chatman et al., 1998; Jehn, 1995; Rosenbaum, 1986). Indeed, W. G. Wagner et al. (1984) found age dissimilarity is associated with less communication between team members, suggesting that social-category differences (e.g., differences in age, gender, race, and/or physical abilities) lead to less interaction between individuals in team environments, which in turn affects team performance (Jehn, 1995). However, it is important to note that social diversity matters most when social-category differences are visible, since the visibility of social-category differences is more likely to invoke stereotypes that influence attitudes and behaviors in work environments (Joshi, Neely, Emrich, Griffiths, & George, 2015; Milliken & Martins, 1996). For example, Duguid & Thomas-Hunt (2015) found that individuals indicated less willingness to work with female coworkers who broke gender stereotypes when these individuals thought that the majority of people do stereotype. This result suggests that the visibility of gender differences and the prevalence of stereotyping behavior impacts employees' willingness to work and interact with other individuals who break stereotypes. Indeed, recent research suggests that individuals use visible cues to categorize others into membership groups, which in turn increases the likelihood of stereotyping and negatively impacts interactions between team members (Joshi et al., 2015; Kulik, Roberson, & Perry, 2007; Milliken & Martins, 1996). Taken together, visible social-category differences among team members might be negatively related to service effectiveness in service queuing systems that utilize teams for interdependent and complex tasks (e.g., teams that provide medical surgery services).

The value in diversity hypothesis (Cox et al., 1991) suggests that diversity of team membership in terms of knowledge and skills is positively related to team outcomes.

The motivating logic behind this assertion is that heterogeneous teams outperform homogenous teams (e.g., Hoffman, 1959; Nemeth, 1986) due to the teams' access to a wider pool of knowledge and skills, increasing the likelihood that a member will have the correct solution to a given problem (Jackson, 1992), increasing the team's capacity for creative problem solving due to alternative perspectives that can lead to novel insights (Nemeth, 1986), and reducing the likelihood of groupthink by providing counter examples that undercut a given assertion (Janis, 1982). Indeed, Van Der Vegt & Bunderson (2005) found diversity in team expertise to be positively related to team performance when team identification was high. The authors suggest that team learning mediates the positive relationship between expertise diversity and team performance. Other studies in the literature support the notion that information-based diversity is positively related to team effectiveness outcomes (Damon, 1991; Homan, Van Knippenberg, Van Kleef, & De Dreu, 2007; Jehn, 1995; Levine, Resnick, & Higgins, 1993). This suggests that service teams that have access to a heterogeneous pool of servers with different functional backgrounds might benefit from having extended access to knowledge and skills and higher capacity for creative problem solving, leading to improvements in key service effectiveness outcomes. In summary, it is important for scholars to consider the role of team diversity in influencing servers' collective performance in teams conducting interdependent and complex tasks.

8.2.2 Theoretical Facet 2: Team Psychological Safety (Team-Level Factor)

Team psychological safety is defined as "a shared belief that the team is safe for interpersonal risk taking" (Edmondson, 1999, p. 354). Psychological safety has been linked to improvements in team effectiveness outcomes by promoting learning behaviors (Edmondson, 1999; Huang, Chu, & Jiang, 2008; Tucker, Nembhard, & Edmondson, 2007) and trust climate (Edmondson & Lei, 2014) at the team level. Studies show that differences in team psychological safety exist among teams within the same organization (Edmondson, 1996, 1999); therefore, it is important to examine the role of team psychological safety in influencing between-team differences in service effectiveness outcomes. For example, Tucker (2007) found that psychological safety was positively associated with frontline system improvement behaviors that were aimed toward reducing operational failures (e.g., missing medical equipment) in medical settings. Moreover, Probst & Estrada (2010) found that psychological safety was positively related to safety practices in both manufacturing and service settings.

Team psychological safety might also moderate the effects of system-level factors on service effectiveness at the team level. Recent OM studies suggest that system workload influences service quality outcomes (e.g., Batt & Terwiesch, 2016; KC & Terwiesch, 2012; Kc & Terwiesch, 2009; Kuntz et al., 2015). In particular, increases in system workload have been linked to higher likelihoods of early discharge, readmission, and mortality in medical settings. However, this negative workload effect might be weaker for teams with higher psychological safety levels. Indeed, Tucker and colleagues found that psychological safety promoted learning behaviors in medical teams which in turn led to lower risk-adjusted mortality rates for infant patients (Nembhard & Tucker, 2011; Tucker et al., 2007). Other studies also show a positive link between psychological safety and team effectiveness outcomes (e.g., Bradley, Postlethwaite, Klotz, Hamdani, & Brown, 2012; Edmondson, 1999; Nembhard & Edmondson, 2006). Therefore, future research would benefit from examining the role of team psychological safety in moderating the effect of system workload and other system-level factors on service effectiveness outcomes at the team level.

Next, I discuss the role of team processes in influencing team outcomes. Team processes are defined as "mechanisms that inhibit or enable the ability of team members to combine their capabilities and behaviors" (Kozlowski & Bell, 2012, p. 430). Kozlowski and colleagues identified three main processes that constitute the core of team mechanisms: behavioral, affective/motivational, and cognitive team processes (Kozlowski & Bell, 2012; Kozlowski & Ilgen, 2006).

8.2.3 Theoretical Facet 3: Behavioral Team Processes (Team-Level Mechanism)

Behavioral team processes are mechanisms that facilitate task organization and goal achievement at the team level by fostering repeated behavioral interactions among team members. These processes include team coordination, cooperation, and communication (Kozlowski & Bell, 2012). Team coordination refers to the process of organizing team members' actions, objectives, and knowledge to attain shared goals (Rico, Sánchez-Manzanares, Gil, & Gibson, 2008). Evidence in the literature suggests that team coordination is a critical process that positively influences team functioning and performance (e.g., Klimoski & Mohammed, 1994; Kraiger & Wenzel, 1997; Levine & Choi, 2004; Marks, Sabella, Burke, & Zaccaro, 2002; Salas, Stagl, & Burke, 2004; Stout, Salas, & Carson, 1994). For example, in a two-person flight simulation task, Stout and colleagues found that members who coordinated their efforts (e.g., by providing information in advance) achieved higher team performance (Stout et al., 1994). In a service queuing context, coordination among team members might have a significant impact on service time and quality when the service outcome is contingent on the integration of multiple actions in a timely and ordered fashion. For example, consider the case of restaurant services where a customer orders a meal. In such a setting, service time (e.g., the time it takes to deliver the order to the customer) and service quality (e.g., meal temperature) depend on the integration of multiple actions that are performed by multiple service providers in the following order: (1) server takes the order from the customer, (2) server passes the order to the kitchen, (3) chef prepares the meal, (4) server picks up the meal, (5) server delivers the meal to the customer. Any delay in the above process might extend service time and reduce service quality (e.g., delays in meal pickup resulting in the food becoming cold). However, coordination among team members might be less of an issue in service settings with non-sequential or independent task structures.

Team cooperation refers to the “willful contribution of personal efforts to the completion of interdependent [team tasks]” (J. A. Wagner, 1995, p. 152). Empirical evidence suggests that cooperation among team members is positively related to team functioning and performance (e.g., Pinto & Pinto, 1990; Smith et al., 1994). Social psychology studies have examined the links between cooperation and social loafing behavior where team cooperation norms have been linked to lower social loafing behavior (Karau & Williams, 1993; Kerr & Bruun, 1983; Latané, Williams, & Harkins, 1979). This suggests that team cooperation behaviors might moderate the influence of queue structure on service time. Recent OM studies suggest that servers are more likely to engage in social loafing behavior in a queuing system with a pooled (rather than parallel) queuing structure (Shunko et al., 2017; Song et al., 2015). However, the queue structure effect might be more evident for teams with weaker cooperation norms. Thus, further investigation of the influence of team cooperation on team service outcomes is warranted.

Finally, team communication refers to the capability of team members to convey relevant information to other members and acknowledge the receipt of such information (Kozlowski & Bell, 2012; Kozlowski & Ilgen, 2006). The literature has offered mixed results regarding the relationship between communication and performance, in which some scholars found a positive relation (e.g., Pinto & Pinto, 1990), negative relation (e.g., Smith et al., 1994), or nonrelation (e.g., Campion, Medsker, & Higgs, 1993) between communication and team performance. However, these discrepancies might be explained by the different motives for communication. For example, higher frequency of communication might reflect higher levels of conflict among team members (e.g., Smith et al., 1994) rather than a desire to share information or request feedback. Thus, when examining communication patterns among team members, one needs to consider both the type and amount of communication. It is also important to note that communication enables the other behavioral team processes of coordination and cooperation (Kozlowski & Bell, 2012). In the restaurant service example, a lack of effective communication between the chef and server might lead to deterioration in service effectiveness outcomes (e.g., chef does not inform server when the food is ready for pickup). In summary, behavioral team processes are key mechanisms that jointly influence team service effectiveness.

8.2.4 Theoretical Facet 4: Affective/Motivational Team Processes (Team-Level Mechanism)

Affective/Motivational team processes are mechanisms that influence team members' collective emotions, compatibilities with other members, and confidence in the team (Evans & Jarvis, 1980; Gibson & Earley, 2007; Jehn, 1995). These processes include team cohesion, team affect, collective efficacy, and conflict (Jehn, 1995; Kozlowski & Bell, 2012). Team cohesion refers to members' attraction and commitment to the team (Evans & Jarvis, 1980; Goodman, Ravlin, & Schminke, 1987). Some scholars distinguish between task cohesion, which reflects members' shared attraction to the team task, and interpersonal cohesion, which reflects members' liking of the team itself (Evans & Jarvis, 1980; Gross & Martin, 1952; Kozlowski & Bell, 2012). Evidence from the literature suggests that team cohesion is an important predictor of team performance, but there have been mixed results regarding whether different team cohesion dimensions have distinct effects on team performance (Beal, Cohen, Burke, & McLendon, 2003; Mullen & Copper, 1994; Zaccaro & McCoy, 1988). In a service queuing context, team cohesion might play a key role in influencing service outcomes when task interdependency is high.

Team affect refers to team members' collective experience of mood and emotions. Scholars suggest that there are two approaches to conceptualize team affect: bottom-up and top-down approaches (Barsade & Gibson, 1998). In a service queuing setting, the bottom-up approach refers to the emergence of individual-level servers' emotions at the team level. In contrast, the top-down approach refers to the contextual influences of higher-level factors (e.g., system workload) on the emotions of team members. Consider, for example, two teams that operate under different levels of system congestion: high vs. low workload. The top-down approach suggests that the shared workload seen by each team has an influence on the emotions of servers within that team. Hence, the emotions of servers in the high congestion team might differ significantly from the emotions of servers in the low congestion team. Other system-level factors might also influence the emotions of team members (e.g., queue structure, routing rule, incentives structure, etc.). Team affect can be negative, positive, or neutral. Negative (positive) team affect has been linked to lower (higher) levels of team performance, respectively (e.g., Cole, Walter, & Bruch, 2008; Grawitch, Munz, & Kramer, 2003). Thus, it is important to understand (a) how different elements of the queuing system might influence team affect (i.e., the top-down approach) and (b) how emotions of individual servers might emerge at the team level (i.e., the bottom-up approach) when studying team service effectiveness.

Collective efficacy refers to members' shared perceptions of the team's capacity to complete a given task successfully (Bandura, 1997; Gibson & Earley, 2007). Collective efficacy has been linked to higher levels of team effort, productivity, and performance (Bandura, 1986; Campion et al., 1993; DeShon, Kozlowski, Schmidt, Milner, & Wiechmann, 2004; Feltz & Lirgg, 1998; Gully, Incalcaterra, Joshi, & Beaubien, 2002; Hodges & Carron, 1992). In a service queuing context, differences among teams in collective efficacy might lead to differences in responses to stress stimuli. For example, a medical team in a metropolitan hospital that has been successful in delivering effective service under different levels of service demands in the past might be more confident and better equipped to deal with sudden surges in service demands than a team in a rural hospital that rarely faces high levels of workload.

Finally, team conflict refers to relational and task conflict among team members. Relational conflict occurs when there are perceived interpersonal incompatibilities among team members. Relational conflict could make team members less satisfied with their team and might lead to team withdrawal behaviors and lower team performance (Amason, 1996; De Dreu & Weingart, 2003; Jehn, 1995; Kozlowski & Bell, 2012). In contrast, task conflict reflects disagreement among team members about task issues. Jehn and colleagues found that task conflict is associated with higher team performance of non-routine tasks (Jehn, 1995; Jehn et al., 1999). However, other scholars questioned the positive impact of task conflict on team performance (e.g., De Dreu & Weingart, 2003). In a service queuing context, relational conflict might negatively affect team communication, coordination, cooperation, and cohesion, which in turn could affect the effectiveness of team outcomes.

8.2.5 Theoretical Facet 5: Cognitive Team Processes (Team-Level Mechanism)

Cognitive team processes are mechanisms that facilitate building team knowledge, processing information, and solving problems. These mechanisms include team learning, team mental models, transactive memory, and macrocognition (Edmondson, 1999; Klimoski & Mohammed, 1994; Kozlowski & Bell, 2012; Milliken & Martins, 1996). In this section, I focus on the role of team learning in influencing service effectiveness at the team level. Readers interested in learning more about the other cognitive team processes are referred to Kozlowski & Bell (2012).

Team learning refers to the "ongoing process of reflection and action, characterized by asking questions, seeking feedback, experimenting, reflecting on results, and discussing errors or unexpected outcomes of actions" (Edmondson, 1999, p. 353). Research in the literature suggests a positive linkage between team learning and various team outcomes (e.g., Edmondson, 1999; Van Der Vegt & Bunderson, 2005; Wong, 2004; Zellmer-Bruhn & Gibson, 2006). For example, in a study of 51 teams in a manufacturing setting, Edmondson (1999) found that team learning behaviors were positively related to self-report, manager, and customer ratings of team performance. Another concept that relates to team learning is information processing, which refers to "the degree to which information, ideas, or cognitive processes are shared, and are being shared, among the [team] members and how this sharing of information affects both individual- and team-level outcomes" (Hinsz, Tindale, & Vollrath, 1997, p. 53).

Evidence from the literature shows that increased information processing leads to positive team outcomes (e.g., Loyd, Wang, Phillips, & Lount Jr, 2013; Sommers, 2006). For example, Philips and colleagues found that increased information processing in diverse teams helped those teams make more accurate decisions and outperform other non-diverse teams that were overconfident in their opinions (Phillips, Liljenquist, & Neale, 2009).

Recent evidence in the OM literature supports the role of learning in influencing service effectiveness at the individual level of analysis (KC et al., 2013; Staats & Gino, 2012). Yet, more research is needed to understand the role of team learning in influencing team-level outcomes. For example, are service quality outcomes more favorable in teams that encourage the reflection on past team errors? Are service effectiveness outcomes more favorable in teams that allow experimentation with new service delivery methods? How do team learning behaviors influence service effectiveness outcomes in the short and long term? These questions among others warrant further investigation.

8.3 Recommendation 3: Consider the Role of Organizational Withdrawal Behavior in Influencing Service Effectiveness

Organizational withdrawal behavior has been a topic of great interest for both academics and practitioners, mainly because workers' withdrawal behaviors could lead to detrimental consequences for organizations (e.g., Hancock, Allen, Bosco, McDaniel, & Pierce, 2013). In this section, I review major theoretical facets of organizational withdrawal behavior. First, I discuss four different categories of withdrawal behaviors: lateness, absenteeism, turnover, and retirement. Next, I differentiate between different types of turnover. Finally, I discuss turnover consequences in service queuing settings.

8.3.1 Theoretical Facet 1: Withdrawal Behaviors

Scholars distinguish between four main categories of organizational withdrawal behavior: lateness, absenteeism, turnover, and retirement (Harrison & Newman, 2012). Lateness is defined as "arriving at work after the start, or leaving before the end, of a scheduled workday", while absenteeism refers to the "tendency to miss scheduled work over a given time interval" (Harrison & Newman, 2012, p. 255-266). These withdrawal behaviors could lead to severe consequences for operational performance in service queuing settings. In such settings, decision-makers strive to balance the demand and supply through optimal scheduling decisions that assume scheduled servers are present at the time service requests arrive. However, lateness or absence of servers could lead to lower effective staffing levels in a given time period, which could negatively impact system performance. Thus, it is important for decision-makers to understand the factors that influence servers work withdrawal behaviors. For example, would higher levels of workload (e.g., busier days or busier times of day) be associated with servers' work withdrawal behaviors? If so, then optimization staffing and scheduling models should account for servers' withdrawal behavior in response to different workload levels to yield more valid and accurate optimal decisions.

Turnover is defined as "movement across the membership boundary of an organization", while retirement is defined as "leaving a career or occupation" (Harrison & Newman, 2012, p. 274-275). Note that unlike lateness and absenteeism behaviors, the target of turnover and retirement behaviors is the job itself rather than the work. Thus, these withdrawal behaviors have implications for the time it takes managers to hire replacement servers, the cost it requires to train the new hires, and the time it takes for the new hires to build levels of experience similar to those of the servers who quit or retired. In the following sections, I discuss the implications of turnover types and consequences on service effectiveness in more detail.

8.3.2 Theoretical Facet 2: Turnover Types

Organizational scholars distinguish between voluntary turnover (i.e., turnover initiated by the employee) and involuntary turnover (i.e., turnover initiated by the organization). Voluntary turnover can occur for a variety of reasons (e.g., quitting to take alternative job) and can be driven by multiple motivational factors (Campion, 1991; Harrison & Newman, 2012; Maertz & Campion, 2004). In service queuing systems, higher voluntary turnover rates might indicate shocks in system that could affect the amount of workload seen by servers in the short run, since these turnover events are unexpected and might lead to lags until replacement servers are hired.

Another distinction made by organizational scholars is between functional turnover (e.g., turnover of poor performers) and dysfunctional turnover (e.g., turnover of good performers), which is based on previous performance levels of quitters that make the quitting behavior more or less desirable to the organization (Campion, 1991). This distinction is important in a service queuing context, since it reflects the degree to which turnover events impact service effectiveness outcomes. In particular, managers might be less (more) worried about the implications of functional (dysfunctional) turnover on service quality outcomes, respectively.

Some scholars also distinguish reduction-in-force turnover (i.e., turnover driven by downsizing factors and is initiated by the organization) from the other types of turnover. The rationale for this distinction is based on the assertion that the firm does not plan to hire replacement employees to fill the positions held by discharged employees, which suggests that the impact of this turnover type on organizational outcomes might differ from other types of turnover (McElroy, Morrow, & Rude, 2001; Park & Shaw, 2013). In a service queuing context, this type of turnover might have implications on the long-term workload levels seen by servers, since the top management is not planning to hire replacement servers.

8.3.3 Theoretical Facet 3: Turnover Consequences

Scholars have examined the consequences of turnover on several dimensions of service effectiveness, including workforce productivity (Armstrong et al., 2010; Arthur, 1994; Bird & Beechler, 1995; Chi & Wang, 2009; Chow, Huang, & Liu, 2008; Cooil, Aksoy, Keiningham, & Maryott, 2009; Detert, Treviño, Burris, & Andiappan, 2007; Donoghue, 2010), quality and safety (Leveck & Jones, 1996; Shaw, Gupta, & Delery, 2005; Shortell et al., 1994; van der Vegt, Bunderson, & Kuipers, 2010), customer satisfaction (Batt & Colvin, 2011; Cooil et al., 2009; McElroy et al., 2001; Mohr, Young, & Burgess Jr, 2012; Ployhart, Van Iddekinge, & MacKenzie, 2011; Ton & Huckman, 2008; Van Jaarsveld & Yanadori, 2011), and financial performance (Angle & Perry, 1981; Bingley & Westergaard-Nielsen, 2004; Bird & Beechler, 1995; Cannella & Hambrick, 1993; Chadwick, Hunter, & Walston, 2004; Chow et al., 2008; Detert et al., 2007). However, there has been little consensus in the literature regarding the shape of the relationship between turnover and service effectiveness outcomes (Shaw, 2011).

First, social and human capital theories suggest a negative linear relationship between turnover and organizational performance (Shaw, Gupta, & Delery, 2005). Social capital theory focuses on the cost of depletion of social capital— "a resource reflecting the character of social relations within the organization, realized through members' levels of collective goal orientation and shared trust" (Leana & Van Buren, 1999, p. 540)— that is driven by turnover events (Park & Shaw, 2013). Therefore, from a social capital perspective, turnover events could lead to extra socialization costs for newcomers and disrupt the social fabric of an organization (Dess & Shaw, 2001; Shaw, Duffy, Johnson, & Lockhart, 2005). In contrast, human capital theory focuses on the costs of depletion of accumulated human capital knowledge and skills that are required for performing the job that would stem from turnover events (Osterman, 1987; Strober, 1990). The core prediction of the human capital theory is built upon the premise that organizations will endure costs to recruit, select, and hire replacement employees who will take time to build human capital levels similar to those of the employees who left the organization. In general, results in the turnover literature support the predictions of the social and human capital theories, where many empirical studies found a negative linear relationship between turnover and organizational performance (e.g., Baron, Hannan, & Burton, 2001; Batt, 2002; Beadles, Lowery, Petty, & Ezell, 2000; Cannella & Hambrick, 1993; Dolton & Newson, 2003; Hausknecht, Trevor, & Howard, 2009; Huselid, 1995; Kacmar, Andrews, Van Rooy, Steilberg, & Cerrone, 2006; McElroy et al., 2001; Morrow & McElroy, 2007; Paul & Anantharaman, 2003). However, the majority of those studies examined either total or voluntary turnover rates, which limits their generalizability because we cannot be sure whether the results will hold true for other types of turnover.

Second, organizational learning and control theories suggest a negative curvilinear relationship between turnover and organizational performance that attenuates under higher levels of turnover. This prediction is built on the premise that organizations with low turnover rates accumulate human capital over a long period of time; therefore, when turnover rates change from low to moderate levels, then human capital will be depleted, and it will take both time and financial resources to replenish the lost human capital to equivalent levels via the new hires. In contrast, when the organization is characterized by high levels of turnover rates, then it will keep losing individuals with lower levels of human capital, which means the new hires can accumulate similar levels of human capital within a short period of time. Hence, organizations face disruption only when the turnover rates change from low to moderate, but not when turnover rates are already high (e.g., Shaw, Duffy, et al., 2005; Shaw, Gupta, & Delery, 2005). Some studies in the literature provide support for the predictions of the organizational learning and control theories (e.g., Alexander, Bloom, & Nuchols, 1994; Shaw, Gupta, & Delery, 2005; Ton & Huckman, 2008). However, these studies looked only at either total or voluntary turnover, which limits the generalizability of the results, as previously discussed.

Third, cost-benefit theories predict an inverted U-shaped curvilinear relationship between turnover and organizational outcomes where turnover is considered beneficial at lower levels and harmful at higher levels (e.g., Abelson & Baysinger, 1984; Dalton & Todor, 1979; Staw, 1980). These predictions are based on the assertion that organizations benefit from low levels of turnover, since it allows them to reduce compensation costs (e.g., base pay, insurance premiums, vacation, etc.), revitalize their work force, and discharge their poor performers (Abelson & Baysinger, 1984; Alexander et al., 1994; Dalton & Todor, 1979). Therefore, this approach predicts that there exists an optimal level of turnover rates where the benefits of turnover overweigh its costs, but that higher turnover rates beyond that level would be disruptive to organizational performance. Empirical tests of this approach are few and show some support for the predictions (Glebbeek & Bax, 2004; Ilmakunnas, Maliranta, & Vainiomäki, 2005; Meier & Hicklin, 2007; Siebert & Zubanov, 2009). However, as with majority of the empirical studies examining the turnover-performance link, most of the turnover measures are for total turnover rates. Hence, one cannot generalize the results to other types of turnover.

In summary, there are three main theoretical predictions about the relationship between turnover and organizational performance: (1) turnover always disrupts organizational performance at all levels of turnover rates and in a linear fashion, (2) turnover disrupts organizational performance in a curvilinear fashion when turnover rates change from low to moderate levels, but the negative effects of turnover are attenuated for higher turnover rates, and (3) turnover rates influence organizational performance in an inverted U-shaped curvilinear fashion where low (high) turnover rates are beneficial (disruptive) to a firm's performance, respectively. Thus, it is important to consider the underlying assumptions of these theories when examining the consequences of turnover in service queuing systems. For example, are disruptions to the social fabric of the work force more of a concern for certain types of queuing systems (e.g., dedicated vs. pooled queuing systems)? Are concerns regarding the depletion of accumulated human capital more relevant to certain types of services (e.g., healthcare vs. call center services)? These questions, among others, need to be considered when examining turnover consequences in a specific service queuing context.

8.4 Conclusion

Research studies in the OM literature have examined the impact of system-level factors on service effectiveness outcomes at the individual and team levels of analysis. Many of these studies recognize the existence of differences among service providers. However, a majority of the studies do not examine the underlying psychological forces that drive these differences. In addition, many OM studies assume that servers react to system-level factors in a uniform fashion. To address these issues, I proposed three recommendations for future research on service effectiveness in service queuing systems that consider the roles of individual-level differences (e.g., motivation- and stress-related differences), team-level differences (e.g., differences in team diversity, psychological safety, or team processes), and organizational withdrawal behavior (e.g., lateness, absenteeism, and turnover) when examining individual- and team-level service effectiveness outcomes.

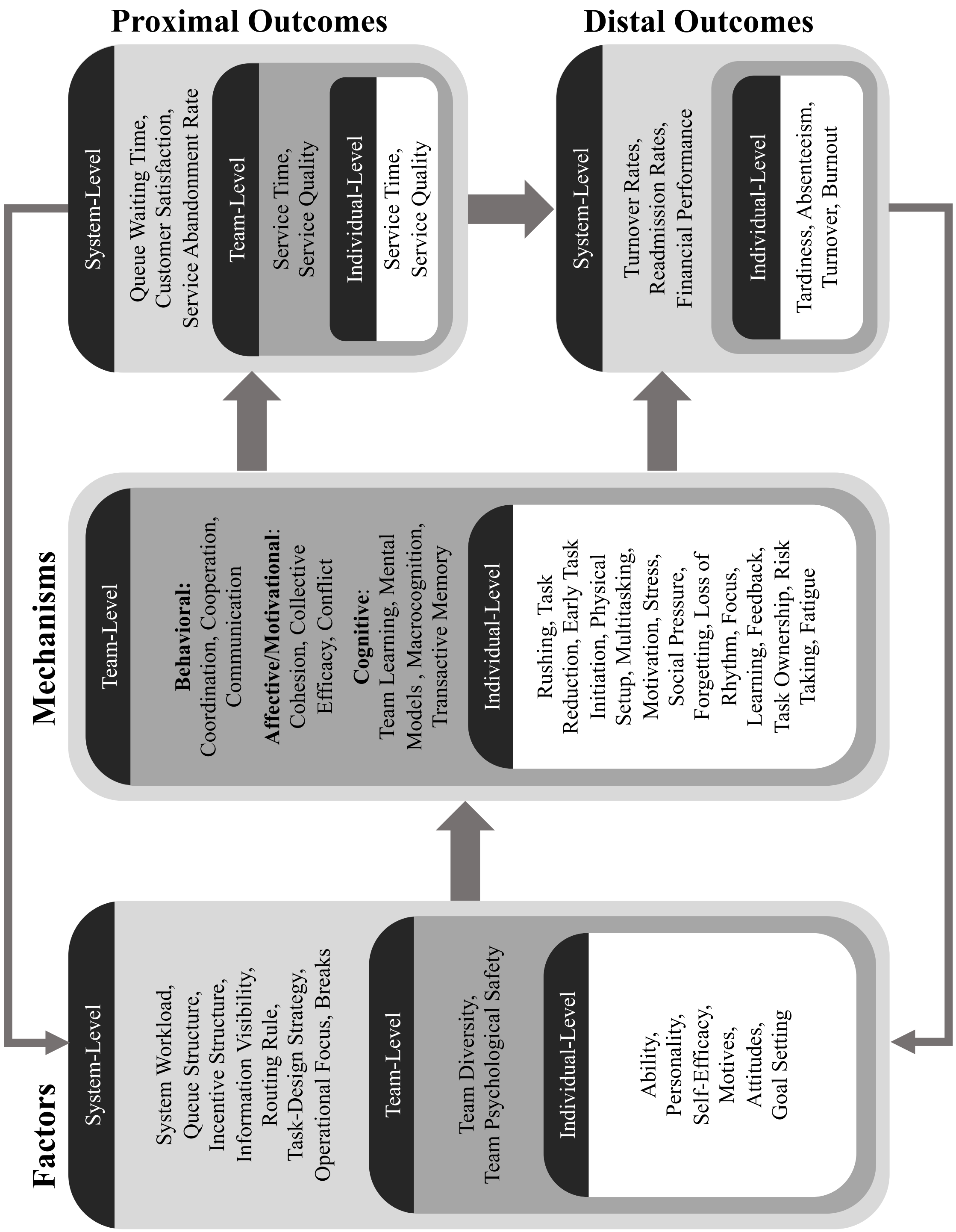

Figure 8.2. A Multilevel Integrative Framework of Service Effectiveness in Queuing Systems

REFERENCES

Abelson, M. A., & Baysinger, B. D. (1984). Optimal and dysfunctional turnover: Toward an organizational level model. *Academy of management Review*, *9*(2), 331–341.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.

Alexander, J. A., Bloom, J. R., & Nuchols, B. A. (1994). Nursing turnover and hospital efficiency: an organizationlevel analysis. *Industrial Relations: A Journal of Economy and Society*, *33*(4), 505–520.

Amabile, T. M. (1988). A model of creativity and innovation in organizations. *Research in organizational behavior*, *10*(1), 123–167.

Amason, A. C. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams. *Academy of management journal*, *39*(1), 123–148.

Angle, H. L., & Perry, J. L. (1981). An empirical assessment of organizational commitment and organizational effectiveness. *Administrative science quarterly*, 1–14.

Armstrong, C., Flood, P. C., Guthrie, J. P., Liu, W., MacCurtain, S., & Mkamwa, T. (2010). The impact of diversity and equality management on firm performance: Beyond high performance work systems. *Human Resource Management*, *49*(6), 977–998.

Arthur, J. B. (1994). Effects of human resource systems on manufacturing performance and turnover. *Academy of Management journal*, *37*(3), 670–687.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American psychologist*, *37*(2), 122.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive perspective. *Englewood Cliffs, NJ: Princeton-Hall*.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Macmillan.

Baron, J. N., Hannan, M. T., & Burton, M. D. (2001). Labor Pains: Change in Organizational Models and Employee Turnover in Young, High-Tech Firms 1. *American journal of sociology*, *106*(4), 960–1012.

Barsade, S. G., & Gibson, D. E. (1998). Group emotion: A view from top and bottom. In *Composition: Research on managing groups and teams* (Vol. 1, pp. 81–102). Stanford, CT: JAI Press.

Batt, R. (2002). Managing customer services: Human resource practices, quit rates, and sales growth. *Academy of management Journal*, *45*(3), 587–597.

Batt, R., & Colvin, A. J. S. (2011). An employment systems approach to turnover: Human resources practices, quits, dismissals, and performance. *Academy of management Journal*, *54*(4), 695–717.

Batt, R., & Terwiesch, C. (2016). Early Task Initiation and Other Load-adaptive Mechanisms in the Emergency Department. *Management Sci. Forthcoming*.

Beadles, N. A., Lowery, C. M., Petty, M. M., & Ezell, H. (2000). An examination of the relationships between turnover functionality, turnover frequency, and organizational performance. *Journal of Business and Psychology*, *15*(2), 331–337.

Beal, D. J., Cohen, R. R., Burke, M. J., & McLendon, C. L. (2003). *Cohesion and performance in groups: a meta-analytic clarification of construct relations.* American Psychological Association.

Berry Jaeker, J. A., & Tucker, A. L. (2017). Past the Point of Speeding Up: The Negative Effects of Workload Saturation on Efficiency and Patient Severity. *Management Science*, *63*(4), 1042–1062.

Berscheid, E. (1985). Interpersonal attraction. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 2, pp. 413–484). New York: Random House.

Bingley, P., & Westergaard-Nielsen, N. (2004). Personnel policy and profit. *Journal of Business Research*, *57*(5), 557–563.

Bird, A., & Beechler, S. (1995). Links between business strategy and human resource management strategy in US-based Japanese subsidiaries: An empirical investigation. *Journal of international business studies*, 23–46.

Borman, W., & Motowidlo, S. (1993). Expanding the criterion domain to include elements of contextual performance. *Personnel selection in organizations; San Francisco: Jossey-Bass.*

Boudreau, J., Hopp, W., McClain, J. O., & Thomas, L. J. (2003). On the Interface Between Operations and Human Resources Management. *Manufacturing & Service Operations Management*, *5*(3), 179–202.

Bradley, B. H., Postlethwaite, B. E., Klotz, A. C., Hamdani, M. R., & Brown, K. G. (2012). Reaping the benefits of task conflict in teams: the critical role of team psychological safety climate. *Journal of Applied Psychology*, *97*(1), 151.

Burke, M. E. (2005). *2005 Benefits: Survey Report.* Society for Human Resource Management.

Byrne, J. A. (1993). The horizontal corporation. *Business Week*, *20*(1993), 76–81.

Campbell, J. P. (1990). Modeling the Performance Prediction Problem in Industrial and Organizational Psychology. In M. D. Dunnette & Leaetta M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed.). Consulting Psychologists Press.

Campion, M. A. (1991). Meaning and measurement of turnover: Comparison of alternative measures and recommendations for research. *Journal of applied psychology*, *76*(2), 199.

Campion, M. A., Medsker, G. J., & Higgs, A. C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel psychology*, *46*(4), 823–847.

Cannella, A. A., & Hambrick, D. C. (1993). Effects of executive departures on the performance of acquired firms. *Strategic Management Journal*, *14*(S1), 137–152.

Chadwick, C., Hunter, L. W., & Walston, S. L. (2004). Effects of downsizing practices on the performance of hospitals. *Strategic Management Journal*, *25*(5), 405–427.

Chatman, J. A., Polzer, J. T., Barsade, S. G., & Neale, M. A. (1998). Being different yet feeling similar: The influence of demographic composition and organizational culture on work processes and outcomes. *Administrative Science Quarterly*, 749–780.

Chi, W., & Wang, Y. (2009). Ownership, performance and executive turnover in China. *Journal of Asian Economics*, *20*(4), 465–478.

Chow, I. H., Huang, J., & Liu, S. (2008). Strategic HRM in China: Configurations and competitive advantage. *Human Resource Management*, *47*(4), 687–706.

Cole, M. S., Walter, F., & Bruch, H. (2008). Affective mechanisms linking dysfunctional behavior to performance in work teams: a moderated mediation study. *Journal of Applied Psychology*, *93*(5), 945.

Colquitt, J. A., LePine, J. A., & Wesson, M. J. (2013). *Organizational behavior: Improving Performance and Commitment in the Workplace.* McGraw-Hill Irwin.

Cooil, B., Aksoy, L., Keiningham, T. L., & Maryott, K. M. (2009). The relationship of employee perceptions of organizational climate to business-unit outcomes an MPLS approach. *Journal of Service Research*, *11*(3), 277–294.

Cortina, J. M., & Luchman, J. N. (2012). Personnel Selection and Employee Performance. In N. W. Schmitt & S. Highhouse (Eds.), *Handbook of psychology* (2nd ed., Vol. 12, pp. 143–183). Hoboken, NJ, USA: John Wiley & Sons, Inc.

Cox, T. H., Lobel, S. A., & McLeod, P. L. (1991). Effects of ethnic group cultural differences on cooperative and competitive behavior on a group task. *Academy of management journal*, *34*(4), 827–847.

Dalton, D. R., & Todor, W. D. (1979). Turnover turned over: An expanded and positive perspective. *Academy of Management Review*, *4*(2), 225–235.

Damon, W. (1991). Problems of direction in socially shared cognition. In L. Resnick, J. Levine, & S. Teasley (Eds.), *Perspectives on socially shared cognition.* Washington, DC: APA Books.

De Dreu, C. K. W., & Weingart, L. R. (2003). *Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis.* American Psychological Association.

Delasay, M., Ingolfsson, A., Kolfal, B., & Schultz, K. L. (2015). *Load Effect on Service Times.*

DeShon, R. P., Kozlowski, S. W. J., Schmidt, A. M., Milner, K. R., & Wiechmann, D. (2004). A multiple-goal, multilevel model of feedback effects on the regulation of individual and team performance. *Journal of applied psychology*, *89*(6), 1035.

Dess, G. G., & Shaw, J. D. (2001). Voluntary turnover, social capital, and organizational performance. *Academy of management review*, *26*(3), 446–456.

Detert, J. R., Treviño, L. K., Burris, E. R., & Andiappan, M. (2007). Managerial modes of influence and counterproductivity in organizations: A longitudinal business-unit-level investigation. *Journal of Applied Psychology*, *92*(4), 993.

Dolton, P., & Newson, D. (2003). The relationship between teacher turnover and school performance. *London Review of Education*, *1*(2), 131–140.

Donnellon, A. (1996). *Team talk: The power of language in team dynamics.* Island Press.

Donoghue, C. (2010). Nursing home staff turnover and retention: An analysis of national level data. *Journal of Applied Gerontology*, *29*(1), 89–106.

Duguid, M. M., & Thomas-Hunt, M. C. (2015). Condoning stereotyping? How awareness of stereotyping prevalence impacts expression of stereotypes. *Journal of Applied Psychology*, *100*(2), 343.

Edmondson, A. C. (1996). Learning from mistakes is easier said than done: Group and organizational influences on the detection and correction of human error. *The Journal of Applied Behavioral Science*, *32*(1), 5–28.

Edmondson, A. C. (1999). Psychological Safety and Learning Behavior in Work Teams. *Administrative Science Quarterly*, *44*(2), 350.

Edmondson, A. C., & Lei, Z. (2014). Psychological safety: The history, renaissance, and future of an interpersonal construct. *Annu. Rev. Organ. Psychol. Organ. Behav.*, *1*(1), 23–43.

Evans, C. R., & Jarvis, P. A. (1980). Group cohesion: A review and re-evaluation. In *Small group behavior* (Vol. 11, pp. 359–370).

Feltz, D. L., & Lirgg, C. D. (1998). Perceived team and player efficacy in hockey. *Journal of applied psychology*, *83*(4), 557.

Garson, G. D. (2012). *Hierarchical Linear Modeling: Guide and Applications.* Sage.

George, J. M. (2007). Creativity in Organizations. *The academy of management annals*, *1*(1), 439–477.

George, J. M., & Harrison, J. M. (2001). Dynamic Control of a Queue with Adjustable Service Rate. *Operations Research*, *49*(5), 720–731.

Gibson, C. B., & Earley, P. C. (2007). Collective cognition in action: Accumulation, interaction, examination, and accommodation in the development and operation of group efficacy beliefs in the workplace. *Academy of management review*, *32*(2), 438–458.

Glebbeek, A. C., & Bax, E. H. (2004). Is high employee turnover really harmful? An empirical test using company records. *Academy of Management Journal*, *47*(2), 277–286.

Goodman, P. S., Ravlin, E., & Schminke, M. (1987). Understanding groups in organizations. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior1* (Vol. 9, pp. 121–173). Greenwich, CT: JAI Press.

Grant, A. M. (2008). Does intrinsic motivation fuel the prosocial fire? Motivational synergy in predicting persistence, performance, and productivity. *Journal of Applied Psychology*, *93*(1), 48–58.

Grant, A. M., & Berry, J. W. (2011). The Necessity of Others is The Mother of Invention: Intrinsic and Prosocial Motivations, Perspective Taking, and Creativity. *Academy of Management Journal*, *54*(1), 73–96.

Grawitch, M. J., Munz, D. C., & Kramer, T. J. (2003). Effects of member mood states on creative performance in temporary workgroups. *Group Dynamics: Theory, Research, and Practice*, *7*(1), 41.

Gross, N., & Martin, W. E. (1952). On group cohesiveness. *American Journal of Sociology*, *57*(6), 546–564.

Gully, S. M., Incalcaterra, K. A., Joshi, A., & Beaubien, J. M. (2002). A meta-analysis of team-efficacy, potency, and performance: interdependence and level of analysis as moderators of observed relationships. *Journal of applied psychology*, *87*(5), 819.

Halbesleben, J. R. B. (2010). The role of exhaustion and workarounds in predicting occupational injuries: a cross-lagged panel study of health care professionals. *Journal of occupational health psychology*, *15*(1), 1.

Halbesleben, J. R. B., & Bowler, W. M. (2007). Emotional exhaustion and job performance: the mediating role of motivation. *Journal of applied psychology*, *92*(1), 93.

Halbesleben, J. R. B., Harvey, J., & Bolino, M. C. (2009). Too engaged? A conservation of resources view of the relationship between work engagement and work interference with family. *Journal of Applied Psychology*, *94*(6), 1452–1465.

Halbesleben, J. R. B., Neveu, J.-P., Paustian-Underdahl, S. C., & Westman, M. (2014). Getting to the COR. *Journal of Management*, *40*(5), 1334–1364.

Halbesleben, J. R. B., & Wheeler, A. R. (2008). The relative roles of engagement and embeddedness in predicting job performance and intention to leave. *Work & Stress*, *22*(3), 242–256.

Halbesleben, J. R. B., & Wheeler, A. R. (2011). I owe you one: Coworker reciprocity as a moderator of the daylevel exhaustionperformance relationship. *Journal of Organizational Behavior*, *32*(4), 608–626.

Hancock, J. I., Allen, D. G., Bosco, F. A., McDaniel, K. R., & Pierce, C. A. (2013). Meta-analytic review of employee turnover as a predictor of firm performance. *Journal of Management*, *39*(3), 573–603.

Harrison, D. A., & Newman, D. A. (2012). Absence, Lateness, Turnover, And Retirement: Narrow and Broad Understandings of Withdrawal and Behavioral

Engagement. In N. W. Schmitt & S. Highhouse (Eds.), *Handbook of psychology* (2nd ed., Vol. 12, pp. 262–291). Hoboken, NJ: John Wiley & Sons, Inc.

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface-and deep-level diversity on group functioning. *Academy of management journal*, *45*(5), 1029–1045.

Hausknecht, J. P., Trevor, C. O., & Howard, M. J. (2009). Unit-level voluntary turnover rates and customer service quality: implications of group cohesiveness, newcomer concentration, and size. *Journal of Applied Psychology*, *94*(4), 1068.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hecker, D. E. (2004). Occupational Employment Projections to 2012. *Monthly Labor Review*, *127*.

Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological bulletin*, *121*(1), 43.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44*(3), 513–524.

Hodges, L., & Carron, A. V. (1992). Collective efficacy and group performance. *International Journal of Sport Psychology*.

Hoffman, L. R. (1959). Homogeneity of member personality and its effect on group problem-solving. *The Journal of Abnormal and Social Psychology*, *58*(1), 27.

Homan, A. C., Van Knippenberg, D., Van Kleef, G. A., & De Dreu, C. K. W. (2007). Bridging faultlines by valuing diversity: diversity beliefs, information elaboration, and performance in diverse work groups. *Journal of Applied Psychology*, *92*(5), 1189.

Huang, C. C., Chu, C. Y., & Jiang, P. C. (2008). An empirical study of psychological safety and performance in technology R&D teams. In *Management of innovation and technology, 2008. icmit 2008. 4th ieee international conference on* (pp. 1423–1427). IEEE.

Huckman, R. S., & Staats, B. R. (2011). Fluid Tasks and Fluid Teams: The Impact of Diversity in Experience and Team Familiarity on Team Performance. *Manufacturing & Service Operations Management*, *13*(3), 310–328.

Huselid, M. A. (1995). The Impact Of Human Resource Management Practices On Turnover, Productivity, And Corporate Financial Performance. *Academy of Management Journal*, *38*(3), 635–672.

Ilgen, D. R. (1999). Teams embedded in organizations: Some implications. *American Psychologist*, *54*(2), 129.

Ilmakunnas, P., Maliranta, M., & Vainiomäki, J. (2005). Worker turnover and productivity growth. *Applied Economics Letters*, *12*(7), 395–398.

Jackson, S. (1992). Team composition in organizations. In S. Worchel, W. Wood, & J. Simpson (Eds.), *Group process and productivity* (pp. 138–176). London: Sage.

Janis, I. L. (1982). *Victims of groupthink* (2nd ed.). Boston: Houghton-Mifflin.

Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative science quarterly*, 256–282.

Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict and performance in workgroups. *Administrative science quarterly*, *44*(4), 741–763.

Joshi, A., Neely, B., Emrich, C., Griffiths, D., & George, G. (2015). Gender research in AMJ: An overview of five decades of empirical research and calls to action thematic issue on gender in management research. *Academy of Management Journal*, *58*(5), 1459–1475.

Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: A meta-analytic review. *Academy of Management Journal*, *52*(3), 599–627.

Kacmar, K. M., Andrews, M. C., Van Rooy, D. L., Steilberg, R. C., & Cerrone, S. (2006). Sure everyone can be replaced... but at what cost? Turnover as a predictor of unit-level performance. *Academy of Management journal*, *49*(1), 133–144.

Karau, S. J., & Williams, K. D. (1993). *Social loafing: A meta-analytic review and theoretical integration.* American Psychological Association.

KC, D. S. (2014). Does Multitasking Improve Performance? Evidence from the Emergency Department. *Manufacturing & Service Operations Management*, *16*(2), 168–183.

KC, D. S., Staats, B. R., & Gino, F. (2013). Learning from my success and from others' failure: Evidence from minimally invasive cardiac surgery. *Management Science*, *59*(11), 2435–2449.

Kc, D. S., & Terwiesch, C. (2009). Impact of Workload on Service Time and Patient Safety: An Econometric Analysis of Hospital Operations. *Management Science*, *55*(9), 1486–1498.

KC, D. S., & Terwiesch, C. (2011). The Effects of Focus on Performance: Evidence from California Hospitals. *Management Science*, *57*(11), 1897–1912.

KC, D. S., & Terwiesch, C. (2012). An Econometric Analysis of Patient Flows in the Cardiac Intensive Care Unit. *Manufacturing & Service Operations Management*, *14*(1), 50–65.

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and social Psychology*, *44*(1), 78–94.

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of management*, *20*(2), 403–437.

Kozlowski, S. W. J., & Bell, B. S. (2012). Work Groups and Teams in Organizations. In N. W. Schmitt & S. Highhouse (Eds.), *Handbook of psychology* (2nd ed., Vol. 12, pp. 412–469). Hoboken, NJ: John Wiley & Sons, Inc.

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological science in the public interest*, *7*(3), 77–124.

Kraiger, K., & Wenzel, L. H. (1997). Conceptual development and empirical evaluation of measures of shared mental models as indicators of team effectiveness. *Team performance assessment and measurement: Theory, methods, and applications*, *63*, 84.

Kulik, C. T., Roberson, L., & Perry, E. L. (2007). The multiple-category problem: Category activation and inhibition in the hiring process. *Academy of Management Review*, *32*(2), 529–548.

Kuntz, L., Mennicken, R., & Scholtes, S. (2015). Stress on the Ward: Evidence of Safety Tipping Points in Hospitals. *Management Science*, *61*(4), 754–771.

Latané, 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*(6), 822.

Lazarus, R. S., & Folkman, S. (1984). *Stress, Appraisal, and Coping.* Springer Publishing Company, New York.

Leana, C. R., & Van Buren, H. J. (1999). Organizational social capital and employment practices. *Academy of management review*, *24*(3), 538–555.

LePine, J. A., Colquitt, J. A., & Erez, A. (2000). Adaptability To Changing Task Contexts: Effects Of General Cognitive Ability, Conscientiousness, And Openness To Experience. *Personnel Psychology*, *53*(3), 563–593.

Leveck, M. L., & Jones, C. B. (1996). The nursing practice environment, staff retention, and quality of care. *Research in nursing & health*, *19*(4), 331–343.

Levine, J. M., & Choi, H.-S. (2004). Impact of personnel turnover on team performance and cognition.

Levine, J. M., Resnick, L. B., & Higgins, E. T. (1993). Social foundations of cognition. *Annual review of psychology*, *44*(1), 585–612.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance.* Prentice-Hall, Inc.

Locke, E. A., & Latham, G. P. (2004). What Should we do About Motivation Theory? Six Recommendations for the Twenty-First Century. *Academy of Management Review*, *29*(3), 388–403.

Loyd, D. L., Wang, C. S., Phillips, K. W., & Lount Jr, R. B. (2013). Social category diversity promotes premeeting elaboration: The role of relationship focus. *Organization Science*, *24*(3), 757–772.

Maertz, C. P., & Campion, M. A. (2004). Profiles in quitting: Integrating process and content turnover theory. *Academy of Management Journal*, *47*(4), 566–582.

Mannix, E., & Neale, M. A. (2005). What differences make a difference? The promise and reality of diverse teams in organizations. *Psychological science in the public interest*, *6*(2), 31–55.

Marks, M. A., Sabella, M. J., Burke, C. S., & Zaccaro, S. J. (2002). The impact of cross-training on team effectiveness. *Journal of Applied Psychology*, *87*(1), 3.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual review of psychology*, *52*(1), 397–422.

Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of management*, *34*(3), 410–476.

McCarthy, M. L., Ding, R., Pines, J. M., Terwiesch, C., Sattarian, M., Hilton, J. A., . . . Zeger, S. L. (2012). Provider Variation in Fast Track Treatment Time. *Medical Care*, *50*(1), 43–49.

McElroy, J. C., Morrow, P. C., & Rude, S. N. (2001). Turnover and organizational performance: a comparative analysis of the effects of voluntary, involuntary, and reduction-in-force turnover. *Journal of applied Psychology*, *86*(6), 1294.

Meier, K. J., & Hicklin, A. (2007). Employee turnover and organizational performance: Testing a hypothesis from classical public administration. *Journal of Public Administration Research and Theory*, *18*(4), 573–590.

Milliken, F. J., & Martins, L. L. (1996). Searching for common threads: Understanding the multiple effects of diversity in organizational groups. *Academy of management review*, *21*(2), 402–433.

Mitchell, T. R., & Daniels, D. (2003). Motivation. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology* (1st ed., Vol. 12, pp. 225–254). Hoboken, NJ: John Wiley & Sons, Inc.

Mohr, D. C., Young, G. J., & Burgess Jr, J. F. (2012). Employee turnover and operational performance: The moderating effect of grouporiented organisational culture. *Human Resource Management Journal*, *22*(2), 216–233.

Morrow, P., & McElroy, J. (2007). Efficiency as a mediator in turnoverorganizational performance relations. *Human Relations*, *60*(6), 827–849.

Mullen, B., & Copper, C. (1994). *The relation between group cohesiveness and performance: An integration.* American Psychological Association.

Nembhard, I. M., & Edmondson, A. C. (2006). Making it safe: The effects of leader inclusiveness and professional status on psychological safety and improvement efforts in health care teams. *Journal of Organizational Behavior*, *27*(7), 941–966.

Nembhard, I. M., & Tucker, A. L. (2011). Deliberate learning to improve performance in dynamic service settings: Evidence from hospital intensive care units. *Organization Science*, *22*(4), 907–922.

Nemeth, C. J. (1986). Differential contributions of majority and minority influence. *Psychological review*, *93*(1), 23.

Ng, T. W. H., & Feldman, D. C. (2012). Employee voice behavior: A meta-analytic test of the conservation of resources framework. *Journal of Organizational Behavior*, *33*(2), 216–234.

O*NET. (2017). *National Center for O*NET Development. 43-4051.00.* Retrieved 2017-06-05, from `https://www.onetonline.org/link/summary/43-4051.00`

Osterman, P. (1987). Turnover, employment security, and the performance of the firm. *Human resources and the performance of the firm*, *275*, 317.

Park, T.-Y., & Shaw, J. D. (2013). *Turnover rates and organizational performance: a meta-analysis.* American Psychological Association.

Paul, A. K., & Anantharaman, R. N. (2003). Impact of people management practices on organizational performance: analysis of a causal model. *International journal of human resource management*, *14*(7), 1246–1266.

Pendem, P., Staats, B. R., Green, P., & Gino, F. (2016). *The Microstructure of Work: How Unexpected Breaks Let You Rest, but Not Lose Focus.*

Phillips, K. W., Liljenquist, K. A., & Neale, M. A. (2009). Is the pain worth the gain? The advantages and liabilities of agreeing with socially distinct newcomers. *Personality and Social Psychology Bulletin*, *35*(3), 336–350.

Pinto, M. B., & Pinto, J. K. (1990). Project team communication and cross-functional cooperation in new program development. *Journal of product innovation management*, *7*(3), 200–212.

Ployhart, R. E., Van Iddekinge, C. H., & MacKenzie, W. I. (2011). Acquiring and developing human capital in service contexts: The interconnectedness of human capital resources. *Academy of management journal*, *54*(2), 353–368.

Probst, T. M., & Estrada, A. X. (2010). Accident under-reporting among employees: Testing the moderating influence of psychological safety climate and supervisor enforcement of safety practices. *Accident Analysis & Prevention*, *42*(5), 1438–1444.

Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology*, *85*(4), 612–624.

Rico, R., Sánchez-Manzanares, M., Gil, F., & Gibson, C. (2008). Team implicit coordination processes: A team knowledgebased approach. *Academy of Management Review*, *33*(1), 163–184.

Rosenbaum, M. E. (1986). The repulsion hypothesis: On the nondevelopment of relationships. *Journal of Personality and Social Psychology*, *51*(6), 1156.

Salas, E., Stagl, K. C., & Burke, C. S. (2004). 25 years of team effectiveness in organizations: research themes and emerging needs. *International review of industrial and organizational psychology*, *19*, 47–92.

Schultz, K. L., Juran, D. C., Boudreau, J. W., McClain, J. O., & Thomas, L. J. (1998). Modeling and Worker Motivation in JIT Production Systems. *Management Science*, *44*(12-part-1), 1595–1607.

Schultz, K. L., McClain, J. O., & Thomas, L. J. (2003). Overcoming the dark side of worker flexibility. *Journal of Operations Management*, *21*(1), 81–92.

Shaw, J. D. (2011). Turnover rates and organizational performance: Review, critique, and research agenda. *Organizational Psychology Review*, *1*(3), 187–213.

Shaw, J. D., Duffy, M. K., Johnson, J. L., & Lockhart, D. E. (2005). Turnover, social capital losses, and performance. *Academy of Management Journal*, *48*(4), 594–606.

Shaw, J. D., Gupta, N., & Delery, J. E. (2005). Alternative conceptualizations of the relationship between voluntary turnover and organizational performance. *Academy of management journal*, *48*(1), 50–68.

Shortell, S. M., Zimmerman, J. E., Rousseau, D. M., Gillies, R. R., Wagner, D. P., Draper, E. A., ... Duffy, J. (1994). The performance of intensive care units: does good management make a difference? *Medical care*, 508–525.

Shunko, M., Niederhoff, J., & Rosokha, Y. (2017). Humans Are Not Machines: The Behavioral Impact of Queueing Design on Service Time. *Management Science*, mnsc.2016.2610.

Siebert, W. S., & Zubanov, N. (2009). Searching for the optimal level of employee turnover: A study of a large UK retail organization. *Academy of Management Journal*, *52*(2), 294–313.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press.

Smith, K. G., Smith, K. A., Olian, J. D., Sims Jr, H. P., O'Bannon, D. P., & Scully, J. A. (1994). Top management team demography and process: The role of social integration and communication. *Administrative science quarterly*, 412–438.

Sommers, S. R. (2006). On racial diversity and group decision making: identifying multiple effects of racial composition on jury deliberations. *Journal of personality and social psychology*, *90*(4), 597.

Song, H., Tucker, A. L., & Murrell, K. L. (2015). The Diseconomies of Queue Pooling: An Empirical Investigation of Emergency Department Length of Stay. *Management Science*, *61*(12), 3032–3053.

Sonnentag, S., & Frese, M. (2012). Stress in Organizations. In N. W. Schmitt & S. Highhouse (Eds.), *Handbook of psychology* (2nd ed., Vol. 12, pp. 560–592). Hoboken, NJ: John Wiley & Sons, Inc.

Staats, B. R., & Gino, F. (2012). Specialization and Variety in Repetitive Tasks: Evidence from a Japanese Bank. *Management Science*, *58*(6), 1141–1159.

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological bulletin*, *124*(2), 240.

Staw, B. M. (1980). The consequences of turnover. *Journal of occupational Behaviour*, 253–273.

Stidham, S., & Weber, R. R. (1989). Monotonic and Insensitive Optimal Policies for Control of Queues with Undiscounted Costs. *Operations Research*, *37*(4), 611–625.

Stout, R. J., Salas, E., & Carson, R. (1994). Individual task proficiency and team process behavior: What's important for team functioning? *Military psychology*, *6*(3), 177.

Strober, M. H. (1990). Human capital theory: Implications for HR managers. *Industrial Relations*, *29*(2), 214–239.

Tan, T. F., & Netessine, S. (2014). When Does the Devil Make Work? An Empirical Study of the Impact of Workload on Worker Productivity. *Management Science*, *60*(6), 1574–1593.

Tan, T. F., & Netessine, S. (2015). *When you work with a Super Man, will you also fly? An empirical study of the impact of the coworkers on workers' performance.*

Ton, Z., & Huckman, R. S. (2008). Managing the impact of employee turnover on performance: The role of process conformance. *Organization Science*, *19*(1), 56–68.

Tubbs, M. (1994). Commitment and the role of ability in motivation: Comment on Wright, O'Leary-Kelly, Cortina, Klein, and Hollenbeck (1994).

Tucker, A. L. (2007). An empirical study of system improvement by frontline employees in hospital units. *Manufacturing & Service Operations Management*, *9*(4), 492–505.

Tucker, A. L., Nembhard, I. M., & Edmondson, A. C. (2007). Implementing new practices: An empirical study of organizational learning in hospital intensive care units. *Management science*, *53*(6), 894–907.

Van Der Vegt, G. S., & Bunderson, J. S. (2005). Learning and performance in multidisciplinary teams: The importance of collective team identification. *Academy of Management journal*, *48*(3), 532–547.

Van Jaarsveld, D. D., & Yanadori, Y. (2011). Compensation management in outsourced service organizations and its implications for quit rates, absenteeism and workforce performance: Evidence from Canadian call centres. *British Journal of Industrial Relations*, *49*(s1), s1–s26.

van der Vegt, G. S., Bunderson, S., & Kuipers, B. (2010). Why turnover matters in self-managing work teams: Learning, social integration, and task flexibility. *Journal of Management*, *36*(5), 1168–1191.

Vinokur, A. D., & Schul, Y. (2002). The web of coping resources and pathways to reemployment following a job loss. *Journal of Occupational Health Psychology*, *7*(1), 68–83.

Vroom, V. H. (1964). Work and motivation. 1964. *NY: John Wiley &sons*, *45*.

Wagner, J. A. (1995). Studies of individualism-collectivism: Effects on cooperation in groups. *Academy of Management journal*, *38*(1), 152–173.

Wagner, W. G., Pfeffer, J., & O'Reilly, C. A. (1984). Organizational demography and turnover in top-management group. *Administrative Science Quarterly*, 74–92.

Wong, S.-S. (2004). Distal and local group learning: Performance trade-offs and tensions. *Organization Science*, *15*(6), 645–656.

Wood, R. E., Mento, A. J., & Locke, E. A. (1987). Task complexity as a moderator of goal effects: a meta-analysis. *Journal of applied psychology*, *72*(3), 416.

Zaccaro, S. J., & McCoy, M. C. (1988). The Effects of task and interpersonal cohesiveness on performance of a disjunctive group task1. *Journal of applied social psychology*, *18*(10), 837–851.

Zellmer-Bruhn, M., & Gibson, C. (2006). Multinational organization context: Implications for team learning and performance. *Academy of management journal*, *49*(3), 501–518.

A. ROBUSTNESS CHECK RESULTS

Table A.1.

Joint Effects of Workload and Motivation on Productivity Level (Robustness Check)

	Dependent variable: Call Wrap-Up Time				
	(1)	(2)	(3)	(4)	(5)
Fixed Effects					
Intercept	58.30***	52.32***	55.11***	55.07***	55.11***
	(6.71)	(7.96)	(8.03)	(7.79)	(7.78)
Level-1 Main Effects (within-agent)					
M.C. Workload$_{NA}$	—	—	0.171**	0.172**	0.157**
			(0.079)	(0.079)	(0.070)
M.C. Workload$_{NA}$ Square	—	—	-0.004**	-0.004**	-0.004*
			(0.002)	(0.002)	'(0.002)
Level-2 Main Effects (between-agent)					
M.C. Intrinsic Motivation	—	—	—	-11.21**	-11.31**
				(5.09)	(5.14)
M.C. Extrinsic Motivation	—	—	—	0.96	0.8
				(4.41)	(4.45)
Cross-Level Interaction Effects					
M.C. Intrinsic Motivation × M.C. Workload$_{NA}$	—	—	—	—	-0.15***
					(0.05)
M.C. Extrinsic Motivation × M.C. Workload$_{NA}$	—	—	—	—	0.03
					(0.05)
Random Effects (Variance Components)					
Residual (within-agent)	6363.51***	6154.83***	6140.11***	6140.13***	6140.33***
	(39.27)	(37.98)	(37.92)	(37.92)	(37.92)
Intercept (between-agent)	2554.91***	2682.20***	2645.94***	2434.56***	2422.46***
	(479.75)	(507.42)	(502.28)	(462.29)	(459.83)
M.C. Workload (between-agent)	—	—	0.2192***	0.2179***	0.1702***
			(0.07)	(0.07)	(0. 056)
Controls					
Service Line F.E.	—	Yes	Yes	Yes	Yes
Number of Agents	—	-0.72***	-0.87***	-0.87***	-0. 86***
		(0.22)	(0.22)	(0.22)	(0.22)
Call Duration (On-Line)	—	0.07***	0.07***	0.07***	0.07***
		(0.002)	(0.002)	(0.002)	(0.002)
Time	—	3.31***	3.21***	3.21***	3.20***
		(0.75)	(0.84)	(0.84)	(0.84)
Time Square	—	-0.44***	-0.43***	-0.43***	-0.43***
		(0.08)	(0.09)	(0.09)	(0.09)
Day of Week F.E.	—	Yes	Yes	Yes	Yes
Overwork$_{NA,K=4}$	—	2.32***	0.155**	0.155**	0.157***
		(0.56)	(0.054)	(0.054)	(0.050)
-2 Log Likelihood (Deviance)	609992.1	608243.8	608191.6	608186.9	608179
Akaike Information Criterion	609998.1	608317.8	608271.6	608270.9	608267
ΔD		1748.3***	52.2***	4.7*	7.9**

Notes: N=52,574 calls/57 agents. The subscript NA indicates the workload and overwork measures are not adjusted for the number of agents assigned to service those calls. Standard errors are in parenthesis. F.E.= Fixed Effects. M.C.= Mean Centered. ΔD= delta deviance. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

VITA

Ahmad M. Ashkanani

Ph.D. Candidate

Purdue University

Krannert School of Management

Email: aashkana@purdue.edu

Education

Ph.D. (expected: August, 2017). *Krannert School of Management, Purdue University*, with emphasis on Behavioral Operations Management.

M.B.A. (May, 2012). *Robert H. Smith School of Business, University of Maryland—College Park*, with emphasis on Operations Management and Marketing.

B.S. (June, 2008). *College of Engineering and Petroleum, Kuwait University*, Computer Engineering.

Academic Appointments

August 2010 - Present. Scholarship holder. *College of Business Administration, Kuwait University*.

September 2009 - August 2010. Teaching Assistant. *College of Business Administration, Kuwait University*.

Research Interests

Behavioral Operations, Healthcare Operations, Service Operations Management, OM/OBHR Interface, Motivation, Stress and Burnout.