

Towards a new definition of unlearning: rethinking individual
level unlearning and its implications for practice change.

by

Njoke Thomas

WP-11-01

Copyright
Department of Organizational Behavior
Weatherhead School of Management
Case Western Reserve University
Cleveland OH 44106-7235
e-mail: ler6@case.edu

Towards a new definition of unlearning: rethinking individual level unlearning and its implications for practice change.

Njoke Thomas

Case Western Reserve University

August 31st, 2011

Abstract

The concept of unlearning has emerged as a potential explanation for cognitive challenges associated with relinquishing outdated practices. Though introduced as both an individual level and organization level construct, far greater attention has been paid to the organizational level construct with little appraisal of the validity of individual unlearning. This paper will provide an overview of the research tradition that spawned the concept of unlearning and identify several unacknowledged assumptions that undermine the popular application of individual unlearning in organizational behavior theory. Integrating the work of theorists in several disciplines, a revised definition of unlearning is proposed. It is reframed as a form of learning in which the individual engages in intentional evaluation of self, task and environment to determine that a change of practice is necessary and possible. This process of evaluation and action is aided by mindfulness and meta-cognitive strategies. Future areas of research based on this new formulation are identified and discussed.

Introduction

Encouraging individuals to relinquish outdated practices is a primary preoccupation of organizations (Grenier, 1987 as cited in Klein, 1991), and yet our understanding of the mechanisms underlying practice change remains extremely limited. To be sure practice change at the individual level is a complex process that integrates knowledge acquisition, skill building and the ability to make specific meaning out of a number of environmental cues. All too often, this process seems to be obstructed by existing practices and habits as has been demonstrated by existing research on practice change in various professional settings (Clark, 2009; Macdonald, 2002; Rushmer & Davies, 2004)

The exploration of these ideas in the literature is often fragmented. There has been much interest in the cognitive aspects of the mechanism by which old practices give way to new. Of these cognitive theories, the most pertinent to this discussion is the concept of unlearning as framed by Hedberg (1981) as the conscious discarding of obsolete memories. Since its introduction to organizational behavior circles, the concept has been moderately popular, particularly among practitioners who have been uncritically receptive to Hedberg's ideas. However the idea has had less traction in academic fields until recently when a number of both theoretical and empirical papers on the topic have appeared (Howells, Mitev & Scholderer, 2010). The proposed need for unlearning as outlined by Hedberg and his colleagues contains a kernel of familiarity that resonates with theorists and practitioners alike, making it an attractive approach to the challenge of relinquishing outdated practices. However, deeper investigation of Hedberg's ideas and the research upon which it is based raises some serious reservations about the concept of unlearning as presently defined. The concept seems to be based on a number of assumptions that are wholly unacknowledged by either Hedberg or proponents of unlearning.

The intent of this paper is to demonstrate the most salient of these assumptions and provide a critical assessment of their veracity. It will begin with a brief overview of the history of the concept, tracing its origin to points prior to Hedberg and marking its trajectory into more recent inquiry. This will be followed by a deconstruction of the assumptions underlying the concept of unlearning as presented by Hedberg. Based on this critique, a reconceptualization of unlearning at the individual level will be proposed and justified. Finally, areas for future inquiry will be explored.

Unlearning in the Organizational Behavior literature

Unlearning was first presented to an organizational behavior audience as just one part of an extensive discussion of the interplay between individual and organizational learning. Almost as an afterthought, Hedberg proposed that the acquisition of new practices was often obstructed by the persisting knowledge of past practices. In essence, old practices needed to be unlearned, to create a space in the individual's memory for new practices to adhere. Seen from this perspective unlearning was an essential and seldom recognized aspect of the learning process. At that juncture, unlearning was formally defined as the ability to discard obsolete knowledge (Hedberg, 1981, p. 18). It was presented as both an individual and organizational level phenomenon, though organizational unlearning was theorized to be dependent on the ability of individuals to unlearn.

The introduction of Hedberg's unlearning model was followed by two distinct eras of theoretical discussion and empirical research; prescriptive and prescriptive (Tsang, 1997). These two eras represent divergent approaches to theorizing on unlearning. Prescriptive theorizing asked and answered the question, "How should an individual or organization unlearn?" This

writing was typically directed toward practitioner audiences, often written by practitioners themselves and offered anecdotal evidence derived from consulting experience. Their stated intent was to improve performance at the individual and/ or organizational level. This kind of writing was most apparent during the mid-80's but maintained a strong presence in the literature well into the late 90's. Two articles in particular, which alongside Hedberg are frequently cited as foundational unlearning literature, epitomize this era; Newstrom, 1983; and Nystrom & Starbuck, 1984.

Newstrom's particular interest in unlearning, which he defined as "the process of reducing or eliminating preexisting knowledge or habits that would otherwise represent formidable barriers to new learning" (Newstrom, 1983), stemmed from its apparent value for enhancing training outcomes. Recognizing that trainers often failed to acknowledge the effects (whether positive or negative) of preexisting knowledge on their attempts to train individuals, he identified specific training goals that would require unlearning for their successful attainment. These included decreasing the frequency of a current behavior (smoking), and replacing existing behavior with a totally novel behavior (adopting a new leadership style). Newstrom's primary concern seemed to be encouraging proponents of practice change to recognize individuals engaged in practice as more than just blank slates awaiting new imprints.

Nystrom & Starbuck (1984) defined unlearning as a process of "discovering the inadequacies of old ideas and discarding them" and identified it as an essential tool for preventing organizational crises. They identified both individual and organizational factors that were necessary precursors to unlearning. While individual required humility, objectivity and the self-confidence to face errors, managers could encourage unlearning by entertaining dissent, and fostering a culture of experimentation. Organizational unlearning could also be achieved by

removing individuals who either held critical knowledge or symbolized an outdated *modus operandi*. For example he proposed the removal of individuals in key leadership roles from the organization (scapegoating) as a means of organizational unlearning. In this conceptualization, practice is seen as more than an abstraction. It is something that resides in individuals. As such, the loss or removal of these individuals effectively makes room for new practices to take root.

It is interesting to note that during this initial era only one article emerged that could be characterized as critical of Hedberg's unlearning model. Klein (1989) called into question the physiological possibility of discarding knowledge. In particular, he expressed reservations about the empirical research underlying the theories from which Hedberg drew his model of unlearning. Ultimately, Klein concluded that the phenomenon Hedberg describe was not a loss of knowledge or ideas but in fact a reflection of increased learning about the nuances of a specific circumstances which then prompts new behavior. His proposal that the unlearning model should be abandoned in favor of his model of parenthetical or contextual thinking was largely unacknowledged during this period of prescriptive theorizing.

Arguably this dearth of critical assessment of the concept of unlearning had the effect of stunting the development of the unlearning concept into theory. Almost two decades after Hedberg's initial proposal very little was understood about the mechanisms underlying the phenomenon of unlearning at either the individual or organization level. Buchen (1998) who hypothesized a role for unlearning in promoting innovation was the first to propose several steps or phases of unlearning which combined both environmental factors and individual level responses. These included; initiation of the process as a result of persistent external intrusion that could not be ignored; a period of denial; gradual recognition of the unique properties of the present situation; and an ultimate utilization of new approaches and perspectives (which signals a

turning point in the unlearning process). Paramount in Buchen's thinking was the idea that organizational success often fostered complacency. Despite the negative valuation often ascribed to failure it could contribute greatly to organizational creativity, if approached with a positive attitude. For Buchen, this receptiveness to mistakes and failure was indicative of an unlearning culture. Buchen's propositions were a notable change in the orientation from the literature of this period. They marked the first step towards a descriptive theorizing body of literature.

Scholarly writing of a descriptive nature asks and answers the question, "How and under what circumstances does an individual or organization unlearn?" This writing is typically directed toward academic audiences, and combines rigorous research methods with systematic data collection to build explanatory theory. It would be some time before this descriptive approach would gain momentum. The interim was characterized by a transitional period which was not entirely the domain of either prescriptive writing or descriptive theorizing. The transition within this era was most apparent in the increasing number of theoretical and empirical articles referencing the concept of unlearning. It was often introduced as an explanatory model to explain a variety of organizational phenomenon including; the relationship between an organization's dominant logic and responsiveness to crises (Prahalad & Bettis, 1986; Bettis & Prahalad, 1995), the role of emotion in individual and organizational behavior change (Pratt & Barnett, 1997), and managerial awareness and perception regarding crises (Sheaffer & Mano-Negrin, 2003). A primary concern of this scholarship centered on the powers of unlearning to inoculate organizations against the effects of unanticipated change and or crisis events.

During this transitional period, unlearning also appeared with greater frequency in reviews of the organizational learning literature (Dodgson, 1993; Hamel, 1991, Hendry, 1996). In these accounts of unlearning, deference to Hedberg's model contributed to a remarkable level

of consistency in the definition of the concept. As a result, the concept established a strong presence in the literature despite limited understanding of the underlying forces that influenced its occurrence at either the individual or organizational level.

In the early 2000's the stage was set for the emergence of truly descriptive studies of unlearning with the development of a 'market-based ecology of unlearning' (Sinkula, 2002). This model, while ultimately focused on the benefits of unlearning at the organizational level acknowledged that unlearning was driven by a set of nested relationships; individuals within organizations and organizations within environments. This was an important contribution as it began to tease apart individual and organizational aspects previously unaccounted for in the unlearning literature. Shortly thereafter, Cegarra-Navarro and Dewhurst (2003) focused even more specifically on the concept of individual unlearning proposing that it consisted of three distinct phases; recognition of problems, alteration of established cognitive patterns and development of new assessment frameworks to be used in the future. Following this lead Cegarra-Navarro and Moya (2005), proposed a theoretical structural model that linked individual unlearning to organizational performance. Of particular interest to the authors was the link between unlearning and the organizational construct of intellectual capital, which they conceptualized as encompassing human, structural and relational capital. They specifically hypothesized relationships between individual and group level unlearning and each of the three component factors of intellectual capital. Most notable here is the introduction of unlearning at the meso level. In fact, they also sought to investigate whether individual unlearning was a necessary antecedent to group unlearning. Their study provided support for their hypothesized relationship between intellectual capital and unlearning and also revealed that individual and group unlearning could occur in either sequential or parallel processes. Furthermore, the

sequencing of these individual and group unlearning processes could have important consequences for organizational performance. In a subsequent study of unlearning in small to medium sized enterprise (Cegarra-Navarro & Dewhurst, 2006) it was determined that individual unlearning was significantly context dependent. Specifically three aspects of the shared organizational context that were identified as essential for positive performance outcomes to result from individual unlearning; individual context, management context and the context of the team. They also identified several important antecedents to unlearning at the level of the individual (tolerance for failure), organizational (psychological safety) and environmental (unpredictability). This particular study discussed for the first time potential negative effects of individual unlearning, proposing that individual unlearning occurring in isolation could have detrimental effects on the individual and ultimately undermine productivity.

The most consistent work to clarify the relationship between individual and organizational unlearning began in earnest in 2004 with a model developed by Windenecht (now Becker) and Hyland at the International Lifelong Learning Conference. Over the past seven years this model has been informed by a number of qualitative and quantitative studies (Becker, 2005; Becker, 2006; Becker, 2007; Becker, 2010; Becker & Delahaye, 2005; Delahaye & Becker, 2006). The most important contribution of this body work has been a clear distinction between individual and organizational knowledge and their methods of 'storage', better delineation of the concept of frames of reference which is identified as critical to the unlearning mechanism, investigating the link between individual personality type and ability to unlearn, and an investigation of the link between organizational culture and a disposition to unlearn.

Another important contribution during this phase came from de Holan & Phillips (2004) who developed a typology of forgetting, in which 'managed unlearning' was defined as a

deliberate process by which individuals discarded organizational schemas and routines to make way for new knowledge. This was particularly vital for the development of the concept as the terms forgetting and unlearning had often been used interchangeably in the foregoing literature. This relationship between unlearning and forgetting was further explored by Azmi (2008) who distinguished between four types of organizational forgetting on the basis of their intentionality and the valence of their outcomes. In so doing she identified unlearning as both positive and planned forgetting, juxtaposing it to sabotage (planned negative forgetting), decay (unplanned positive forgetting) and negligence (unplanned negative forgetting). Azmi ultimately proposes that organizations adopt a learn-unlearn-relearn (LUR) strategy as a means of maintaining an organizations culture of flexibility, agility, steadfastness and tactfulness. This model highlighted a body of literature on the concept of relearning which emerged in the early 00's. It was always implicitly understood that unlearning did not take place for its own sake thus the relearning was implied. This new concept could serve as either an alternative to unlearning or as Azmi proposed the final step in a three part sequence. While this model was created with organizations in mind and not individuals per se it is just as important to understand the extent to which this sequence occurs at the individual level.

As with the initial era, there has only been one notable challenge to unlearning in this descriptive era. Building on Klein's initial objections, Howells et al. (2010) argue that empirical research predating Hedberg had clearly established that individuals could not unlearn, in the literal sense of the word. They further challenged that several of the examples provided in both prescriptive and descriptive literature did not adhere to their own definitions of unlearning. Their specific concern is the lack of a discernible example of the deletion of existing knowledge. In each of the examples they cite, they demonstrate that the observed phenomena could also be

explained by additional learning about contextual factors. For example, their primary critique of Starbuck's anecdote of unlearning in the cold-war era Swedish Navy is that it was not so much the change in Swedish administration as it was the introduction of new technology that facilitated a more accurate picture of the environment. Thus, the incident recounted was ameliorated by the Swedish Navy's ability to achieve greater understanding through additional learning. Similarly, their critique of De Holan and Phillips qualitative study of expatriate Canadian managers working in Cuba, challenges the characterization of a need for 'managed forgetting'. Howells et al. contend that in reality what is being documented is a need for these expatriates to achieve a greater understanding of their new environment. What has been labeled as unlearning is thus an opportunity for more nuanced learning than was originally required to complete the same tasks in their native settings.

While these challenges seem to reflect the minority opinion on the concept of unlearning they still have serious implications for the promotion of unlearning as a means of facilitating practice change in organizations. These concerns will be explored in greater detail in subsequent sections. However it warrants noting the descriptive phase of the unlearning literature though more theoretically robust than the prescriptive phase ultimately fell short of either engaging in critical assessment of emerging ideas or establishing a clear distinction between micro and macro levels of unlearning.

Recurring themes in existing unlearning literature

To summarize, there are three distinct processes of unlearning that have been invoked if not fully articulated in the literature; individual unlearning, aggregated individual unlearning in the context of organizations and organizational unlearning. As previously noted, organizational

unlearning can occur in the absence of individual unlearning through the removal of key individuals or organizational structure.

Scholars of unlearning seem to agree on two key points; that it is a process as opposed to a discrete event and it is closely coupled with the acquisition of new knowledge (Becker, 2007). That is, unlearning is never undertaken for its own sake. Beyond this there seems to be little consensus.

For the better part of three decades since Hedberg first proposed his unlearning model, the distinction between individual and organizational unlearning has remained unclear. This is most apparent in a comparison of several articles that introduced a definition of organizational unlearning in the expression of their respective theoretical arguments (Tsang & Zahra, 2008). On further inspection, a number of these articles were describing individual unlearning within the context of organizations as opposed to a unique phenomenon of unlearning at the organizational level. This lack of distinction may be attributed in part to the intent of these articles several of which seemed to be focused the end point of organizational performance and as a result spent more time in articulating the measures of organizational performance. Much more time needs to be devoted to outlining an operational definition and measures of unlearning specifically at the individual level particularly considering how much of the existing literature focuses (whether intentionally or inadvertently) on the individual level phenomenon.

The origins of the unlearning concept

When Hedberg proposed individual and organizational unlearning as interdependent concepts, he borrowed substantially from a well-established tradition of experimental

psychology research on forgetting. In fact he based his assertions primarily on two articles which he identified as promoting “theories of unlearning” (Hedberg, 1981, p. 18). This body of research originated at the turn of the nineteenth century when Muller and Pilzcher demonstrated that consecutively memorized sets of nonsense syllables competed with each other for primacy at the time of recall (Anderson & Neely, 1996). This phenomenon, more specifically described as “the competition between memory items for a shared retrieval cue” (Anderson & Neely, 1996) or interference, generated significant interest and ultimately sparked the ‘classical interference era’. Commencing roughly in 1900, this seventy year period was highly generative both in terms of theory development and empirical research. This section will provide an overview of the theories attributed to this era that either set the stage for the emergence of unlearning or were otherwise critically influential in the development of the concept of unlearning.

Among the earliest theories was Thorndike’s ‘law of decay’ (Bjork, 2003; Postman & Stark, 1965) which proposed practice was essential for maintaining the memories generated through learning. Thus, it was disuse that led to forgetting. This theory had intuitive appeal as it seemed to reflect the subjective experience of forgetting (Bjork, 2003). However, the concept of decay was subsequently deposed by McGeoch who refuted the notion that forgetting was solely a function of use, noting that the nature of activities in the intervening period was also a critical factor (Bjork, 2003). McGeoch’s alternative theory of associative learning held that information was not lost from memory in any absolute sense, but under circumstances the ability to recall could become noticeably impaired (Bjork, 2003). According to McGeoch, each item in memory was associated with a specific cue, when more than one memory was mapped onto the same cue the competition that inevitably arose, decreased the likelihood of either of the memories being recalled.

The idea of associative learning would contribute immensely to the tradition of interference research. Researchers were particularly interested in two phenomena; proactive and retroactive interference. Proactive interference captures the extent to which the first memory associated with a stimulus or cue inhibits the association of future memories to that cue. Retroactive interference captured the extent to which new memories suppressed the recall of previously associated memories (Anderson & Neely, 1996; Bjork, 2003).

Anderson & Neely (1996) assert that the appropriate investigation of the interference phenomenon in experimental settings would require manipulation of three key factors; (1) the cues with which target memories are associated; (2) the cues that prompt memory retrieval and; (3) the link between cues and targets of interest to other items in memory. During this era, researchers seem to have achieved this balance by employing the classical paired-associate methodology in empirical research. Under this paradigm, studies were characterized by the memorization and recall of paired associates of verbal stimuli and responses. Study participants, who were typically male college students, were usually assigned a list of paired verbal stimulus and response to memorize until they demonstrated a criterion level of recall. At this point they were provided a second list to memorize which might consist of a pairing of the original stimuli with a new response. These trials would extend over the course of several hours or in some cases several days, but for the most part the duration of the experiments were relatively short. Throughout the course of these trials, researchers collected data on a variety of factors, including the number of trials that were necessary to memorize the new pairing to criterion and the extent of recall of the original list following this interpolated learning.

Employing this tradition of verbal interference studies, Melton & Irwin (1940) noted that in the course of attempting to recall interpolated responses research participants generated a

number of erroneous intrusions from the original pairwise listing of responses. This demonstration of interference between the two lists was no longer groundbreaking at this point in history. However, it was noteworthy that these overt intrusions were later linked to a significant decrease in the recall of the original responses beyond what could be attributed to mere forgetting (as measured by a control group). They surmised that these incidents of erroneous recall, and the recognition that in this particular context the responses were inappropriate, elicited a subtle negative internal response that served to weaken associations to the first list responses. They labeled this phenomenon unlearning and in a subsequent study (Melton & von Lackum, 1941) noted that the intrusive recollections delayed the rate of interpolated learning. It is from these interactions that emerged the idea that the full acquisition of new knowledge may rely on the unlearning of competing knowledge. The resultant two-factor theory of interference proposed by Melton suggested both competition between original and interpolated responses at the time of recall and a resultant unlearning of the original responses in the process of assimilating new responses into memory.

The subsequent thirty year period of research on the two factor theory of unlearning focused on the verification of various aspects of the theory including the hypothesis that unreinforced intrusions were the key mechanism through which unlearning occurred (Petrich, 1975; Postman, Keppel & Stark, 1965), the importance of the level of differentiation between original and interpolated responses, and the proportional contribution of proactive and retroactive interference to the process (Postman, Stark & Fraser, 1965).

Of the many findings that emerged from this line of inquiry was a growing conviction that the suppression of old memories was not absolute. Barnes and Underwood (1959) demonstrated that apparent suppression of recall was never more than 50%. Several other studies

also demonstrated the possibility of achieving full recall of the new list without suppressing the original information (Postman & Underwood, 1973). In other cases, interpolated learning could occur without apparent intrusions from the original list, and on these occasions a positive relationship was demonstrated between recall of both lists. These findings suggested that there was no interference between lists (Postman & Underwood, 1973). They also called into question the necessity of suppressing old knowledge as well as the possibility of fully extinguishing unwanted memories. It is worth noting that pursuant to these findings, Postman & Underwood's only firm conclusion was that additional research was essential to address the many inconsistencies and contradiction that had emerged in the preceding 70 years of interference research. Unfortunately, this conservative position on the phenomenon of unlearning was not captured in Hedberg's propositions a decade later. The following section will provide a more critical discussion of the interpretation of these foundational studies as demonstrated in both Hedberg's re-presentation of unlearning and that of the theorists and researchers who followed in his wake.

Deconstruction of the unlearning concept

Implicit in the proposition of unlearning in the organizational behavior literature are a number of nested suppositions which are at no time acknowledged by Hedberg in the course of putting forth his own theory of unlearning. These suppositions can be classified into three major categories: existential, methodological, and relevance.

Existential assumptions

Hedberg presents as fact that it is possible for individuals to extinguish memories. In support of this assertion he discussed two articles from the verbal interference literature. As the foregoing review of interference theory demonstrates neither of these were foundational studies (Anderson & Neely, 1996; Bjork 2003). Both were written at the tail end of the classical era of verbal interference studies. Furthermore, there seems to be substantial discrepancies between Hedberg's interpretation of these articles and their actual content. For example, Hedberg attributes to Postman & Stark (1965) the assertion that "response are deleted and replaced by new ones which are coupled with the old stimuli". However, the authors seem to be proposing a *reduction* in the availability of first list responses. The authors also describe this phenomenon as iterative, resulting from the process of erroneous generation of first list responses and weakening of their association with the stimulus. Yet Hedberg interprets their findings as an "abrupt phenomenon" (Hedberg, 1981, p. 18).

Hedberg's interpretation of Postman & Underwood (1973) is equally problematic. The most apparent problem is that the authors are described as proposing a theory of unlearning. At the core of this supposed theory, is a gradual decrease in the expectation of the original responses and gradual increase in the expectation of interpolated response in conjunction with the index stimulus (Hedberg, 1981, p. 18). In fact, the authors perform little if any theorizing in this article. Their stated intent is to provide a review of current empirical studies and theoretical critiques of the two-factor theory of unlearning (Melton & Irwin, 1940). Whereas Hedberg interprets this article as concluding that the most critical aspect of the unlearning activity is the perception of stimuli (Hedberg, 1981, p. 18), the authors actually review two competing theories about the importance of the stimulus component in the unlearning process. While they acknowledge the

potential value of these theories, they ultimately determine that there was insufficient empirical evidence to support any definite conclusions.

Melton and Irwin also reiterated that the unlearning mechanism involves the reduction in availability of original responses, as opposed to an absolute deletion of the association. By their own admission, the fact that (i) no original response is ever completely extinguished, and (ii) original responses seemed to exhibit variable resistance to suppression from interpolated responses made it difficult for them to accept that the acquisition of interpolated responses was dependent on the unlearning of original responses. If anything this article seems to refute several of the ideas that subsequently would be held as central to the notion of unlearning.

An additional concern with the concept of unlearning is the issue of intentionality. Hedberg himself does not specifically differentiate unlearning from forgetting but his definition does emphasize the intentionality of the unlearning process. Individuals and organizations must be deliberate in their identification of obsolete knowledge and in discarding them. However none of the studies cited in his article support such an assertion. These studies were not designed to capture an intentional process, which becomes increasingly apparent on further examination of the methodology of the earliest unlearning studies.

Methodological assumptions

The aforementioned discrepancies in Hedberg's interpretation of foundational studies certainly warrant greater scrutiny, but Klein offers an even more fundamental critique of unlearning. He contends that the methodology used in these studies is inadequate to the task of demonstrating unlearning. He takes particular issue with the stimulus-response paradigm employed in verbal interference studies, proposing that a research participant's designations of

stimulus and response may not necessarily adhere to the researcher's designation (Klein, 1991). From his perspective; there is nothing to prevent participants from reversing the associations of response and stimulus in their own memorization mechanism. Similarly, the research participant may be registering several stimuli as key in the course of the study that may not be accounted for in the interpretation of subsequent recall events. This is perhaps an unacknowledged limitation of the classical studies.

Bouton (1993) highlights another methodological problem that merits more careful scrutiny, arguing that verbal interference studies conflate memory retention and memory retrieval. Interestingly enough, this was an assertion made previously by McGeoch in his critique of Thorndike's law of decay. Bouton proposes that the observed difficulty in relinquishing practices lies not in the ability of a subject to acquire a new response, but in their ability to accurately select the appropriate response for the context at hand. He theorized that a stimulus unto which multiple responses have been mapped becomes ambiguous, compelling subjects to seek out additional cues to determine which response is appropriate.

Both Klein and Bouton argue for greater attention to the role of context in instigating memory retrieval. This idea of context appears to be more critical to understanding barriers to practice change than has been previously recognized in the unlearning literature, though this omission is slowly being addressed in the more recent literature (Becker, 2007; Becker, 2010).

To be clear, none of the studies that have been used as the basis of the unlearning concept as presented by Hedberg can actually prove that memories have been extinguished. In fairness, the technology to make such determinations simply did not exist during that time frame. It is not

clear if present technologies such as MRI scans might now be able to capture subtle changes in brain activity that could be distinguished as either learning or unlearning activity

As mentioned previously, the fact that memories are (apparently) never completely extinguished had already been established at the time of Postman and Underwood's review. A possible explanation for this resilience may be found in the domain of neurobiology. Research on the role of myelination in skill development demonstrates that the process of deep practice (analogous to the memorization process in classic interference tests) involves the insulation of neural pathways with sheets of myelin. Over time, continued practice results in thicker insulation of the pathway, which results in faster and more coordinated transmission of neural impulses. The relevant aspect of the myelination process is that it is somewhat irreversible. At best an abandoned pathway simply undergoes a halt in myelination, succumbing over time to typical wear and tear but not likely to disappear completely. So, the information as encoded in the brain as a specific neural pathway remains intact.

As described by Coyle, this process of myelination moves at a slow pace. Given that duration of many of these studies was seldom longer than a few hours it highly improbable that these critical alterations to the physical structure of participant's brains would have had an opportunity to occur. One would imagine though that deeply engrained practices, perpetuated over long periods of time would be demarcated by highly developed neurological pathways. This final point does call into question the generalizability of the studies on which the unlearning concept is based.

Relevance assumption (generalizability of foundational studies)

Howells et al. (2010) first presented the argument that Hedberg does not make a compelling case that the mechanisms revealed by the classic verbal interference studies can and

should be extrapolated to behaviors in organizations. However, their exploration of this specific concern is quite limited. This section will explore three immediately apparent challenges to the generalizability of these studies to practice change situations; these relate to complexity, duration and desirability.

What distinguishes the mechanisms studied in verbal interference research is their relative lack of complexity in the spectrum of human practices. It is understandable that a process as simple as verbal memorization and recall, afforded the researchers a considerable amount of flexibility. The research population need not be sophisticated, though there is a minimum criteria of literacy. This simplicity of design may certainly have allowed for a high level of fidelity across multiple repetitions of the research protocol. However, it is the very simplicity of the process that is cause for concern here. It does not seem appropriate that the process of memorization and recall of simple words or nonsensical syllables could be considered comparable to for example the process of memorizing and recalling on command two different sets of filing systems. The ability to acquire new information (whether conflicting or congruent with existing information) is potentially a modest part of a complex process that may also include knowledge integration and skill building.

It cannot be overlooked that the practices that are most difficult to relinquish are ones that an individual has used for an appreciable amount of time (Clark, 2009). Consequently it would be ill-advised to ignore the relevance of memory duration in these investigations. Typically duration of the verbal interference experiments could be measured in minutes, hours at most. Thus, it seems likely that the processes demonstrated would be more illustrative of short term memory as opposed to the long term memory associated with entrenched practices (Anderson, 1996). Interference research on this particular issue has demonstrated that different interference

paradigms dominate the long and short term memorization processes. Specifically the impact of retroactive interference is more apparent when retention intervals such are short, conversely proactive interference is more prominent when retention intervals are extended (Postman, Stark, & Fraser, 1968).

Finally, Hedberg failed to address the lack of emotional investment in the knowledge being acquired and/ or personal meaning ascribed to the experimental activities. Once again, the simplicity of the study design while conferring several advantages limits its practical implications. Research participants in these particular studies have very little incentive to privilege one set of responses over the other. If anything they might be intrinsically motivated to provide the right response. Arguably, this relative detachment from the experimental activity places the research participant at odds with individuals who may actively struggle to relinquish practices in real life situations.

As the limitations of Hedberg's proposition become more apparent it is tempting to attribute some causality to the narrow scope of literature on which his ideas are based. There were several concepts related to unlearning prevalent in the literature in the decades leading up to his proposition (Anderson & Neely, 1996), yet it is not clear that any of these ideas influenced his thinking. Perhaps he might have developed a more robust conceptualization of unlearning had he integrated ideas emerging from other quarters of memory and behavioral research. Verbal interference studies appear to be the mainstay of his understanding of unlearning yet these represent just one of several interference paradigms in simple associative learning (Bouton, 1993). It is useful to think of these paradigms in three basic categories; those that focus exclusively on retroactive interference (e.g. extinction), those that focus exclusively on proactive interference (e.g. latent inhibition), those that focus on the interplay between retroactive and

proactive interference (e.g. counter-conditioning). Verbal interference falls into this last category. However, much of the work in all three categories of interference paradigms has employed animal conditioning research methods (both classical and operant).

Within the animal conditioning literature, the concept of unlearning has been most apparent in the extensively studied extinction model of retroactive interference (Bouton, 1993). Extinction is defined as “the loss of learned performance that occurs when a Pavlovian signal or instrumental action is repeatedly presented without reinforcement” (Bouton, 2002). Throughout the literature on experimental psychology, some theorists have used the terms unlearning and extinction interchangeably (Barnes & Underwood, 1959; Postman & Underwood, 1973) but the latter typically refers to more complex conditioned responses. The similarities between unlearning in verbal interference and extinction in animal conditioning are significant. Some theorists have proposed that extinction was a critical step towards unlearning particularly in light of assumptions that extinction involves the destruction of the original association between conditioning stimulus and desired behavior (Rescorla & Wagner, 1972). There is reason to believe that the complexity of the experimental activities in animal conditioning studies provide a far better starting point from which to extrapolate to practice change. In addition Bouton (1993) has argued for the relevance of animal conditioning studies to treatment development in clinical psychology as well as cognitive behavioral disorders related to unwanted thoughts, emotions, and/ or behaviors.

Re-conceptualizing unlearning

Our actions are guided by what we know, and this same knowledge that can keep us tied to actions that are no longer fruitful. It is clear then if our actions are to change attention must be paid to identifying and transforming the underlying knowledge base. The predominant conceptualization of unlearning has been, and continues to be that of an intentional process that requires both awareness of the validity of one's knowledge and the ability to dispose of whatever is found to be obsolete. However the empirical evidence upon which theories of unlearning are based provide little support for either intentionality or the disposability of knowledge. The theoretical and empirical legacy of unlearning does make a compelling case for the occurrence and importance of knowledge interference. Individuals are not blank slates and therefore existing knowledge does need to be addressed in order to facilitate the cognitive, affective and behavioral mechanisms associated with adopting a new practice. Based on the foregoing discussion, it is proposed that more attention should be focused on the manifestations of unlearning at the individual level. Furthermore, it is suggested that a key revision be made to the most commonly cited definition of unlearning. This reframing of the concept advocates for a less literal application of the word, consistent with past assertions that unlearning is not the opposite of learning; instead it is a special case of learning. Thus, unlearning is redefined as a form of learning in which the individual engages in intentional evaluation of self, task and environment to determine that a change of practice is necessary and possible. This process of evaluation and action is aided by mindfulness and meta-cognitive strategies.

This new definition addresses issues of intentionality and contextually appropriate responses that are referenced but insufficiently explored in the existing literature. As discussed earlier, Hedberg conceptualization of unlearning attributes to Postman and Underwood the

conclusion that attention to stimulus is a critical aspect of the unlearning process. Despite a lack of support for this assertion in the literature quoted, it is worth noting that many of the strategies prescribed by unlearning proponents are based on the premise that individuals become stuck due to desensitization to key changes in their environment. For example Starbuck (1996) asserts that unlearning is contingent on an element of surprise, events that violate expectations. Similarly, Buchen (1998) proposed that an essential step in unlearning is the ability to recognize the unique aspects of the situation at hand. It follows then that an unwillingness or inability to read contextual cues fosters this insistence on choosing inappropriate responses to stimuli.

Both Klein (1991) and Bouton (1993) have proposed independent theories that reframe unlearning as enhanced attunement with context. Klein proposed the use of the term *parenthetical learning*. He argued that the term unlearning is inappropriate in part because it connotes a process that is antithetical to learning when in fact it is learning with a greater sensitivity to contextual meaning that is taking place. The individual is actually learning that stimuli can have contextual meaning. Similarly, Bouton provides an alternative context dependent mechanism for the unlearning phenomenon. He proposes that it is memory retrieval and memory retention that guides performance output (desired behavior). Further, memory retrieval is guided by contextual stimuli. As such it is attention to contextual stimuli that ultimately governs performance. Based on this reasoning, one would expect that a deeper understanding of individual unlearning would require greater attention needs to individual capacity for contextual thinking. Thus both mindfulness and meta-cognitive strategies are identified as important facilitators of individual unlearning. Such a link has been largely unexamined in the unlearning literature.

Flavell defined metacognition as “knowledge concerning one’s own cognitive processes or anything related to them” (Flavell, 1976 p. 232). He designated four main components of metacognition; meta-cognitive knowledge, meta-cognitive experience, goals and strategies. He further classified meta-cognitive knowledge into person, task and strategy variables. Person variables, which include an individual’s knowledge of their own learning processes, and task variables, which encompass the necessary knowledge to complete a given task, both seem to be vital factors in the experience of unlearning at the individual level.

In relating metacognition to experiential learning, Kolb & Kolb (2009) discuss several aspects of their theory that correspond with the person and task variables of meta-cognitive knowledge. The most pertinent to a reconceptualization of unlearning is the concept of learning self-identity, which as demonstrated by Dweck and Leggett (1988) can be either fixed or incremental. Individuals with an incremental learning identity have confidence in their ability to learn, particularly that it can be developed and expanded to incorporate novel areas of expertise. Conversely, individuals with a fixed learning identity perceive of their capacity to learn as finite and are thus resistant to any new knowledge that falls outside of these apparent boundaries.

The idea of a learning identity is particularly important when one considers the resource costliness of the learning process. Learning requires “conscious attention, effort and time on task” (Kolb & Kolb, 2009). It is to be expected that a commitment of such resources would be contingent on an individual’s belief that they are in fact capable of learning. Furthermore, there would need to be a desire to engage in the learning process. Even a person who believes that they can learn must have some incentive, whether intrinsic or extrinsic to engage in learning. Absent a conscious desire to engage in new learning about the changing circumstances of practice an individual will be disinclined to change. It is worth noting that the meta-cognitive strategies for

creating a learning self-identity proposed by Kolb & Kolb mirror the antecedents of unlearning identified by several theorists in the prescriptive era. The most ubiquitous of these proposed factors was psychological safety as realized through the promotion of an environment where experimentation is embraced and mistakes or failure are reframed as learning opportunities. Similarly the metacognitive strategies outlined for creating a learning identity include; redefining relationship to failure; managing emotional responses to failure and risking loss. Future research should explore the role of psychological safety in facilitating both mindful behavior and the meta-cognitive strategies that are essential for unlearning.

With respect to the task variables of metacognitive knowledge, it is also important to consider the circumstances under which an individual is no longer cognizant of the nature of the task at hand. Bargh and Chartrand (1999) explain that a task that occurs with sufficient frequency and reliability of outcomes can easily become an automatic response requiring little or no cognitive engagement. Ironically, the same processes that facilitate this automaticity, conceals it from consciousness. As a result, it is often difficult to be aware that automatic processes are taking place. Particularly when, as Nelson (1996) maintained, individuals are often imperfect evaluators of their own cognition. In addition, Clark (2009) estimated that individuals are only consciously aware of 30% of germane mental processes when they operate on a familiar task. This suggests that much of the underlying processes associated with individual practices are outside of our conscious awareness. Clark also observed that automated practices have a high level of accuracy and propose that this combination of reliability and resistance to change may provide an adaptive function. Though they may require a significant investment of resources to establish (Anderson, 2003), automated practices are far less costly in application than conscious thought. Thus, the relegation of frequently repeated practices to automatic control permits the

conscious mind to engage in other processes. In light of this high front end investment, they reason that the resistance to change may act as a failsafe that prevents important practices from being arbitrarily displaced by practices of lesser significance. It may be worth investigating whether practices that are deemed as less significant are easier to unlearn than practices deemed fundamental to the individual's existence.

The proposed emphasis on individual unlearning is by no means intended to invalidate existing scholarship on organizational unlearning. To the contrary more inquiry is needed to illuminate the mechanisms of organizational unlearning that is independent of individual unlearning. The concept on unlearning would be better served by a clear distinction between these two levels of unlearning.

The foregoing discussion has demonstrated that much remains to be understood about the cognitive aspects of practice change, particularly as it related to learning in its many forms. The application of the unlearning concept in the resolution of this persistent unknown reflects one of the greatest virtues of organizational behavior scholars, our ability to recognize untapped potential in the work of other disciplines. In this sense we may be exemplars of sustainable scholarship. However, the thoughtful reframing of old ideas into new theories requires vigilance, determined inquiry and holistic thinking. The unfolding story also reveals our abiding challenge in achieving these goals. The most apparent shortcoming of popular application of the unlearning concept has been the lack of holistic and integrative thinking. As such, this article may serve as an object lesson in the development of theory. With respect to the primary objective of clarifying the link between unlearning and practice change it is hoped that the questions raised herein will prompt more critical and integrative scholarship.

References

- Anderson, M. L. (2003) Embodied cognition: a field study. *Artificial Intelligence*, 149, 91–130.
- Anderson, M.C & Neely, J.H. (1996) Interference and inhibition in memory retrieval. In E.L. Bjork & R.A. Bjork (Eds.) *Memory* (pp. 237-313). New York: Academic Press.
- Azmi, F. T. (2008). Mapping the learn-unlearn-relearn model: Imperatives for strategic management. *European Business Review*, 20(3), 240-259
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54, 462-479.
- Barnes, J.M. & Underwood, B.J. (1959) Fate of first list associations in transfer theory. *Journal of Experimental Psychology*, 58, 2.
- Becker, K. L. (2005) Individual and organisational unlearning: directions for future research. *International Journal of Organisational Behaviour*, 9(7), 659-670.
- Becker, K. L. (2006) Unlearning: A people development issue for sustainable change and innovation. In: 7th International CINet (Continuous Innovation Network) Conference, 8-12 September, Lucca, Italy. Accessed on 2/17/11 at <http://eprints.qut.edu.au/10788/>
- Becker, K.L. (2007) Unlearning in the workplace : a mixed methods study. PhD thesis, Queensland University of Technology. Accessed on 2/27/11 at <http://eprints.qut.edu.au/16574/>
- Becker, K.L. (2010) Facilitating unlearning during implementation of new technology. *Journal of Organizational Change Management*, 23(3),251 – 268
- Becker. K.L & Delahaye, B.L. (2006) Unlearning as a lifelong learning strategy: An important pathway for transitions. In *Proceedings Lifelong Learning Conference*, pp. 26-31, Yeppoon, Queensland. Accessed on 2/27/11 at <http://eprints.qut.edu.au/7101/>
- Bjork, R. A. (2003). Interference and forgetting. In J. H. Byrne (Ed.), *Learning and memory* (2nd ed., pp. 268–273). Farmington Hills, MI: Thompson-Gale.
- Bouton, M. E. (1993) Context, time and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, 112 (1), 80-99
- Bouton, M.E. (1994) Conditioning, remembering and forgetting. *Journal of Experimental Psychology: Animal Behavior Processes*, 20(3), 219-231.
- Bouton, M.E. (2000) A learning theory perspective on lapse, relapse, and the maintenance of behavior change. *Health Psychology*, 19(Suppl), 57-63

Bouton, M.E. (2002) Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Society of Biological Psychiatry*, 52, 976-986

Bettis, R.A. & Prahalad, C.K. (1995). The dominant logic: Retrospective and extension. *Strategic Management Journal*, 16, 5–14.

Buchen, I.H. (1998) An organizational New Year's resolution: To unlearn. *National Productivity Review*, 18(1), 1–4.

Cegarra-Navarro, J.G. & Dewhurst, F.W. (2003). Unlearning as a prior step in the creation of intellectual capital in the organizational context: an empirical investigation. In *4th European Conference on Knowledge Management*, 18-19 September, Oxford University.

Cegarra-Navarro, J.G. & Dewhurst, F.W. (2006). Linking shared organizational context and relational capital through unlearning: An initial empirical investigation in SMEs. *Learning Organization*, 13, 49–62.

Cegarra-Navarro and Moya (2005). Business Performance Management and Unlearning Process. *Knowledge and Process Management*, 12 (3), 161–170

Clark, R.E. (2009). Resistance to change: unconscious knowledge and the challenge of unlearning in D.C. Berliner and H. Kupermintz (Eds.). *Changing Institutions, Environments and People*. Mahwah, NJ: Lawrence Erlbaum Associates. Accessed on 5/13/11 at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.99.1116&rep=rep1&type=pdf>

Coyle, D. (2009) *The talent code: unlocking the secret of skill*. Random House Publishing Group.

Delahaye, B., & Becker, K. (2006) Unlearning: A revised view of contemporary learning theories?. In Proceedings Lifelong Learning Conference. Accessed on 2/27/11 at <http://eprints.qut.edu.au>

Dodgson, M. Organizational learning: A review of some literatures. (1993). *Organization Studies*, 14, 375–94.

Dweck, C.S. & Leggett, E.L. (1988) A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2) 256-273.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911.

Hamel, G. Competition for competence and inter-partner learning within international strategic alliances. (1991). *Strategic Management Journal*, 12 (Summer Special Issue), 83–103.

- Hedberg, B. (1981). How organizations Learn and Unlearn. In P. Nystrom & W.H. Starbuck (Eds.) *Handbook of Organizational Design (Voll.)* London: Cambridge University Press.
- Hendry, C. (1996). Understanding and creating whole organizational change through learning theory. *Human Relations*, 49, 621–41.
- Howells, J., Mitev, N. & Scholderer, J. (2010) A Skeptical look at the use of the concept of unlearning on organisational analysis. Accessed on 3/7/11 at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1309525
- Kolb, A.Y., & Kolb, D.A. (2009). The learning way: meta-cognitive aspects of experiential learning. *Simulation Gaming: An Interdisciplinary Journal*, 40 (3), 297-327.
- Nelson, T.O. (1996). Consciousness and metacognition. *American Psychologist*, 51(2), 102-116.
- Klein, J.I. (1989) Parenthetical learning in organizations: Toward the unlearning of the unlearning model. *The Journal of Management Studies*, 26(3), 291-309.
- Macdonald, G. (2002). Transformative unlearning: safety, discernment and communities of learning. *Nursing Inquiry* 9 (3) 170-178.
- Martin de Holan, P. & Phillips, N. (2004). Remembrance of things past? The dynamics of organizational forgetting. *Management Science*, 50, 1603–13.
- Melton, A.W. & Irwin, J.M. (1940) The Influence of Degree of Interpolated Learning on Retroactive Inhibition and the Overt Transfer of Specific Responses. *The American Journal of Psychology*, 53 (2), 173-203.
- Melton, A.W. & von Lackum, W. J. (1941) Retroactive and Proactive Inhibition in Retention: Evidence for a Two-Factor Theory of Retroactive Inhibition. *The American Journal of Psychology*, 54 (2), 157-173.
- Nelson, T.O. (1996). Consciousness and metacognition. *American Psychologist*, 51(2), 102-116.
- Newstrom, J.W. (1983) The Management of Unlearning: Exploding the “Clean Slate” Fallacy. *Training and Development Journal*, 37(8), 36-39.
- Nystrom, P.C. & Starbuck, W.H. (1984) To Avoid Organizational Crises, Unlearn. *Organizational Dynamics*, 12(4) 53-65.
- Petrich, J. (1975) Storage and retrieval processes in unlearning. *Memory and Cognition*, 3(1), 63-74.
- Postman, L., Keppel, G., & Stark, K. (1965). Unlearning as a function of the relationship between successive response classes. *Journal of Experimental Psychology*, 69 (2), 111 – 118.

Postman, L. & Stark, K. (1965). The role of response set in tests of unlearning. *Journal of Verbal Learning and Verbal Behavior*, 4, 315-322.

Postman, L., Stark, K. & Fraser, J. (1968). *Journal of Verbal Learning and Verbal Behavior*, 7, 672-694.

Postman, L. & Underwood, B.J. (1973). Critical issues in interference theory. *Memory and Cognition*, 1, 19-40.

Prahalad, C.K. & Bettis, R.A. The dominant logic: A new linkage between diversity and performance. *Strategic Management Journal*, 1986, 7, 485–501.

Pratt, M.G. & Barnett, C.K. (1997) Emotions and unlearning in Amway recruiting techniques: Promoting change through ‘safe’ ambivalence. *Management Learning*, 28, 65–88.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W F. Prokasy (Eds.), *Classical conditioning: Current research and theory* (pp. 64-99). New York:Appleton-Century-Crofts.

(accessed online on 8/1/2011 at <http://homepage.mac.com/sanagnos/rescorlawagner1972.pdf>)

Rushmer, R. & Davies, H. T. O. (2004). Unlearning in health care. *Quality and Safety in Health Care*, 13(Suppl II):ii10–ii15.

Sheaffer, Z. & Mano-Negrin, R. (2003). Executives’ orientations as indicators of crisis management policies and practices. *Journal of Management Studies*, 40, 573–606.

Sinkula, J.M. (2002). Market-based success, organizational routines, and unlearning. *Journal of Business and Industrial Marketing*, 17, 253–69.

Starbuck, W.H. (1996) Unlearning ineffective or obsolete technologies. *International Journal of Technology Management*, 11(7/8), 725-737.

Tsang, E.W.K. (1997). Organizational learning and the learning organization: A dichotomy between descriptive and prescriptive research. *Human Relations*, 50, 73–89.

Tsang, E. W. K., & Zahra, S.A. (2008). Organizational unlearning. *Human Relations*, 61(10), 1435 – 1462.