Performance measurement as a driver for learning and improvement in organizations

Jason Martin (jason.martin@liu.se)
Linköping University

Mattias Elg (mattias.elg@liu.se)
Linköping University

Andreas Wallo (andreas.wallo@liu.se)
Linköping University

Henrik Kock (henrik.kock@liu.se)
Linköping University

Abstract

The purpose of this study is to contribute to better understanding about levels of learning found in the practices of performance measurement. We create a conceptual model merging established workplace learning theory with operations management purposes of performance measurement. By integrating discretion and organizational logics we interconnect purposes of either control or improvement with four levels of workplace learning: Reproductive, rule-oriented, goal-oriented and creative learning. We empirically elaborate our model using case related critical incidents involving performance measurements. The findings suggest that a conceptual understanding of workplace learning theory meliorates the design and practice of performance measurements for improvement purposes.

Keywords: Performance measurement, improvement, learning.

Introduction

There are two overall purposes with performance measurement; control and improvement (e.g. Neely, 1999; Neely et al., 2005, Neely & Al Najjar, 2006). In this study, improvement constitutes a main imperative for performance measurement, with learning as a key issue to consider (Behn, 2003). Would a better understanding of learning in performance measurement support improvement efforts? We believe so since failed improvement efforts can be attributed to learning inability within operations (e.g. Baumard & Starbuck, 2005; Cannon & Edmondson, 2005). In their study, Tucker and Edmondson (2003) claimed that 93% of the solutions to daily performance problems were not addressed, communicated or institutionalized on a higher organizational level. It seems that learning is an essential ability (Bohmer, 2009) and even forms a core habit (Lucas, 2015) in performance improvement. Hence, in order to understand how practices of performance measurement may support improvement efforts we must gain a further understanding of learning processes.
In the influential study by Flynn et al. (1990), operations management research is argued as mainly atheoretical and should be empirically driven. On the other hand, Walker (2015) argues for a need to widen the theoretical base in operations management research. In this study, we follow the latter in our approach of situating performance measurement within a workplace learning framework (Billet, 2001; Ellström, 2011). We argue for a need to further understand the diverse and complex ways of learning and how workplace learning theory should have salience in developing operations management in general and performance measurement practices in particular. The demand for improvement may prescribe suitable learning processes to guide and improve the design of performance measurement methods. The purpose of this study is thus to contribute to a better understanding of the opportunities for various levels of learning found in the practice and purpose of performance measurement. We will extend our investigation in two steps: first, we will develop a conceptual model for learning processes in performance measurement. Second, we will empirically elaborate on the conceptual model by drawing on data from a case study conducted within a “performance measurement rich” context.

**Purposes of performance measurement: Control vs. learning and improvement**

Within the performance measurement literature, the issue of defining and understanding the different purposes and roles of performance measurement in organizations has been a recurring theme (e.g. Bond, 1999; Neely et al., 1999). According to Behn (2003) it is possible to discern at least eight different purposes of performance measurement: *evaluate, control, budget, motivate, promote, celebrate, learn, and improve*. Similar compilations of performance measurement purposes are also found in other studies (e.g. Elg, 2007; Franco-Santos et al., 2007; Pidd, 2012). In this paper, we are particularly interested in the learning and improvement purposes.

Performance measurement vis-à-vis learning and improvement have been addressed in several studies within the operations management field (e.g. Dervitisiotis, 2004; Franco-Santos et al., 2007). There are also studies that have employed theory from other research fields to study performance measurement and learning (e.g. Elg et al., 2013; Bond, 1999, Dossi & Patelli, 2010). However, in applying a learning perspective on performance measurement, questions on *what is learnt* and *by whom* are rarely addressed and in previous research there are seldom attempts to distinguish between different meanings of the concept. Rather, learning in performance measurement research seems to be something of a “black box”, often left without further problematization.

According to Behn (2003), learning and performance measurement is about determining what is working, or not, and then figuring out how to make improvements based on these new insights. However, learning and improvement from performance measurement is by no means easy as performance improvement does not follow automatically from measuring (Bourne, 2008). “Measurement just keeps the score” Bourne (2008, p.68) states. One problem in making performance measurement have impact on operations management is the risk of data overload (Neely, 1999). A second problem is when an abundance of performance measurements creates a risk of confusion, inhibiting learning. A third problem is choosing the right performance measurements aimed at optimal improvement. Learning often is triggered by disruption (Dewey (1933/1998) and the unexpected (Behn, 2003), making it difficult to design the particular performance measurements best facilitating learning and improvement at any given time or context.
A workplace learning theory framework

We will in this paper draw on a workplace learning perspective (Billet, 2001; Ellström, 2011) to discuss learning in relation to performance measurement. As used here, learning implies both formal training activities (e.g. courses) and informal learning. The latter concept refers to learning as a continuous activity that is inherently integrated in work (Ellström, 2011; Zuboff, 1988). Theoretically, this distinction between formal and informal learning is, in some respects, parallel to the distinction between learning as acquisition and learning as participation as proposed by Sfard (1998).

There are many ways of conceptualising different types of learning (cf. Argyris & Schön, 1978; 1996; Engeström, 1987). In this paper we mainly draw on Ellström’s (2010) distinction between the logic of production and the logic of development, which, in turn, is based on March’s (1991) theory of exploitation and exploration. The dominant theory of the logic of performance focuses on routine action obtained through adaptive learning, where rules and instructions are followed for known problems or situations. According to the logic of production, learning is primarily instrumental and is valued when it promotes effective action, leading to a learning environment characterised by security, standardisation, exploitation and consensus. Adaptive learning thus happens on a more routine basis often aiming at improving something already mastered. The logic of development, in contrast, is characterised by reflective and alternative thinking, risk taking, critical reflection and the desire to experiment (Ellström, 2010; 2011). Uncertainty and divergence are seen as potential generators of exploration in developmental learning rather than as threats or inconveniences. According to Ellström (2011), these two logics should be perceived as co-existing and complementary.

Whether learning is adaptive or developmental requires a discussion on the conditions in the workplace that may enable or constrain learning, i.e. the learning environment (Coetzer, 2007; Fuller & Unwin, 2004). These conditions can be structural i.e. constituted by material, cultural or social structures in the organization, or related to the character of work processes and practices. Furthermore, the conditions can be seen as “objective” work conditions, or workplace affordances (Billet, 2001), provided by an employer; e.g. the learning potential of the work tasks, the work organization, and the available learning resources. Certain “subjective” factors also seem to affect the individual’s engagement in learning activities; e.g. self-confidence, motivation, values, previous learning experiences, learning readiness (Billet, 2001).

When it comes to learning in performance measurement, the learning potential of the task, and especially the discretion that the task entails, is of particular interest to explore. Discretion refers to the degrees of freedom, or scope of action, that the learning subject has as to how the task is defined, how the methods for solving the task are chosen, and how the results should be evaluated (Ellström, 2001). In combining these factors – task, method, and results – four levels of learning can be distinguished (see Table 1):

Table 1 - A taxonomy of levels of learning as a function of the discretion that exists in the learning environment (based on Ellström, 2001).

<table>
<thead>
<tr>
<th>Levels of learning</th>
<th>Adaptive learning</th>
<th>Developmental learning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reproductive</td>
<td>Productive type I</td>
</tr>
<tr>
<td>Aspects of the</td>
<td>(1)</td>
<td>(2) (rule-oriented)</td>
</tr>
<tr>
<td>learning situation</td>
<td>Reproductive</td>
<td>Productive type I</td>
</tr>
<tr>
<td>Tasks</td>
<td>Given</td>
<td>Given</td>
</tr>
<tr>
<td>Methods</td>
<td>Given</td>
<td>Given</td>
</tr>
<tr>
<td>Results</td>
<td>Given</td>
<td>Not given</td>
</tr>
</tbody>
</table>
In Table 1, reproductive learning (1) represents the lowest level of learning, which occurs when the tasks, methods and results are predetermined. This level of learning is sufficient and necessary in many activities, but its primary function is in the formation of competences for solving routine assignments. The next level of learning is called productive learning and it comes in two slightly different types, rule-oriented (2) and goal-oriented (3). Rule-based productive learning is characterized by a higher degree of discretion when it comes to the evaluation of results. The learning subject has to evaluate the outcomes and make minor changes to the methods used to solve the task. Goal-oriented productive learning has, in addition to the higher degree of discretion regarding results, also a higher degree of choice and use of methods. So even though the task may be given, the learning subject can experiment with ways of solving the problem at hand. The last level is called creative learning (4) and it represents the highest level of learning. In this type of learning the learner has to use his or her own authority not only to evaluate results or choose methods, but also to define the task itself, i.e. to diagnose the situation. In this type of learning the learner is free to question the definitions of tasks and problems posed by the environment (e.g. colleagues or management), and to act to transform standardized or institutionalized solutions. Creative learning, thus, occurs when individuals or groups of individuals within an organization begin to reflect on and question established ideologies, routines, structures and practices (Ellström, 2001).

A conceptual model
In order to further understand the relation between learning and performance measurements we propose a conceptual model (Figure 1) building on our workplace learning framework. The model builds on the distinction between the logic of production and the logic of development (Ellström, 2011) and the four levels of learning as a function of discretion (Ellström, 2001).

![The Performance Measurement Learning model](image)

By the logic of production, learning from performance measurement is primarily valued if it can contribute to operational efficiency. There is a strong imperative for reliability, and learning is strongly connected to the task of reducing variation in operations (Ellström, 2011). Learning from performance measurement can thus be said to focus on elimination of errors and aimed at organizational compliance with operational routines and rules. According to the logic of production, the context supports learning conditions aimed at incremental improvement, enhancing and making current operations more efficient, i.e. adaptive learning (Ellström, 2001). According to the logic of development, the context supports exploration of variety instead. Any deviation or
divergence from routine is viewed as a possibility for knowledge development and as an incentive for change and radical improvement. Failures are accepted and a certain level of risk taking is encouraged. The most critical component in a developmental context is organizational acceptance and encouragement for critical thinking and critical reflection. Learning from performance measurement by the logic of development can thus be characterised as developmental.

In the model we integrate Ellström’s (2001) four learning levels as performance measurement learning levels. The main condition separating the levels of learning is discretion of the agents involved. As used here, “agent” is an intentionally acting subject or collective, which occupy a structural position (Danermark et al., 2010). Agent discretion, in the form of bounded conditions for participation in operational and organizational development, has been widely addressed in the literature (e.g. Ellström, 2001; March & Olsen, 1976; Norros, 1995). Discretion necessitates not only the ability and will of the individual agent to participate but also the right to do so, in the form of organizational acceptance and encouragement (Kock, 2007). As such, agent discretion is viewed as a combination of cognitive (individual) and contextual (structural) factors creating conditions for participation. We use Ellström’s (2001) operationalization of discretion as tasks, methods and results:

- Tasks are directives expressing needs, desired outcomes and intended purposes of the performance measurement. Examples include targets, rules, regulations, orders, professional needs and/or insights.
- Methods specify the performance measurement procedures. These can be routine-based, “one-off” designs and/or experimentally developed.
- Results are performance measurement outcomes which can be specified with clear intentions, or developed according to unforeseen contingencies or arising needs.

In reproductive performance measurement learning there is little, if any, agent discretion in the choice of performance measurement tasks, methods and results. In rule-oriented performance measurement learning there is more freedom on behalf of agents having influence over results. In goal-oriented performance measurement learning, there is, in addition to freedom over results, also freedom concerning the actual performance measurement methods used. And finally, in creative performance measurement learning, agents have autonomy not only in relation to the performance measurement methods, results, but also in relation to the actual task of performing performance measurement.

**Research method**
The case study approach forms the methodological basis for the research reported in this paper (Eisenhardt, 1989; Voss et al., 2002). We aimed for a first attempt to build a theory of various forms of learning in contexts of performance measurement. The elements of the proposed theory are based on four components: key dimensions spanning up four various learning conditions; definition of types of learning environments that determine levels of learning, the domain in which the theory is applicable and specific predictions that can be made regarding forms of learning in different contexts of performance measurement. In line with Voss et al. (2002) the present work takes as its starting point that it is necessary to have a large and rich amount of primary data in order to make sense of the key dimensions and the types of environments that frame learning in performance measurement.

Our selection of case was an orthopedic and rheumatologic department located in a county council in southeast Sweden, an operations management context in which performance measurement is widely used. The yearly healthcare budget for the
organization is €15 million. The department has about 150 employees in nine sections treating both planned and acute patients. Each year, approximately 30,000 patients visit the department and about 3,000 surgeries are carried out. This case is considered to be best in class when it comes to utilization of performance measurement in both production and developmental work (Elg et al., 2013). We view this case as instrumental (Stake, 1995). Thus, we are not interested in its intrinsic qualities but rather its abilities for us to understand and develop our theoretical ideas through analytical generalization (Firestone, 1993). What, then, makes this case worth studying? In previous studies we have identified the organization’s wide variety of activities linked to performance measurement (Elg et al., 2013). This includes higher management’s strategic initiatives, internal strategic goal deployment, continuous follow-up in management meetings, various forms of improvement work and individual’s own initiatives for decision making. By and large, this enables us a single-case with great potential for studying many variations in which learning may be linked to performance measurement.

We used Critical Incident Technique (Flanagan, 1954) as a means to identify critical events related to the application of performance measurement (see also Elg, 2013). In interviews respondents were asked to elaborate on incidents with respect to (1) what was going on in this situation; (2) the types of performance measurements being used; (3) who was involved; and (4) short-term results and the perceived long-term effects. All interviews were recorded, transcribed and we extracted a total of 73 critical incidents. The critical incidents were first reported in Elg et al. (2013).

Respondents for the study were selected in order to maintain a high degree of variation in critical incidents. The following respondent groups within the orthopedic unit were included: Medical doctors (MD) (n=3), nurses (n=3), assistant nurses (n=2), care unit managers (n=2, also with background as nurses), financial manager (n=1), administrative unit manager (n=1), administrators (n=2), the clinical department manager (n=1, also MD), organizational developer (n=1, physiotherapist), and physiotherapist (n=1). In total, 17 respondents were interviewed. We viewed each critical incident as a specific type of environment that creates various conditions for learning in performance measurement. The analytical procedure followed a structured approach where our theoretically proposed framework guided the coding of each incident. This supported us in validation and further refinements of our key dimensions and definition of types of learning environments that determined levels of learning.

**Analysis of critical incidents**

In the analysis of the total number of critical incidents (n=73), 59 incidents were discovered where learning environments and conditions for different levels of learning conditions could be adequately categorized (for a sample of typical examples, see Table 2). In 14 critical incidents, contextual and/or individual factors inhibited any observable condition for learning and/or support for improvement. Descriptive examples of critical incidents where no conditions for learning and/or improvement could be found include critical incident (CI) no. 38, in which the sheer numbers of quantitative performance measurements produced within the studied organization creates an overload of information which led to lost opportunities for learning. CI no. 54 describes how the visualization of quantified performance measurements is not enough in creating necessary conditions for any kind of learning and/or improvement. In CI no. 68, MD’s describe performance measurements as a nuisance, irritating staff and discouraging any kind of conditions for learning and/or improvement.
Table 2 - Illustration of typical Critical Incidents (CI).

<table>
<thead>
<tr>
<th>Learning levels in performance measurement</th>
<th>Reproductive performance measurement learning</th>
<th>Rule-oriented performance measurement learning</th>
<th>Goal-oriented performance measurement learning</th>
<th>Creative performance measurement learning</th>
</tr>
</thead>
<tbody>
<tr>
<td># of CI’s</td>
<td>18</td>
<td>16</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>Example 1</td>
<td>CI no. 1: routine, survey based performance measurements (M), measuring compliance with pre-surgery hygiene routines (T) resulting in information on how to follow routine procedures (R).</td>
<td>CI no. 12: routine screening of patient journals (M) in order to book revisits (T) resulting in team discussions which helped create new booking system and changed routines (R).</td>
<td>CI no. 10: nurse booking new visits checking stipulated and guaranteed max waiting time of 90 days (T). No specific measurement method specified. She compares MD schedule with incoming referrals (T) resulting in new routine (R).</td>
<td>CI no. 35: MD use of national quality registries. By own initiative, MD check registries on professionally grounded hypotheses (T). Procedure decided by the MD (M). Result dependent on professional interest and critical reflection (R).</td>
</tr>
<tr>
<td>T=task</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M=method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R=result</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In total, 59 (80%) of the total number of CI’s (n=73) comprising the empirical material could be plotted in the model. These CI’s could further be elaborated and categorized in accordance to the particular learning types each of the CI’s supported. 8 CI’s (13%) could be categorized as performance measurement supporting creative learning. 34 CI’s (58%) were categorized as performance measurement creating conditions for reproductive or rule-oriented learning (low degrees of freedom) and 17 CI’s (29%) were categorized as performance measurement supporting goal-oriented learning.

**Reproductive learning**

Within reproductive learning, task, method and result are given in beforehand and within such a production logic context there is generally a strong alignment between task, method and result. The CI’s reveal examples with tasks specified according to routines and/or management directives (e.g. CI no. 1, and 60). Performance measurement is initiated as an organisational management action in order to follow standards and regulations, either locally or centrally governed. The methods are managed, designed, and implemented top-down and the desired results are set, often in accordance with the task. The staff have little, if any, influence or choice on what, how, and when the performance measurements should be executed.

**Rule-oriented learning**

Within rule-oriented learning, the CI’s reveal a higher agent discretion level concerning results and outcomes. The effects and outcomes of performance measurements are not specified or defined in any detailed sense. CI examples reveal dynamic team discussions where alternatives for solutions are discussed before decisions are made, and there is also room for individual influence. However, the alignment between task and method remains, and the examples reveal adherence to routines and pre-designed measurement methods such as complying with requirements, and using pre-set screening methods.

**Goal-oriented learning**

CI’s with goal-oriented learning show performance measurements which are usually decided “top down”, specifying particular tasks, such as a specific target or goal to achieve. The choice of performance measurement method and responsibility for results are, however, delegated to the staff. The featured example (CI no. 10) describes a
particular problem, as defined by a higher level (e.g. regulation or target), in this case a guaranteed waiting time of 90 days. The nurse responsible for executing the performance measurement has full discretion in the choice of measurement design and execution and also in how to handle the results. There is little alignment between task, method and result.

Creative learning
In creative learning the agent has full discretion over task, method and results. Our CI examples reveal a high degree of integrity and responsibility in the definition of task, the design of the method, and the handling of results. The MD in CI no. 35 makes all these choices and assumes a high degree of responsibility for the whole performance measurement process. From an agent perspective, our CI’s show a dominance of highly qualified professionals (i.e. MD’s) who can be assumed to have strong professional legitimacy and authority in the organization, granting liberties about also high levels of responsibility and freedom. The structural context (development logic) may be assumed to contribute to agent discretion, providing conditions for higher learning readiness.

Discussion
The purpose of this study was to contribute to a better understanding of the various levels of learning found in the practice of performance measurement. In utilizing established workplace learning theory we created a conceptual model combining learning theory and organizational management purposes. Four levels of learning determined by contextual and discrentional distinctions were identified: Reproductive, rule-oriented, goal-oriented and creative learning. Next, we empirically elaborated the conceptual model utilizing a case with critical incidents involving performance measurement activities. Our empirical investigation showed that it is possible to map performance measurement activities according to the learning conditions they create or support. The findings presented in this paper suggests that research on operations management benefit from theory development involving theories and concepts from other, related research fields. The workplace learning literature pointed to key issues interconnecting needs for learning and improvement with performance measurement.

We found that performance measurement always results in some form of informal learning process, though not necessarily resulting in improvement. Even the 14 CI’s we identified as not creating learning conditions for improvement certainly contain data referencing to other kinds of learning at risk of being quite detrimental to improvement. As Ellström (2010) points out, the processes of learning, change, and improvement need not always be positively related. The conscious design of performance measurements as means for learning thus forms a powerful tool in successful operations management aiming for continuous improvement. However, our study also shows that learning levels may interchange, back and forth, within a particular performance measurement activity, sometimes even allowing simultaneous learning levels. Therefore, design of performance measurement needs continuous reflection and conscious balancing of learning levels (Ellström, 2011).

Another finding is how organisational legitimacy and authority affects agent discretion and possibilities for creative learning. A majority of the respondents describing performance measurement supporting creative learning were MD’s or staff positioned as team leaders or organizational developers. This may imply an increased learning readiness due to better access structures (March & Olsen, 1976) and professional and positional legitimacy (Dowling & Pfeffer, 1975), and poses the question how operations management practices can give further impetus for empowerment, involvement and commitment within broader staff groups.
Performance measurement design has great potential for greater impact on improvement efforts in operations management. Mapping operations management practices (such as performance measurement) and the different conditions for learning provides practitioners with an opportunity to harness and manage inevitable learning participation (Sfard, 1998) and “hidden” potential for organizational improvement. Understanding workplace learning theory in performance measurement helps to detect dysfunctional learning environments, thus offering the possibility to break inhibiting hierarchies of learning and overcoming learning obstacles and constraints.

Critique may be levelled at the small empirical foundation and sampling in the study. We use a single case situated in a healthcare context in which it could be argued to exist a positivist and evidence-focused paradigm, governing performance measurement and learning. However, an underlying rationale in the sample choice was its richness and variety in performance measurement methods and practices which makes the case fill illustrative purposes. In conclusion, the study provides an evolved theoretical perspective on learning processes in relation to performance measurement, thus bridging a perceived gap in traditional operations management research. The study shows that utilizing and incorporating “untraditional” theory in an operations management perspective, create possibilities to accelerate improvement efforts in operations management.

Since the research presented in this paper is limited to data retrieved within a healthcare context, additional research will determine generalizability and comparability both to similar healthcare contexts, but also to other domains, such as industry and other public areas. Further research focusing on operation managers as potential learning facilitators may contribute to further knowledge on operations management practices and conditions for learning to support improvement. In this respect, contextual, structural, professional and cognitive factors in order to understand effective agency may constitute valuable research topics.

References


