How the Task-Set Influences Implicit and Explicit Learning

Sarah Esser & Hilde Haider

University of Cologne

Corresponding Author

Dr. Sarah Esser

Department of Psychology
University of Cologne

Sarah.esser@uni-koeln.de

Running Head: How the Task-Set Influences Implicit and Explicit Learning

Keywords: Implicit Learning, Implicit / Explicit Knowledge Representations, Conscious Awareness, Action Control, Theory of Event Coding, Associative Strength, Global Workspace Theory, Higher Order Thought Theories, Unexpected Event Hypothesis

Abstract

This article aims to provide a conceptual framework for implicit sequence learning that addresses two issues, both of which are important for understanding how implicit learning can guide our everyday behavior in coherence with our current task-sets and goals.

The first part of this article is concerned with the implicit sequence learning mechanism itself. Here, we try to develop a theoretical view on implicit learning by refining the critical concept of the "dimensions" implicit sequence representations refer to. By integrating the basic assumptions about uni- and multidimensional implicit learning with current theories on action control, we target current topics in implicit sequence learning research. This includes the question how implicit sequence knowledge is represented, whether it is helpful to differentiate between motor and perceptual implicit learning, the degree to which implicit knowledge can be said to be acquired and used in an automatic fashion and which role task-sets have in implicit learning.

In the second part, we ask for the mechanisms underlying the transition from implicit to explicit knowledge. Here, we review hitherto existing theoretical views and evaluate their compatibility with current scientific concepts of consciousness (the Global Workspace Theory and Higher Order Thought Theories). In this context, we introduce the Unexpected Event Hypothesis (Frensch et al., 2003) in an elaborated form and discuss its advantage in explaining the emergence of conscious knowledge in an implicit learning situation.

One important characteristic of many of our daily actions is their sequential character. Organisms adapt to sequential structures or statistical regularities inherent in the environment and use the acquired knowledge to predict future events (e.g., Clark, 2013; Rescorla & Wagner, 1972). Interestingly, most of this knowledge develops by repeatedly performing an action sequence or perceiving perceptual contingencies without any intention to learn them. Usually, neither this ongoing learning process nor the knowledge resulting from it is consciously aware. For example, we can walk our way from our office to the cafeteria while checking our phone, not really watching our steps or our surroundings that tell us when to turn left, take a stair or open a door. We also can detect a wrong note in a song we know well, even when we lack musical education. How do we learn about these regularities, how are they represented, which kind of information is extracted from the environment, how complex, abstract, transferable or flexible is this knowledge? These are all concerns of the research in the field of implicit sequence learning.

Since implicit (sequence) knowledge is usually treated as unconscious knowledge, it might easily be associated with the typical characteristics ascribed to unconscious (or automatic) processes, mainly going back to the work of Schneider and Shiffrin (1977; Shiffrin & Schneider, 1977). This inspired the investigation of the independence of implicit learning from attentional resources, its lack of controllability and inflexibility. Additionally, much effort has been made to rigorously examine and debate whether it is justified to classify implicitly acquired knowledge as unconscious (Newell & Shanks, 2014; Overgaard, 2017; Peters & Lau, 2015; Shanks & St John, 1994). Thereby implicit learning has often been contrasted against explicit learning and knowledge (Dienes & Berry, 1997; Jiménez, Vaquero, & Lupiañez, 2006).

This approach of investigating implicit learning surely has been very productive and appropriate to establish a core understanding of the underlying probably unconscious learning processes. However, in the last ten years, an interest in the interaction between conscious and unconscious processes has increased in various research paradigms concerned with unconscious processing (Kiefer, 2012; Kiefer, Adams, & Zovko, 2012; van Gaal, de Lange; & Cohen, 2012; van Gaal et al., 2014;). For implicit learning research, this shift in interests predominantly led to a stronger focus on the role of selective attention and task-sets (Abrahamse et al., 2010; Jiménez & Méndez, 1999).

Task-sets are commonly defined as the configuration of cognitive processes that enable us to bind the appropriate responses to the selected characteristics of stimuli in order to achieve a set goal (see Sakai, 2008, for a review). This includes, for example, the guiding of attention (Longman, Lavric, & Monsell, 2016), allocation of short-term memory resources (Sheremata & Shomstein, 2017), the inhibition of irrelevant aspects of a stimulus (Dreisbach & Haider, 2009), goal maintenance and resolution of conflict (Kane & Engle, 2003). Simply put, a task-set determines what aspects of our actions and the environment we represent, and consequently, also what we consciously perceive while pursuing a certain goal.

It is an old idea in psychology, leading back to Ebbinghaus (1885), that the representations we briefly hold in our consciousness will become associated, even though the process of association and the resulting knowledge of sequential information might not be conscious in itself. Yet, the role of the conscious representation of a task has played a very minor role in implicit learning research for a long time and has only recently been acknowledged as an important factor in explaining which sequential information will be learned implicitly and which won’t (Abrahamse et al., 2010). Even though gradually more evidence for the role of task-sets in implicit learning is accumulating, there is no up-to-date model that combines older and current results. One of the most influential and comprehensive models of implicit learning came from Keele et al. (2003). This model still contains valuable concepts. Nevertheless, it is not specific enough in some points in order to explain the interaction of conscious and unconscious processes involved in implicit learning.

In this article, we will incorporate some of the core assumptions of the Dual System Model from Keele et al. (2003) and extend them by including concepts from current theories of action control, most importantly the Theory of Event Coding (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Hommel, 2004, 2009, 2015; Shin, Proctor, & Capaldi, 2010), in order to account for the role of task-sets, respectively conscious task-representations. While the first part of this article will be concerned with the question of how conscious, or consciously available, information influences the implicitly learned representations, the second part of this article covers the opposite question: How can implicit knowledge become a conscious, explicit representation?

In the Dual System Account, this latter question is discussed only marginally and lacks reference to overarching conceptualizations of unconscious and conscious processing. Yet, the mechanisms how implicit knowledge can become explicit knowledge have received less attention even though many empirical findings show that implicit knowledge often but not always transforms into explicit knowledge at some point. The conditions under which explicit knowledge arises should be captured by a model of implicit learning. Moreover, this question is of theoretical and practical importance. On the theoretical side, when which unconscious representation will be selected to be represented consciously, how this selection is realized and which functions separate unconscious from conscious representations are some of the most controversial and enigmatic topics in psychology. These mechanisms of consciousness are difficult to access and often approached by priming studies (e.g. Kouider & Faivre, 2017; Kouider, de Gardelle, Sackur, & Dupoux, 2010; Lau & Rosenthal, 2011 Overgaard, 2003). Priming studies provide an ideal opportunity to study the differences between conscious and unconscious processing on a trial-by-trial basis and help to fathom the role of attention, signal strength, and beliefs. Implicit learning paradigms can be a valuable extension to these commonly used paradigms; they provide access to the question how the cognitive system can learn about its own internal states. So far, there is not much research on how the system changes from a state of not knowing that it knows something to knowing that it knows something. In the second part of this article, we aim to summarize the existing literature on this problem. We will evaluate the advantages and disadvantages of the proposed theoretical accounts and discuss the open questions. We will include these into our framework of implicit learning.

To address these questions, our article contains two main sections. In the first, we focus on implicit learning. Here, we start with a brief overview of the history of implicit learning research in order to understand the progressions made in the conceptualization of implicit learning mechanisms and current research interests. We then review the different theoretical views on what contents can be learned in an implicit learning situation and how these contents are represented. Lastly, we will discuss the question how attentional respectively task set mechanisms are related to implicit learning. In the second part, we will then turn to the question of how implicitly acquired knowledge can become a conscious representation.

Conceptualization of Implicit Learning

The following section will mainly incorporate research from the Serial Reaction Time Task (SRTT; Nissen & Bullemer, 1987), because this paradigm constitutes a very simple and versatile way to examine implicit learning processes and, in conjunction with that, also provides a large amount of studies.

In the most basic form of the SRTT, participants see marked locations on the screen which are mapped to spatially corresponding keys on the keyboard. On each trial, a location lights up and the participant's task is to respond with the corresponding key. Unbeknownst to participants, the stimuli on the screen, and thereby also the required motor commands, follow a systematic sequence. Typically, participants show a substantial learning effect in their performance data (latencies and errors). At the same time, participants usually are not able to express their sequence knowledge in a subsequent direct test when asked to either express their knowledge verbally (Rünger & Frensch, 2010) or to use it strategically (Destrebecqz & Cleeremans, 2001). This dissociation between task performance and verbally expressible knowledge is typically interpreted as implicit knowledge without concurrent explicit knowledge.

This simple version of the SRTT has been extended to designs allowing to test rather flexibly pure perceptual sequence learning or pure motor learning (e.g., Eberhardt, Esser & Haider, 2017; Gaschler, Wenke, Frensch & Cohen, 2011; Goschke & Bolte, 2013; Haider et al., 2011; 2012, 2014) on the one hand. On the other hand, it also allows to implement different sorts of sequences, such as deterministic or probabilistic sequences (e.g., Jiménez & Méndez, 1999). This extension of the SRTT makes it possible to answer also most of the research questions put forward in the field of statistical learning (Goujon, Didierjean & Thorpe, 2015; Perruchet & Pacton, 2006).

A Brief Overview of the History of Implicit Learning: Abstract Rule Knowledge versus Single Associative Transitions

Concerning the content of implicit learning, one question researchers are interested in is its representational format. Two conceptually different views can be distinguished here: One assumes that abstract rule knowledge is acquired. The other one proposes a simpler, associative learning mechanism. The account of abstract rule learning is mainly based on research within the AGL paradigm. Historically, this view is the oldest as it leads back to the studies of Reber (1967) and builds on findings which show that a new but structurally equal sequence in a transfer task also profits from the sequence learned during training (e.g. Francis, Schmidt, Carr, & Clegg, 2009; Reber, 1989). However, it has been debated whether such transfer effects, seemingly based on abstract rule knowledge rather lead back to at least partially explicit sequence knowledge (Gomez, 1997) or whether some form of statistical learning which is sensitive to perceptional features, respectively to repetitions, is responsible for such results (Conway & Christiansen, 2006; Gomez, Gerken, & Schvaneveldt, 2000). Even though more recent findings, which involve refined methodological designs, suggest that it is possible to acquire implicit abstract rule knowledge (Ling, Li, Qiao, Guo, & Dienes, 2016; Jiang, Zhu, Guo, Ma, Yang, & Dienes, 2012), the debate of implicit learning of abstract rules is not settled. . A more common assumption is that implicit learning involves the learning of single associative transitions, mainly governed by transition probabilities (e.g. Howard, Howard, Dennis, & Kelly, 2008; Remillard & Clark, 2001). This view became increasingly popular in the 1990s and has a somewhat stronger connection to the then upcoming SRTT paradigm, with which it has been demonstrated that implicit learning data can be modeled with simple recurrent network architectures (Christiansen, Allen, & Seidenberg, 1998; Cleeremans & McClelland, 1991). The view that implicit learning rather relies on knowledge about transitional probabilities than on extracting abstract rules from the training material can be extended by the assumption that this type of learning gradually leads to the development of information chunks (see Perruchet & Pacton, 2006, for a discussion).

What is learned in implicit learning?

Assuming that the learning of associations between single elements constituting a sequence plays a significant role in implicit learning leads to another question: What is the informational content of these associations? Looking at the traditional SRTT, as it has been introduced by Nissen and Bullemer (1978), participants are confronted with at least two concurrent sequences. The stimuli on the screen and the finger movements, respectively the to-be-pressed buttons follow the same sequence. This leads to at least four possibilities which associations are acquired in such a situation. First, implicit learning might be a motor process, leading to R-R associations. Second, perceptual knowledge about the succession of the stimuli in the form of S-S associations might be acquired. Lastly, stimulus and response information might become integrated in either the form of S-R or R-S associations (Abrahamse et al., 2010; Ziessler & Nattkemper, 2001).

Implicit R-R learning might be the most self-evident assumption about the content of implicitly learned representations. Examples for sequential activities within the motor domain, like typing on a keyboard or playing an instrument, most easily come to one’s mind, while at the same time, motor knowledge has the rather salient characteristic of not being easily accessible to verbalization (Wise & Willingham, 2009). Drawing on the work about nonimplicit motor sequence learning (e.g. Keele & Summers, 1976; Willingham, 1998), implicit R-R learning was among the first proposed mechanisms.

Despite this main focus on implicit motor learning, there have been attempts to show perceptual learning as early as in the 1990s. Howard, Mutter and Howard (1992) demonstrated learning of a stimulus sequence that was merely observed, i.e. no concurrent sequential responses were made. Mayr (1996) found learning of a stimulus sequence that was uncorrelated to the response sequence. For the following closer inspection of implicit learning, it is noteworthy that implicit perceptual learning can be differentiated in two different kinds: Visual implicit learning (Haider, Eberhardt, Kunde, & Rose, 2012; Haider, Eberhardt, Esser, & Rose, 2014; Turk-Browne, Scholl, Chun, & Johnson, 2009) and visuo-spatial implicit learning (Howard et al., 1992; Mayr, 1996; Remillard, 2009). While visual implicit learning can mean the learning of a color- or a shape-sequence, visuo-spatial implicit learning refers to the learning of stimulus locations. Since the identity of a stimulus comprises its color and other visual characteristics as well as its location, these two learning types both fall under the category of S-S learning. Apart from implicit learning in the visual domain, perceptual implicit learning has also been investigated in the auditive (Dienes & Longuet-Higgins, 2004; Weiermann & Meier, 2012) and haptic (Abrahamse, van der Lubbe, & Verwey, 2008; Kim, Johnson, Gillespie, & Seidler, 2014) domain, but to a much lesser extent.

The suggestion of an S-R learning mechanism within implicit learning goes back to the work of Willingham, Nissen, and Bullemer (1989). They found that participants who were trained with a motor sequence in which the keys were mapped to colors appearing at random locations did not show any transfer of this motor knowledge when the colors where removed and responses instead were cued by stimulus position. It was concluded that implicit learning is neither a purely perceptual nor a purely motor process and that “condition-action statements” (Wilingham et al., 1998, p. 1058) are acquired (for a current view of S-R learning, see Schwarb & Schumacher, 2010).

The last implicit learning mechanism, we would briefly like to introduce here is R-S learning which is often equalized with response-effect (R-E) learning. In several experiments, it has been shown that maintaining a certain response sequence while manipulating the contingency of the stimuli affects implicit learning (e.g. Ziessler, 1998; Ziessler & Nattkemper, 2001). Moreover, it has been shown that task-irrelevant additional effect stimuli (i.e. effect-tones) enhance implicit learning processes (Hoffmann, Sebald, & Stöcker, 2001; Stöcker, Sebald, & Hoffmann, 2003). Based on recent findings, it remains debatable whether R-S and R-E learning should be treated synonymously. R-E learning, outside of the implicit learning field, is strongly linked to intentional action control (ideomotor principle, Hommel et al., 2001) and it has been suggested that the acquisition of R-E associations depends on the interpretation of the stimuli being caused by one's own actions (Herwig & Waszak, 2009, 2012; but see Gaschler & Nattkemper, 2012, for a different interpretation).

To summarize, even within an associative account, there are many different proposals for implicit learning mechanisms. In earlier works, studies often aimed to isolate one single mechanism which is responsible for implicit learning. This led to seemingly contradictory findings. What was needed was a more flexible framework that allowed learning of a multitude of different sequences, including perceptual and motor learning. An important step in this direction was made by Keele et al. (2003) with their Dual-System Model.

Keele et al.'s Dual System Account

The Dual-System Approach of Keele et al. (2003) is based on viewing implicit learning as a multidimensional process. Consequently, in their model, there is not one central learning mechanism, like R-R or S-S learning, which is the only or most important one in sequence learning. Rather, different implicit learning processes can occur in parallel. The authors assume that implicit learning is realized by two independent systems. One is the unidimensional system which acquires knowledge unconsciously and independent of attentional resources. The other system is multidimensional; it is said to build up knowledge which is accessible to consciousness and relies on attentional resources.

The unidimensional system consists of multiple encapsulated modules which operate in parallel. Each of these modules is specialized for information along a single dimension (e.g. response locations, colors of the stimuli, etc.). The modules of the unidimensional system work independently from each other. They process the information they are specialized for and build up associations between all available predictable events along this dimension. This input specificity grants independence of attention and enables the parallel learning of multiple sequences even if they are not correlated.

The multidimensional system works differently. This system is able to integrate information between different modules and to build up associations across these dimensions. This enables the system to build up knowledge about sequences that consist of more than one dimension and that are only informative when both dimensions are considered (for example when the location of a stimulus predicts its next color). In order to protect this system from overload, attention functions as a filter mechanism. Only attended information is granted access into this system. The Dual-System Model has been a great contribution to understanding implicit learning as a multimodal, flexible process. The introduction of learning modules which work along specific dimensions has been especially helpful in understanding lots of former, seemingly contradictory, empirical findings.

However, since the model was first introduced further research points to two important issues in which regard the model could be improved. The first point already has been shortly addressed by Keele et al. (2003) themselves: The term dimension could profit from further specification which kind of information can be processed within one dimension. The second point concerns the postulated role of attentional mechanisms for implicit learning. More recent research has shown that the question whether implicit learning is dependent on attention, needs some finer specification which attentional mechanisms are or are not involved in implicit learning (Jiang & Chun, 2001; Jiménez & Mendèz, 1999; Musz, Weber, & Thompson-Shill, 2015). In the following section, we will provide a theoretical outlook that consolidates both issues by refining the definition of what a “dimension” is and how implicit learning depends on selective attention, respectively on a person’s task set.

Defining the term “Dimension”.

Keele et al. (2003) stated that the term dimension can be used interchangeably with modality, even though they also acknowledge that some modalities might also consist of more than one dimension. The general idea is that dimensions, viewed as modalities, refer to a defined, fixed network in the brain. For example, learning a color sequence might involve the V4 and the left fusiform gyrus (Bartels & Zeki, 2000; Simmons, Ramjee, Beauchamp, McRae, Martin, & Barsalou, 2007), while motor learning involves cortico-striatal-cerebellar pathways (Rose, Haider, Weiller, & Büchel, 2002; Tzvi, Münte, & Krämer, 2014). It remains somewhat unclear whether Keele et al. (2003) distinguish between different dimensions within the motor system like, for instance, a hand- vs. a foot dimension. Thus, dimension remains a rather abstract term. Few other efforts have been made since then to further define the term dimension within implicit sequence learning. Abrahamse et al. (2010) suggest that a dimension should be “regarded as equivalent to a specific type of feature, either at the stimulus level (e.g., shape) or at the response level (e.g., response location)” (p. 614). This distinction between stimulus level and response level corresponds well to the very common procedure of separating perceptual implicit learning (S-S) from motor implicit learning (R-R) as this distinction roots in the same conception of having learning mechanisms which are divided between the processing of stimuli and responses.

Slightly different, Goschke and Bolte (2012) assume that dimensions are based on separable attributes which are not specific to stimuli or responses. As an example for the blurred line between stimulus and response level, they state that the encoding of a stimulus location also activates a spatially compatible response code. The problem of implicit location learning makes it clear that a more specific definition of what constitutes a dimension in implicit learning is needed.

Based on the idea of loosening the distinction between a stimulus and a response level, Eberhardt, Esser, and Haider (2017) and Haider, Esser, and Eberhardt (2018) recently proposed to understand the term dimension as feature codes, irrespective of whether these belong to the stimulus or to the response. The concept of feature codes is derived from research in the field of action control, in particular from the Theory of Event Coding (TEC; Hommel et al., 2001; Hommel, 2015)

Hommel and colleagues (2001) postulate that action and perception are represented in the same format. A central assumption of the TEC is that actions and perceptions are both represented by their distal events in the form of consciously available feature codes. Feature codes consist of various proximal sensory and motor representations which have been associated and integrated over a person's learning history and therefore are distributed over the whole brain. Elsner and Hommel (2001) suggested a two-phased learning mechanism behind the development of a feature code. Whenever one interacts with the environment, one's own actions necessarily lead to various consciously accessible, distal sensory effects, represented in multiple corresponding proximal, unconscious modules. For example, moving left is associated with a multitude of visual, auditive and proprioceptive effects, of which each single effect is represented in distributed modules over the entire brain. Over the course of repeated actions, the motor commands that produced these effects and their proximal sensory representations will be bound together, constituting the feature code "left". As a result, such a feature code is the bi-directional, multimodal association between actions and their sensory effects. This bi-directionality is supported by the finding that endogenously anticipating or exogenously perceiving one single sensory element of the feature code results in the activation of the whole feature code, including an activation of the corresponding motor commands (Kunde, Hoffmann, & Zellmann, 2002). Conversely, executing the corresponding action leads to an activation of all associated sensory effects, making their detection much easier (Craighero, Fadiga, Rizzolatti, & Umilita, 1999; Wykowska, Schubö, & Hommel, 2009). This justifies the assumption of the TEC that within a feature code, no qualitative difference is made between the involved sensory or motor modules.

What do these assumptions of the TEC mean for the implicit learning research and the understanding of what a dimension is? Our proposal is that in implicit learning, a dimension is defined as the processing of an abstract feature code which is an integrated structure of different modalities, including motor and perceptual features of an action. This means that the representation of a task, respectively the task-set, has a significant role for the question what associations will be built implicitly. The same sequence usually can be defined by various different feature codes. For example, when one learns to play the piano, one might first code one's actions by the location of the keys. With progress in training, the keys become associated with the tones they produce and actions can then be coded by their effect-tones. The idea that environmental events can intentionally be coded by various different features is called *intentional weighting* in the TEC (Hommel et al., 2001; Memelink & Hommel, 2013). A piano novice will code the same sequence (piece of music) as a sequence of response-key locations, while the advanced player will represent a sequence of tones. Implicit learning by feature codes means that all the different modules that correspond to a certain feature code (or dimension), whether they process motor or perceptual information, will acquire implicit knowledge. Whenever this code is activated, the different corresponding modules will contribute to task performance. We will make this point clearer by giving two examples of implicit learning where the relevant dimension, or feature code, is "location", respectively where it is "color".

While learning of response locations is a widely accepted finding, it is, up to now, controversially discussed whether there is such a thing as pure perceptual location learning. Even when motor commands are controlled, it is difficult to exclude the involvement of eye-movements (Higuchi & Saiki, 2017; Mayr, 1996). Pure perceptual location learning should be defined as learning of attentional shifts without the involvement of any motor programming (Coomans, Deroost, Vendenbossche, Van den Bussche, & Soetens, 2012; Marcus, Karatekin, Markiewicz, 2006; Remillard, 2009). This seems to contradict the assumption of a very strong connection between visual location processing and motor programming as the perceived location of an object is usually used to guide the respective motor command (e.g., Goodale & Milner, 1992). Nevertheless, the processing of motor locations can be independent of the visual location information (Eisenberg, Shmuelof, Vaadia, & Zohary, 2011; Hardwick, Rottschy, Miall, & Eickhoff, 2013). Vice versa, also visual location processing can be independent of motor programming (Colby & Goldberg, 1999; Kesner & Rogers, 2004; Zimmer, 2008). If dimension is equalized with modality, as suggested be the model Keele at al. (2003), pure perceptual location learning and pure motor location learning should be possible and accordingly, two parallel, uncorrelated perceptual- and motor-location sequences should be learnable.

Viewing dimensions as abstract feature codes leads to different predictions. Perceiving a stimulus on the, for example, left position on a screen activates different modules which, over a person's learning history, have been integrated into the abstract feature code "left". This includes visual as well as motor-related information and, therefore, having two uncorrelated, concurrent perceptual- and motor "location" sequences should lead to interference and a lack of implicit learning. Thus, the prediction whether or not a person can concurrently learn two uncorrelated location sequences depends on the definition of the term dimension.

Another example for different hypotheses arising from the modality-based and the feature-code based accounts of dimensionality could be learning of a color-sequence. So far color-sequence learning is clearly regarded as perceptual learning in the implicit learning literature. However, within the TEC account, it has been shown that actions can be coded by the various effects they produce in the environment (Elsner & Hommel, 2001). It follows that a motor response sequence might very well be coded not only by response locations but also by other perceptual features, for instance, by color. It has been shown that neurons in the primary motor cortex quickly develop a sensitivity toward sensory features (e.g. color) when these features become relevant to the response (Zach, Inbar, Grinvald, Bergmann, & Vaadia, 2008). Therefore, implicit learning of a (perceptual) color sequence could be perturbed by a concurrent motor sequence and vice versa, if the participant is brought to code their responses by colors as well (Gaschler, Frensch, Cohen & Wenke, 2012; Wenke & Frensch, 2005).

The last example not only makes it clear that implicit learning based on feature codes blurs the line between perceptual and motor learning, but moreover suggests that implicit learning can be a flexible process, depending on the currently given task-set. This reliance on task-sets goes against the common assumption – also proposed in the model of Keele et al. (2003) – that implicit learning is a fully automatic and attention-independent process in which knowledge is acquired about unidimensional sequences whenever a person interacts with their environment. In the following section, we will discuss in which way implicit learning of feature codes seems to depend on task-sets and how far it can be said to be independent of attentional processes.

Implicit Learning of Feature Codes and its Relation to Attention

Leading back to the fundamental work by Nissen and Bullemer (1987), implicit learning has often been investigated under dual-task conditions. It has been extensively debated whether the implementation of a second, concurrent task reduces the amount of attentional resources which can be allocated to the implicit learning task and thereby interferes with implicit learning processes (Cohen, Ivry, & Keele, 1990; Frensch, Lin, & Buchner, 1998). The general conclusion of these studies was that implicit learning is independent of attentional resources (Frensch, Wenke & Rünger, 1999).

In an important study, Jiménez and Méndez (1999) later suggested that even though attention defined as the amount of mental effort dedicated to a task does not affect implicit learning, the selective function of attention might nevertheless play an important role. The sequential relation between a stimulus shape and a stimulus position was only learned if the stimulus shape was instructed to be attended to by a secondary task. Keele et al. (2003)

integrated the research on attentional demands for implicit learning under dual-task conditions and concluded that as long as implicit learning is unidimensional, even the selective function of attention is unnecessary, because no integration or exchange of information needs to be coordinated and the single modules can work in parallel. On the contrary, the selective function of attention is said to be needed for multidimensional learning to occur because a task-set is needed to specify which elements in the environment correlate with each other. Otherwise, in a real-life setting, multidimensional learning would hardly be possible due to the multitude of uncorrelated information disrupting meaningful correlated information.

Slightly different, within our proposal in which implicit learning depends on feature-codes, we assume that task-sets play an important role even for unidimensional learning. We agree on the assumption of Keele et al. (2003) that a multitude of unidimensional sequences can implicitly be learned in parallel without any attentional supervision. However, according to our proposal, the important constraint is that they need to rely on non-overlapping feature codes.

In an implicit learning task, a sequence could, for example, consist of a succession of colors and response-locations. Decorrelating these two dimensions can lead to parallel learning of both sequences (Haider et al., 2012), without the need for selective attention to the two features because there is no dimensional overlap and both sequences can only be coded by the respective feature that defines them. However, once a sequence can be described by more than one feature, it might become important which feature a person codes as being response relevant. Moreover, if two uncorrelated sequences overlap in the feature codes that describe them (e.g. stimulus location and response location), it would be necessary to code one of the sequences by a different feature code. For instance, one sequence would need to be coded by color and not by location, in order to maintain parallel learning of both sequences. That is, we assume that intentional weighting is important as it enables participants to separate abstract feature codes of otherwise overlapping sequences.

In the following section we will provide empirical evidence for the assumptions, that (1) implicit learning relies on abstract feature codes and that (2) the task-set determines by which features a sequence is coded.

Empirical evidence:

So far, not many studies have explicitly investigated the role of the task-set on implicit learning. One of the first studies to address this issue and to connect the ideas of the TEC and implicit learning came from Gaschler et al. (2012). In their experiment, two groups of participants performed an SRT with the same sequence. The sequence consisted of four colorless symbols. Each symbol was assigned to a corresponding key on a computer keyboard. In addition, these four keys were labeled with four differently colored stickers. The only, but important, difference between the two conditions was the instructions the participants received. Prior to the SRT training, one group was asked to memorize the mapping of shapes to key positions (diamond = outer left key, circle = outer right key, etc.; spatial condition) while the other group was asked to learn the mapping of shapes to key colors (diamond = red key, circle = yellow key, etc.; color condition). Due to the fixed color labels on the four keys, the sequence consisted of two correlated dimensions: A response-location and a color sequence. This was true for both conditions, regardless of which feature was emphasized in the instructions. After a training phase the keyboard was exchanged. On the new keyboard, the four colors of the training phase were arranged differently. For the test phase, all participants were given the instruction to respond to the shape stimuli with their respective colors (diamond = red key, etc.; Experiment 2b). This instruction left the previously presented color sequence intact while at the same time eliminating the previous response location sequence. Gaschler et al. (2012) found that only those participants which had been instructed to respond to the shapes with the colored keys before training (color instruction group) showed a preserved learning effect in the test phase. The group for which the keys had been labeled by their spatial positions before training did not show any color sequence knowledge. Thus, Gaschler et al. (2012) were able to show that the instructed task-set, respectively the coding of the sequence, plays a significant role in the question which information about a sequence is implicitly learned. Note that this result would not have been predicted by any theoretical account claiming that unidimensional learning is independent of selective attention or independent of the instructed task-set (e.g. Keele et al., 2003).

In a series of two studies (Eberhardt et al., 2017; Haider et al., 2018), we investigated the assumption that intentional weighting is necessary to resolve the conflict between two sequences that share an overlapping feature code. As mentioned earlier, learning of a parallel stimulus- and response-location sequence should not be possible because coding either the stimuli or the responses by “location” should activate sensory as well as motor modules. There are two studies which showed parallel learning of a stimulus- and an uncorrelated response-sequence (Mayr, 1996; Deroost & Soetens, 2006), which would refute our assumption and rather fit a narrower definition of dimensions as local modules. However, both studies had a design that allowed the response-sequence to be coded by object-identity instead of response-location. Our studies directly tested this alternative explanation.

In both studies (Eberhardt et al., 2017; Haider et al., 2018), there was a response-location sequence and an uncorrelated stimulus-location sequence. In order to achieve this decorrelation, participants received a colored target stimulus in the upper part of the screen that occurred at a certain location (one out of seven locations). In the lower part of the screen, six colored response squares were presented horizontally. These were spatially mapped to certain response keys. On each trial, participants were instructed to find the response square containing the target color and to press the mapped response key. The arrangement of the color of the response squares changed from trial to trial so that there was no fixed color-response mapping.

In the first series of experiments (Eberhardt et al., 2017), the participants always received a spatial stimulus-location sequence. In the first experiment, participants concurrently experienced either an uncorrelated response-location, an uncorrelated stimulus-color sequence, or no additional sequence. After the training, the stimulus-location sequence was replaced by a pseudo-random sequence. Participants acquired knowledge about the stimulus-location sequence when they received either a parallel stimulus-color sequence or no other sequence. In contrast, the stimulus-location sequence was not learned when the training material also contained a response-location sequence. This fits the results of Gaschler et al. (2012), who found no learning of the color sequence when participants were not instructed to code their responses by color; however the response-location sequence was always learned, regardless of the instructions. Taken together, these findings point to the interpretation that participants by default code their key-presses by location. Once they use the location coding for their responses, this code is not free for the coding of the stimulus locations anymore.

In the second experiment of this series, Eberhardt et al. (2017) went one step further. In two conditions, the participants received, in addition to the stimulus-location sequence, a dual-sequence. This dual sequence consisted of a color-sequence and a fully correlated response-location sequence (response keys were mapped to the color). The only difference between the two conditions was the instruction prior to the training. In the response-location condition, the participants were asked to code the keys in terms of their locations, whereas in the color condition, they were told to code the keys in terms of their colors (Gaschler et al., 2012). As in the first experiment, participants only showed learning of the stimulus-location sequence if they had coded their responses in terms of the colors. This further supports our assumption that the feature code “location” can only be used for either the response- or the stimulus-location sequence, hence that there is no strict difference between motor- and perceptual-learning.

Unidimensional implicit learning seems to be based on abstract feature codes that refer to motor- as well as to perceptual features, which have been integrated into abstract feature codes over a person’s learning history. Further, the second experiment supports the idea that task-sets have an important role, even in unidimensional implicit learning. Two sequences can be learned in parallel without the need for any selective attentional mechanism as long as they do not overlap in their feature codes. However, the task-set can modulate not only multidimensional but also unidimensional implicit learning; if a sequence can be coded by various features, a task-set can determine which of the features of a sequence is learned. If two sequences can be coded by the same feature, learning of both sequences can be realized by assigning different, non-overlapping feature codes to both sequences. This is in line with former results on intentional weighting. Hommel (1993), for example, demonstrated that the Simon effect (Simon & Rudell, 1967) can be modulated by the intentional coding of the task. The interference between input (the task irrelevant stimulus location) and output (a spatial response) could be eliminated if the stimulus-location was coded in a way that made it compatible with the location of the response.

A second series of experiments (Haider et al., 2018) provided evidence for the assumption that implicit learning relies upon abstract feature codes that incorporate sensory and motor information associated with the respective feature. Here, the goal was to show that an implicitly learned sequence of stimulus locations transfers to a new sequence of response locations. In this experiment, only an induction phase prior to training varied between the two experimental conditions. In the response-induction condition, the induction phase required participants to respond to one out of six different stimulus locations on the screen with spatially corresponding response keys on the keyboard. Stimulus locations appeared in random order, no sequence was present. This should establish an activation of the feature dimension “location”. In the color-induction condition, the participants were asked to count the appearance of one stimulus color and to enter the amount at the end of the block. Thus, these participants should be more likely to activate the feature dimension “color” upon perceiving the stimuli. In the subsequent training phase, all participants then were asked to judge if the target and the response stimuli were equal or different in color. Unbeknownst to all participants, the locations of the targets followed a sequence. In the test phase, one out of six different digits appeared at the center of the screen, cueing with which of the six response keys the participant should respond. Only the participants who were induced to code the target stimuli in terms of the location (in the response-induction condition) responded faster when the response locations in the test-phase matched the stimulus-location sequence in the training phase. Participants who were induced to code the stimuli in terms of their color (color-induction condition) did not show a transfer effect. Thus, this experiment shows that it is possible to express a location sequence in terms of key-presses when during training this sequence was only observed as a sequence of stimulus locations on the screen. This fits the assumption that the term dimension refers to abstract feature codes that do not exclusively belong to either the stimuli or the responses.

Taken together, the empirical findings support our suggestion that abstract feature codes are important to implicit learning processes. In addition, implicit sequence learning depends on the task-set that determines which dimensions of a sequence are task relevant.

Intermediate Summary

Historically, implicit sequence learning was rooted in two rather distinct areas, one being the automatization of motor programs and the other being the unintentional and effortless acquisition of grammatical rules. It has evolved now to a field that is concerned with the ability to assimilate to any sequential, respectively statistical information to which we are exposed to in our everyday lives. This pays respect to our ability to anticipate future events, even in situations in which the informational structure is highly complex.

We consider the conceptualization of unidimensional implicit associations being built along abstract feature codes to be an important extension of Keele et al.'s (2003) narrower conceptualization of dimensions as processing modules. Being strongly influenced by the concept of feature codes as proposed by the TEC, this new definition takes a person's learning history as well as the individual perception of an experimental task or real-life situation into account. Abstract feature codes are consciously available representations of distal events, composed of information from various proximal modules, containing sensory as well as motor-related information which correspond to an external event (Hommel et al., 2001). On the one hand, this view on implicit learning keeps the assumption of it being an automatic and unintentional process upright, while, on the other hand, it puts more weight on the conscious and subjective representation of the situation. As long as a person represents different features of an event (e.g. an experimental task), implicit learning about these features will take place in parallel. The task-set might require the participants to respond to the locations of stimuli with corresponding keys. When furthermore, the colors of the stimuli follow an uncorrelated sequence, but color is not response-relevant, consciously perceiving that there are colors will be enough for activating color as a feature code and implicit learning of the color sequence will take place (Goschke & Bolte, 2007; Haider et al., 2014,). In this sense, unidimensional implicit learning is automatic and independent from selective attention. However, implicit learning becomes more flexible within this view because events usually can be described by more than one feature and in such cases it becomes important by which features these events are consciously represented.

So far, building on the model of Keele et al. (2003), we have elaborated why implicit learning models will profit from a framework that takes the current conscious representation of a task into account. Abstract, hierarchically higher feature codes determine which modules are involved in the processing of a task. The different modules that are bound to one feature code comprise perceptual and motor information, this allows transfer of implicit knowledge from perceptually learned sequences to motor responses and vice versa.

What we have not touched yet, is the question under which conditions an implicitly learned sequence will become an explicitly represented sequence. During training each element of the sequence will be held in consciousness for a brief period of time, yet the associations between these elements often but not always stay unconscious. What is needed is a model of implicit learning that clarifies the conditions under which explicit knowledge will develop. The multidimensional system proposed by Keele and colleagues (2003) not only describes the necessary conditions for multidimensional implicit learning to occur, but further also aims to explain how explicit learning in an implicit learning situation can occur. However, the mechanisms behind this transformation from implicit to explicit knowledge mostly remain untouched and restricted to the assumption that multidimensional learning involves the ventral processing pathway (Grafton, Hazeltine, & Ivry, 1995; Hazeltine, Grafton, & Ivry, 1997), which is assumed to represent categorized information (Goodale & Milner, 1992; Goodale, 2011). We further believe that the explanation of explicit sequence learning in an implicit learning situation requires additional input from scientific theories of consciousness and furthermore is not exclusively related to multidimensional learning.

On the Emergence of Explicit Knowledge in an Implicit Learning Situation

The second section of this paper is concerned with the question under which circumstances and by which mechanisms an implicitly learned (sequence) representation can become a conscious representation. Other than the question by which mechanisms implicit learning is acquired, the question about the emergence of conscious knowledge has far less been researched and there are only a handful of models trying to explain this phenomenon. However, this question is of importance for various reasons. There is an obvious practical use when knowing how explicit knowledge will develop. Explicit sequence knowledge can have various benefits on performance; for example, it enables a person to gain flexible strategic control over their knowledge (Haider, Eichler, & Lange, 2011; Tubau, López-Moliner, & Hommel, 2007) and to transfer their knowledge to different task contexts (Jiménez et al.,2006). It can allow shortcuts, if these are possible (Haider & Frensch, 2005, 2009; Haider, Frensch & Joram, 2005), and it is usually associated with faster performance (Haider et al.,2011; Jiménez et al., 2006; but see Tanaka & Watanabe, 2017, for circumstances where performance is hindered by explicit knowledge). Furthermore, knowing how to promote the development of explicit knowledge can be helpful in several, for example, educational contexts (Pacton, Fayol & Perruchet, 2005; Pacton, Perruchet, Fayol & Cleeremans, 2001).

There are also more fundamental reasons why understanding the mechanisms underlying the transition from implicit to explicit knowledge might be important. It requires to evaluate general theories of unconscious as well as of conscious processing in order to create a model which is scientifically testable and compatible with the current state of empirical findings. In this relation, research on implicit learning can contribute to identifying strengths and weaknesses of theories about unconscious and conscious processing. This discussion will be picked up in more detail in the following section.

How to Conceptualize the Transition from Implicit to Explicit Sequence Knowledge

The explanation of how explicit knowledge arises from an implicit learning situation strongly depends on the conceptualization of consciousness. As already implied in the former part of this article, we make the presumption that unconscious and conscious processes can be separated, and moreover, that they can be investigated within an SRTT or similar tasks. It should of course be noted that there is still a debate on whether there is truly any evidence in psychological research for unconscious processing (see e.g. Newell & Shanks, 2014; Peters & Lau, 2015). These debates should generally not be neglected. From a methodological viewpoint, they bring very important and productive criticism leading to valuable improvements in the assessment of (un-)conscious knowledge (Rothkirch & Hesselmann, 2017). However, we do not debate the unconscious status of implicit knowledge in greater depth within this article.

Assuming that unconscious and conscious states should theoretically and can empirically be separated, there are two scientifically promising theories which attracted a lot of research and which both are important for conceptualizing the transition from implicit to explicit knowledge: These are the Global Workspace Theory (GWT; Baars, 1997; Baars & Franklin, 2003; Dehaene & Changeux, 2011; Dehaene & Naccache, 2001) and the Higher Order Thought Theory (HOTT; Dienes & Perner, 1999; Lau & Rosenthal, 2011; Rosenthal, 2012; see Deheane, Lau, & Kouider, 2017, for an argumentation why both, the GWT and HOTT are important for consciousness studies). In the following, we will shortly introduce the basic concepts of these two theories and will elaborate how they have already been applied in the field of implicit learning. Moreover, we will try to illustrate which problems both theories have when they are used to explain how explicit knowledge arises from an implicit learning situation. Lastly, we aim to discuss in which direction future research should go, in order to tackle these issues.

Global Workspace Theory

The GWT is a prominent functional and neuroscientific theory of consciousness. The basic assumption of the GWT is that the brain contains a multitude of functionally highly specialized subsystems working in parallel. Information in these areas is unconscious, there is no phenomenal- (Chalmers, 1995; Block; 2007), micro-consciousness (Lamme, 2006) or anything alike associated with information processing in these networks. Per se, these subsystems (networks) work encapsulated, that means they exchange information only within hard-wired or acquired pathways to fulfill their specialized task. This specialization enables the brain to handle a massive amount of input in parallel (Baars, 1997).

Nevertheless, coherent interaction with the environment requires serial output and therefore a mechanism is needed that selects information and puts it into the focus of attention. Here, the theory postulates a global workspace (GWS) mechanism which provides the necessary infrastructure, neurologically mainly realized by thalamo-cortical long-distance neurons of the prefrontal and the anterior cingular cortex (see Baars, Franklin, & Ramsøy, 2013 for a detailed elaboration of the neuronal architecture). The GWS is able to select relevant information, prevents interference, allows the encapsulated modules to exchange information and flexibly establishes temporary networks between these modules (Dehaene & Naccache, 2001).

The GWT uses a blackboard metaphor for imagining how the GWS works. When a module gets selected to enter the GWS, it can broadcast its content to any other network in the brain. Other modules can use this information from the blackboard and process it in their specified function. The information of the broadcasted module is no longer encapsulated. It is now said to be amodal because it is no longer bound to the specialized processes of the module it originated from, but instead is now processed in a broad context of unconscious subsystems. These subsystems include, for example, perception, language, intentions, selfconcepts, expectations, memory, and also exclusive access to working-memory function (Baars, 1997, 2005; Baars et al., 2013; Baars & Franklin, 2003; Cowan, 2010; Persuh, LaRock, & Berger, 2018; Schwager & Hagendorf, 2009). Neuroimaging shows that this de-capsulation of information is accompanied by a neurological “ignition”, a sudden, strong activation of a vast variety of cortical and subcortical regions (Dehaene & Changeux, 2011; Dehaene & Naccache, 2001, Rose, Haider & Büchel, 2010; Wessel, Haider & Rose, 2012).

Crucial to the GWT as a functionalist theory of consciousness is that conscious processing is equalized with the global accessibility of information and the thereby enabled options of using this information. There is no specific mechanism or place where consciousness is “created”. The GWT suggests a stochastic bottom-up variation-selection mechanism for explaining how the most relevant information is selected from the enormous amount of unconscious information (“Neural Darwinism”, Changeux & Dehaene, 1989). Every unconscious module constantly competes for access to the GWS (variation component), while the GWS sets a selection function depending on current goal states. Only one module or coalition of modules will show the strongest activation in the context of the current goal-state-dependent content of the GWS and will therefore win the competition for global broadcasting (Shanahan & Baars, 2005).

Global Workspace Theory and the emergence of conscious knowledge in implicit learning.

How can the GWT be applied to implicit learning research and the explanation how explicit knowledge develops in an implicit learning situation? At first sight, the conception of unconscious processing within the GWT seems to fit very well to the common conception of implicit sequence learning taking place in encapsulated modules. With our suggestion of implicit learning being mediated by feature codes, some more explanation is needed to clarify how implicit learning via feature codes still fits to this conception of encapsulated unconscious processing (here, implicit sequence learning).

With its roots in ideomotor theory (James, 1890; Greenwald, 1970), the TEC is a useful framework to explain how consciously available information, such as an intentional action plan or task set, can gain control over otherwise consciously unavailable processes. Within the TEC, it is assumed that the process of intentional weighting leads to an increased activation of the unconscious modules which correspond to the abstract feature code. This is compatible to the idea put forward by Dehaene and Naccache (2001) that unconscious processing is not restricted to predisposed circuits; instead conscious information within the GWS has the capability to establish temporary connectivities between different unconscious modules. Within these temporary circuits, information can be exchanged in an unsupervised, automatized fashion. Even though its routes have been established by conscious task-sets or intentions, the informational exchange itself is unconscious because it does not interact with other modules or sends information back to the GWS. Thus, in terms of implicit learning via abstract feature codes, the code itself is consciously represented which leads to a higher activation of the unconscious modules corresponding to this code. However, the learning taking place in the distributed modules itself is unconscious. Hence, our proposed model of implicit learning is still compatible with the conception of unconscious processing in the GWT.

Taken together, access to the GWS and thereby the conscious state of information is regulated solely by the competition between the representational strength of the unconscious modules. As a consequence for the transition from implicit to explicit sequence knowledge, it needs to be explained how implicit, encapsulated information can reach a representational strength high enough to win the competition for access to the GWS. A simple and suitable answer to this might be that with increasing training, the associative strength of implicitly learned information increases, making it a question of time and practice trials until implicitly learned representations win the competition for access to the GWS and turn into explicit representations. This has, with an explicit reference to the GWT, been suggested by Cleeremans and Jiménez (2002). They proposed three different factors which influence the quality of a representation: (1*) Stability*, i.e. the time a certain activational pattern can be maintained, (2) *strength*, i.e. the number of modules involved and their respective activational strengths, and (3) *distinctiveness*, i.e. the extent of overlap between representations within a functional network (see Kinsbourne, 1996, for a similar position). While implicit learning first leads to very weak representations, with practice these representations gradually gain quality and can become explicit. Within their framework the transition from implicit to explicit knowledge is a gradual process, involving a gradual increase in control as well as gradual change in subjective experience. With reference to Block (1995) the authors furthermore remark that there is a differentiation between access- and phenomenal consciousness within their framework.

The gradual change in control and experience assumed by Cleeremans and Jiménez (2002), however, does not fit the GWT. According to the GWT, consciousness with all its subjective and behavioral components is an all-or-none phenomenon (Dehaene & Changeux, 2011; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Del Cul, Baillet, & Dehaene, 2007; Kouider et al., 2010). This is supported by research in implicit learning which aims to examine the point in time a person becomes able to verbalize their acquired knowledge or use it in a strategic way. So far, it rather seems as if there is a certain moment or rather a short time window of “insight”, where a person switches from an unconscious to a conscious knowledge state. Haider et al. (2011) provided evidence that only participants who showed a sudden drop in their RT during learning were able to verbalize their knowledge by the end of training. The RT-drop seemingly reflects the moment where participants switched from stimulus- to plan-driven control (Tubau et al., 2007). Moreover, neuroimaging data like the sudden coupling of gamma-band activity, respectively increases of the BOLD-signal in the ventrolateral prefrontal cortex, the medial and ventrolateral prefrontal cortex and the ventral striatum have been shown to precede such an RT-drop (Rose et al., 2010; Wessel et al., 2012). These changes might reflect the sudden “ignition” of cortical activity which, as postulated by the GWT, accompanies the transition from an unconscious to a conscious state (Dehaene & Changeux, 2011; Dehaene & Naccache, 2001).

It might of course nevertheless be possible that the increasing quality of a representation plays a very important role in developing conscious knowledge of implicitly learned representations. It could be assumed that with increasing quality of a representation, its activation becomes stronger and therefore the likelihood gradually increases for this representation to win the competition for access to the global workspace. Once this threshold is passed, the representation becomes conscious in an all-or-none manner.

There are two further problems with the solution that the transition from implicit to explicit knowledge depends solely on the gradually increasing representational quality. The first has also been acknowledged by Cleeremans and Jiménez (2002) themselves: Representational quality might be a necessary, but not a sufficient condition for explicit, conscious knowledge to occur. First, it seems hard to imagine that implicitly acquired information can reach a higher representational strength than any of the other competing modules by a pure bottom-up process. As mentioned before, the GWT postulates a variation-selection mechanism. That means that in most cases the module that gets selected for processing within the GWS has the higher activation in the context of the current selection-function represented in the GWS. This way, current goal states have the option to provide top-down enhancement of the activation of potentially relevant modules. Selection based on pure bottom-up activation can happen in case of an alarming stimulus in the environment, e.g. a loud noise, but seems highly unlikely for a relatively weak implicit learning signal (see e.g., change blindness; Simons, Franconeri & Reimer, 2000). Cleeremans and Jiménez (2002) stated that attention and integration of information play an important role in determining whether any sufficiently strong representation will eventually enter a state of conscious processing. This supposed involvement of top-down mechanisms has not been elaborated any further. It therefore remains questionable how a top-down enhancement of implicit information could occur without a concurrent goal state that sets a fitting selection function for the modules processing implicitly learned information. Hence, if one aims to describe a mechanism for the emergence of explicit knowledge in an implicit learning situation based on the assumptions of the GWT, it needs to be explained how the system gets into a state in which the encapsulated module containing implicit information provides the fitting information.

Lastly, the second problem here might be even more profound. Even if it was explained how a goal state for which the implicit modularized representation provides the most useful information, it is yet unclear whether implicitly acquired information can simply leave its encapsulated state so that other modules, which then potentially gain access to its content, can interpret this information in their own specialized way. Usually, when we learn something explicitly, different modules are activated concurrently. If, for example, we learn about a new object, we have visual input, concurrently active with verbal and semantic information. When, instead, we learn something implicitly, like a sequence of movements or shapes, there never has been a connection between the sequential, module-specific input and any other information processing system. It seems unclear whether other modules can interpret this information simply by gaining access to it, so that, for example, a person becomes able to verbalize that they have learned a certain sequence.

An alternative to this could be that implicitly learned sequence knowledge stays in its encapsulated form and never gains access to the GWS itself. Instead, it is conceivable that explicit sequence knowledge needs to be acquired by developing a whole new representation via explicit reasoning functions. In this case, it would not have to be explained how an implicit representation in an encapsulated module can reach an activation high enough to win the competition against other modules or how top-down attention might be directed towards that module. Neither would it have to be explained how and whether other modules can interpret the formerly encapsulated module without ever formerly being co-activated with it. What instead would have to be explained is how the GWS gets into a goal state that mobilizes the relevant subsystems to learn the sequence explicitly. This will be discussed in more detail in a later section. However, beforehand we discuss the HOTTs as another important consciousness theory with the potential to explain the transition from an implicit into an explicit knowledge state.

Higher-Order Thought Theory

The Higher-Order Thought Theory (HOTT) in its most popular form goes back to the work of Rosenthal (1997; Dienes & Perner, 1999). The HOTT is concerned with the metacognitive aspects of consciousness. In its core, it differentiates between first-order and second-order (or higher-order) states. First-order states refer to simple input-output rules of any sensory or motor system. This can be understood in analogy to the parallel working modules in the GWT. Encapsulated, respectively implicitly learned information can be seen as a first-order state which per se is unconscious. Not only the human brain, but any simple or complex machine which shows discriminatory performance has first order states (e.g. perceiving light of a certain wavelength results in the output of detecting red). Consciousness, according to the HOTT, crucially depends on developing higher-order knowledge about this first-order knowledge. Put simply, consciousness means knowing that one knows. This comprises the ability for self-reflection, self-reference and a propositional attitude (e.g. “I *know/believe/guess* that it is red that I see”, “It is *I*, who sees red”, “it *is red* that I see”, Dienes & Perner, 1999). What is needed for consciousness is a mechanism that allows the brain to draw inferences about its own internal first-order states and about how these relate to states in the environment. Different theoretical suggestions and models have been put forward to describe the learning process behind the acquisition of higher-order knowledge about first-order states (Fleming & Daw, 2017; Lau, 2008; Lau & Rosenthal, 2011).

Higher-Order Thought Theory and the emergence of conscious knowledge.

In his recent work, Cleeremans (2008, 2011, 2014) has applied HOTTs to the question how explicit knowledge develops in an implicit learning situation: Through interaction with the environment, a first-order representation is developed, gradually improving in quality, as originally assumed by Cleeremans and Jiménez (2002). The crucial addition to their former stance and the new implementation of HOTTs is that the acquired first-order information is never conscious; it is labeled as knowledge *within* the system. For consciousness to arise, the first-order information needs to be redescribed as a metarepresentation; that is, knowledge *for* the system (Clark & Karmiloff-Smith, 1993). The first-order representation itself becomes an object of a representation for higher-order systems. This higher-order system receives input from the first-order systems and learns that the state the first-order system has changed and thereby develops a higher-order attitude towards the first-order knowledge (e.g. "knowing that …", "hoping that …", "seeing that …"). This higher-order representation is assumed to be a new representation involving a broad pattern of activation over different processing units which is only indirectly shaped by the changes of the connection-weights within the first-order system. The proposed learning mechanism behind the first- and the higher order learning system is the same; both systems gradually improve the quality of a representation with each learning trial. Pasquali, Timmermans, and Cleeremans (2010) have investigated the relation between first-order sensitivity and higher-order awareness measured by the Post-Decision Wagering Task (PDWT; Persaud, McLeod, & Cowey, 2007) within different paradigms (i.e. Blindsight, Iowa Gabling Task, and an AGL Task). These results supported the assumption that the higher-order representations gradually improve with the learning progress of the first-order system.

Surely, there are debates how exactly the relation between first-order knowledge and a meta-cognitive learning mechanism should be modeled with most of the suggested models being based on bottom-up signal-detection theories (Barrett, Dienes & Seth, 2013; Fleming & Lau, 2014; Maniscalco & Lau, 2012, 2016). What they all have in common is the gradual development of higher-order, respectively conscious knowledge. The conscious state of a representation changes from guessing, which is equalized with being unconscious about a first-order representation, to knowing, which is equalized with being conscious about a first-order representation (Dienes & Scott, 2005; Sandberg, Timmermans, Overgaard, & Cleeremans, 2010).

Applied to the question how explicit sequence knowledge develops in an implicit learning task, this leads to one question: With gradually increasing metacognitive knowledge, when is a person able to state that they have detected that there is a sequence hidden in the task and consequently is able to report that sequence? Imagine a person not only gradually giving more correct responses in an SRTT (via first-order learning) but also noticing that their perception of the task gradually changes to knowing the next response in advance (via higher-order learning). Does the moment this person notices that they know the answers equal the moment the person knows that there is a sequence and what this sequence is? This question is related to the aforementioned problem based on the GWT, whether there can be an immediate access to the contents of the first order representation or whether a new representation needs to develop via explicit, conscious reasoning processes.

A rather simple higher order learning mechanism as proposed by Cleeremans (2014) might indeed provide an important basis for a cognitive system to determine what first order state it is currently in. We agree with the assumption that meta-cognitive learning plays a significant role in gaining conscious insight into otherwise unconscious information processing. Yet, we think there are a few problematic aspects about this simple explanation that need to be considered. First, the mechanism described by Cleeremans and colleagues (Pasquali et al., 2010) is tested in situations where a person is directly asked to evaluate the correctness of their responses. It is questionable whether such an evaluative process of one's own behavior happens automatically and in parallel when there is no external instruction to do so (as there is in a subsequent PDWT). While, as shown in the first part of this article, the research implies that implicit learning processes can happen in parallel (Goschke & Bolte, 2012; Haider et al., 2012, 2014, 2018), it is not granted that higher-order learning processes can happen in parallel for all implicit learning processes. It might be that higher-order learning processes rely on an intention, respectively selective attention to evaluate one specific behavioral output. In this case, it needed to be explained how the system decides which first-order representations are used to develop higher-order representation.

Assuming that higher-order learning processes happen automatically and in parallel to the acquisition of first-order knowledge, as it seems to be implied by HOTTs, leads to another problem: The higher-order learning process informs the system that knowledge has been acquired, but knowing that one knows (instead of guessing) the correct response seems close but not equal to knowing that there is an underlying sequence. Rather, on a subjective level, knowing that one knows the correct answer will most likely be a surprise, resulting in wondering why one knows the correct answer. Even coming to the conclusion that the reason is an underlying sequence, this further does not directly imply that the person knows the exact regularity without further inferential processes (Scott & Dienes, 2010).

A third question that arises in the light of such a simple higher-order learning mechanism is whether learning to know that one's own responses are increasingly often correct is the only option to develop explicit from implicit sequence knowledge. There are a few studies showing that explicit knowledge seems to develop whenever an unexpected change in one's own behavior occurs (Esser & Haider, 2017; Haider & Frensch, 2005, 2009; Haider et al., 2011; Rünger, 2012; Rünger & Frensch, 2008; Schwager, Rünger, Gaschler & Frensch, 2012). This includes, for example, noticing premature responses before the next stimulus occurs (Haider & Frensch, 2009), sudden changes in the sequential structure which lead to slower reaction times and an increased amount of errors (Rünger & Frensch, 2008), or changes in the perceived fluency of the task performance (Esser & Haider, 2017).

On the theoretical side, the problem here is to explain how such findings can be explained by a higher-order learning process that works with a simple bottom-up, gradual strengthening mechanism. A multitude of parallel higher-order learning processes would be needed to be active at the same time to evaluate all the different aspects of the task (i.e. error rates, reaction times, fluency, etc.) for just one single sequence. Moreover, such results imply sensitivity for sudden changes. This implies that a learning process involving expectations, predictions and violations thereof, rather than simple associative strengthening need to be considered. On an empirical side, this is further supported by the above mentioned studies, which used different manipulations for balancing the associative strength between conditions but manipulated whether small or large violations of expectations occurred. For example, Esser and Haider (2017) showed differences in the emergence of explicit knowledge when the structure of the task led to noticeable differences in the fluency of the task material, while the amount of regular and irregular sequential trials was equal to a condition that could not experience such differences in the experienced fluency. A simple bottom-up higher-order learning mechanism does not include the size of prediction error (here, the sudden changes in fluency) as a signal. In the following section, we will propose a model which includes ideas of the HOTTs and the GWT in order to respond to the formerly described problems.

The Unexpected Event Theory

The Unexpected Event Theory (UEH) was originally proposed by Frensch et al. (2003). It is based on an idea by Dienes und Perner (1999) who stated that the observation of one's own behavior is a central aspect for the emergence of explicit knowledge from an implicit learning situation. The UEH aimed to improve the explanation how and when implicit learning can trigger an inferential process.

In its original form, the UEH shared the assumptions of Cleeremans and Jiménez (2002) as well as of Keele et al. (2003) on implicit learning. In the UEH it is assumed that implicit learning is a byproduct of interacting with the environment. By repeatedly interacting with sequential information, associative weights will gain strength and implicit representations will develop. It is further assumed that due to its encapsulated nature and the circumstance that there is no intrinsic conscious property of implicit knowledge, there has to be an additional mechanism that can transform it into explicit knowledge.

In the context of the first part of this article, this view has not changed in its core. What has been added is that the term "encapsulation" now refers to less hard-wired processing units, but allows these units to interact within the boundaries of the abstract feature codes that correspond to perceivable events in the environment. However, as outlined in the first section, the knowledge within the involved modules construing one abstract feature code remains encapsulated and the question remains which mechanism grants access to this knowledge.

The crucial idea of the UEH is that explicit sequence knowledge can only develop when a person unexpectedly notices a change in their own behavior. This can trigger an intentional search for the sequence. In an implicit learning situation, interaction with the task leads to continuous improvement of the responses to the stimuli; they become more accurate and faster. It can be this improvement or, for example, the feeling that the task becomes more fluent or easy, that there is a certain rhythm in one's own responses, or even an external event that directs the participant's attention towards noticing an underlying pattern and triggers following search processes. These search processes do not necessarily lead to a detection of the sequence if another explanation seems more likely to account for the unexpected event (Haider & Frensch, 2005). Generally speaking, the UEH comprises a monitoring process which constantly compares expected and actual experiences. This comprises internal, experiential, as well as external, behavioral deviations from one's expectation. This process allows detecting unexpected changes and initiates an attributional process for the detected conflict in order to adjust its predictions and reestablish coherence between the distant environment and one's proximal model of it. Comparable monitoring-models have been established in neurocognitive models of conflict-detection and adaption (Botvinick, 2007; Botvinick, Braver, Barch, Carter, & Cohen, 2001), metacognitive control (Koriat, 2000, 2012, 2015), or memory (Whittlesea, 2002; Whittlesea & Williams, 2000).

It is an aim of this article to elaborate the processes behind the original proposal of the UEH a little further and to point to open questions which should be addressed by future research. We believe that first further elaboration is needed concerning the mechanism that allows the detection of unexpected events. In this relation HOTTs and research on metacognitive, higher-order learning processes are very important. What is needed is a mechanism which allows a comparison between the expected metacognition we have of a given situation and the experienced metacognition.

Right now, there are several different models aiming to explain the relation between the first-order signal (here, the implicit knowledge) and the metacognitive evaluation of these signals. Usually these rely on simple signal-detection models (Cleeremans, Timmermans, & Pasquali, 2007; Galvin, Podd, Drga, & Whitmore 2003; Lau & Rosenthal, 2011; Del Cul, Dehaene, Reyes, Bravo, & Slachevsky, 2009). The problem with these models is that they are often pure bottom-up models that do not take several important top-down factors into account which have often been shown to influence metacognitive decisions. This includes for example the use of heuristic cues (e.g. fluency, luminance) which have no direct relation to the first-order knowledge the metacognitive judgement is relating to (Hoyndorf & Haider, 2009; Koriat, 2007; Wilbert & Haider, 2012). It further includes the role of previous experiences, for example, with similar situations, successes and failures, or general knowledge about one's own performance capacities.

There are, however, a few promising suggestions modelling the relationship between first-order knowledge and metacognitive judgements with Bayesian learning (Fleming & Daw, 2017; Sherman, Seth, Barrett, & Kanai, 2015). One advantage here is that Bayesian models allow metacognitive learning via predictive coding (Clark, 2013; Friston, 2010). The evaluation of one's own behavior respectively knowledge leads to a first hypothesis of what metacognitive experience is expected in the next, similar situation. This prediction is used to be compared with the current experienced metacognitive judgement and, in turn, the resulting error-signal is used as a bottom-up learning signal for the next, more precise prediction.

For implicit learning and the development of explicit knowledge, this means that any person has a certain expectation about their own performance in an SRTT, based on previous experiences with similar situations. The sequential material hidden in the task usually leads to behavior different from the expected one. These deviations from expectation will be used to adjust the metacognitive model (Esser & Haider, 2017). We assume that what is important for the development of explicit knowledge is the size of the prediction error and the strength of the a-priori hypothesis. When a person is participates in an SRTT training, they will have certain prior expectations on how their performance in this task should be like (e.g. determined by prior similar participation in experimental studies or by general knowledge how well their eye-hand coordination is in computer guided tasks). Many deviations from the expected metacognitive judgement of the situation can easily be used to adjust the model via this bottom-up error signal. For example, faster responses, fewer errors, increasing fluency are compatible with mere practice effects and only slight, gradual adjustments of the metacognitive models are the result. However, large prediction errors would lead to a stronger change of the metacognitive model. It might be functional to evaluate the situation and wonder whether a new, different model should be applied to the situation, instead of making rather drastic changes to the current model. So far, there is not much research on how metacognitive models are selected in a given situation and under which circumstances a model is replaced with a new or different one or when instead the current model will be adjusted. Collins and Frank (2013) suggested a Bayesian “context-task set” model. In this model, an inference is made in every single learning trial about whether the current task-set is still applicable to the current situation or whether there are yet unknown rules that should influence the task-set and therefore, a new model should be applied. This model also uses arbitrary context cues to determine whether the current situation is indicating a new, unknown task context or whether previously acquired metacognitive models should be used and adjusted.

We believe that this is the point, where the GWT has an important role. A large prediction error about the current metacognitive evaluation of the situation is a signal with a high likelihood of entering the GWS. The person becomes conscious about not experiencing what they expected to experience. This conscious state allows the involvement of other cognitive subsystems in order to evaluate whether a different model should be applied to the situation. Via hypothesis testing it can be detected that there is a sequence in the task. It is however also possible that other, seemingly more likely, explanations suffice to explain the unexpected experience and the search is terminated before the sequence is found (Haider & Frensch, 2005). What is important here is that we assume that this search process, if successful in finding the sequence, leads to a new, explicit representation of the sequence. It is not the acquired implicit knowledge itself which becomes conscious. This knowledge remains encapsulated.

These assumptions solve some of the formerly described problems behind explanations based solely on the GWT or HOTTs. Concerning the GWT, the UEH does not need to explain how an implicit representation can gain a signal-strength strong enough to win the competition against all other unconscious modules or how top-town attention can be directed to this encapsulated knowledge. This problem is solved because it is a conflict situation stemming from a strong conflict between the expected and the experienced metacognitive judgements which has a high likelihood of winning the competition against other parallel processes. Neither does it need to be explained whether and how different subsystems can interpret the information from formerly encapsulated modules without ever being synchronized with this information before. Instead, we assume that a new explicit, multi-modal representation of the sequence is created through active search processes.

Concerning the HOTT-based explanation of the emergence of explicit from implicit knowledge, we believe that the improvement the UEH offers lies in the assumption how implicit, first-order knowledge and explicit higher-order knowledge are related. An account where metacognitive judgements depend on a predictive learning process which does not only base its predictions on the first-order bottom-up signal, but also on heuristic cues, previous knowledge and experiences with similar situations, can help to explain different empirical findings. This includes, for example, premature responses (Haider & Frensch, 2009), changes in the underlying sequence (Schwager et al., 2012) and changes of experienced fluency (Esser & Haider, 2017). All these results are difficult to explain with a pure bottom-up mechanism relying on gradual strengthening. Furthermore, large prediction errors and the processes they are assumed to trigger fit the data suggesting that explicit knowledge seems to develop in a sudden moment of insight (Haider et al., 2011; Rose et al., 2010; Schwager et al. 2012; Wessel et al.,2012), rather than developing gradually. Assuming a gradual development of explicit knowledge requires further elaboration on how this gradual development matches subjective experience and how the transition from knowing that one knows to knowing the exact sequence occurs.

Conclusion and Future Directions

In the two main sections of the article, we aimed to provide a framework of implicit (sequence) learning. According to this framework, the current consciously available task set determines by which feature codes the interaction with the environment is represented and consequently which sequential features of a task are learned. Conversely, the implicitly learned knowledge can influence the current conscious representation of a task and thereby can lead to the development of explicit knowledge. We assume that the metacognitive model a person has about their own behavior in a given situation (e.g. how fast, how precise, how difficult or fluent a task should be) adapts to the task by comparing the predicted and the experienced metacognitive judgement in any given situation. The behavioral changes resulting from implicit learning may not fit the current metacognitive model (i.e. responses might suddenly be much slower than expected when the sequence is exchanged with new, random material). If so, this violation serves as a trigger to evaluate whether a new metacognitive model of the situation should be applied.

In further detail, in the first part of this article, we argued that the task set defines the features by which the actions and stimuli in the environment are coded. That means, we do not share the assumption of Keele et al. (2003) that implicit learning in the unidimensional system is fully independent of attention and that implicit knowledge is built within single dimensions that can be defined as processing modules (Keele et al., 2003; e.g. a module for shape processing, a module for response location processing). Instead and based on recent findings of Gaschler et al., (2012), Eberhardt et al. (2017), and Haider et al. (2018), we proposed to consider implicit learning to be based on modules that are temporarily bound together to an abstract feature code. Feature codes are integrated, multimodal structures comprising all the sensory and motor codes that have been associated with a certain perceivable event in the distal environment during a person's learning history (Hommel et al., 2001).

This perspective on implicit learning has two important implications: First, it loosens the strict distinction between stimulus- and response-related implicit learning. This especially helps to better understand, for instance, implicit location learning, which has been subject to a controversial debate of the perceptual- or motor-dependent processes behind it (Marcus et al., 2006; Willingham, Wells, Farrell, & Stemwedel, 2000). The second important implication here is that this view makes implicit learning more dependent on a person's task set. Implicit learning is still automatic and capacity independent in the way that it can still take place in parallel and in the presence of concurrent other tasks. Nevertheless, selective attention plays an important role in our conception, as implicit learning is assumed to be restricted by the abstract feature codes a person represents in their current task set. The suggestion of implicit learning being influenced by conscious representations of the task set fits well with more recent research from other areas of unconscious processing. For example, it has been shown that the behavioral effect of unconsciously presented primes can be influenced by a person's task set (Kiefer, 2012; Kiefer et al., 2012; Kouider & Dupoux, 2004; Kunde, Kiesel & Hoffmann, 2003).

Even though there are first empirical findings supporting this view, more research should be encouraged by this article. What is needed, are more experiments investigating the applicability of this concept to other dimensions of implicit learning. For example, whether listening to an auditive sequence can lead to the acquisition of a response-location or color sequence depending on the currently instructed task-set. Moreover, it should be tested whether not only the acquisition but also the expression of implicitly acquired knowledge can be influenced by a given task set. Can a task set control which implicitly acquired knowledge is recruited and which is not by activating or inhibiting certain feature codes? A third interesting aspect could be whether a task set can hinder the automatic acquisition of a sequence, when the task set directs attention to sequence irrelevant aspects of the task.

The second section of our article dealt with the emergence of explicit knowledge in an implicit learning situation. In this part, we deviate from the model of Keele et al. (2003). In their model, explicit learning is restricted to multidimensional learning; an assumption which has, to the best of our knowledge, never been directly tested and does not seem to fit empirical findings showing the option of explicit learning of any unidimensional sequence. Rather, our suggestions are in closer relation to models directly concerned with the emergence of explicit sequence knowledge. These models basically all rely on associative strengthening, either with relation to the GWT (Cleeremans & Jiménez, 2002) or to HOTTs (Cleeremans, 2014). It is a big advantage of these models that they are very parsimonious assumptions with reference to the two most important current frameworks on the emergence of conscious processing. Nevertheless, both accounts have some weaknesses. On the one hand, they include some theoretical assumptions in need of further elaboration. On the other hand, they do not fit the results from some studies on the emergence of explicit knowledge without further assumptions. Both aspects require further explanations with which these accounts might end up less simple. Also, both views have yet to be tested empirically in experiments, as so far, they only rely on simulation data (Cleeremans, 2014).

The UEH is not a new account of the emergence of explicit sequence knowledge and it has already brought up converging empirical evidence. Here, we elaborated the UEH a little further, bringing together important aspects from the HOTTs and the GWT. We assume that metacognitive learning processes are an important part in the emergence of explicit sequence knowledge. However, we do not think that simple associative, bottom-up learning processes will be sufficient to explain the empirical findings pointing to a role of unexpected events in the emergence of explicit knowledge (Esser & Haider, 2017; Gaschler, Marewski, Wenke & Frensch, 2014; Haider & Frensch, 2009; Rünger & Frensch, 2008; Schwager et al., 2012).

We propose that a predictive Bayesian account of metacognitive learning might be more promising. Within such a framework, it can be modelled that not only the current bottom-up first-order signal but also top-down factors, such as heuristic cues and previous experiences with similar situations, are the basis for a prediction of the metacognitive judgement of a given situation. This prediction is then compared to the experienced metacognitive judgement and its error signal is used as a learning signal for developing a more accurate metacognitive model. Small prediction errors might lead to a gradual change of the model. Yet, large prediction errors (in combination with strong a-priori hypotheses) can serve as a signal that the current model is not suitable for the given situation and a different model should be applied. This view corresponds to other current models of metacognitive judgements outside of the field of implicit sequence learning research, which also point to the shortcomings of prominent pure bottom-up models of metacognition (Collins & Frank, 2013; Fleming & Dew, 2017; Scott, Dienes, Barrett, Bor, & Seth, 2014).

Our proposal of explicitly integrating the role of metacognition in the UEH needs further experimental investigation: First, it should be tested whether the predicted metacognitive judgements can be manipulated not only by the strength of the first-order signal but also by differences between the expected and the actual experienced metacognitive judgment. Second, the size of the prediction error of metacognitive judgements as well as the strength of the a-priori hypothesis should be manipulated to test its relation to the emergence of explicit knowledge. Third, we proposed that large prediction errors serve as a conscious signal to trigger explicit search processes. These search processes are assumed to lead to a new explicit representation, independent of the implicit representation. Also this latter assumption needs to be tested empirically. In a study by Schwager et al. (2012), it has been investigated whether a change of the underlying sequence serves as an unexpected event, leading to conscious knowledge of the new, unpracticed sequence. Even though their results were not clear cut, their results provide an interesting experimental starting point, to test whether new representations are created in an explicit search process. This line of research should be pursued further.

Last but not least, we believe that the development of a comprehensive model of implicit learning and explicit sequence learning, is important for several reasons. Within the field of implicit learning research, there are so far only very few theoretical perspectives on the emergence of explicit knowledge. They all draw on metacognitive learning, respectively higher-order thought theories (Cleeremans, 2008, 2011; Dienes & Perner, 1999; Perruchet & Vinter, 2002; Scott & Dienes, 2010). We agree with the assumption that metacognitive learning processes are important to explain how explicit representations can develop in an implicit learning situation. However, we think that further development of these models will profit from current advances in research on how the relation between first- and higher-order knowledge can be modeled. Bayesian approaches seem promising insofar as they allow including the role of prior expectations about a task as well as the prediction errors that arise from the violation of these expectations.

A better understanding of the transition from implicit to explicit sequence knowledge can provide interesting contributions to the broad and difficult field of consciousness theories itself. Implicit learning paradigms create the unique experimental situation where unconscious knowledge does not need to be induced by week signal strength. By definition, implicitly acquired knowledge is knowledge that has never been conscious before and therefore has never been embedded into a broader context of different subsystems, which might change the way the information is processed unconsciously. Therefore, implicit sequence learning paradigms are well suited to investigate how the brain develops knowledge about its own internal states.

The development of metacognitive knowledge is a concern of many different and often separated research fields which all provide different contributions. For example, research on decision-making or on perception is governed by bottom-up signal-detection models (Galvin et al., 2003; Pleskac & Busemeyer, 2010), cue-utilization is prominent in memory research (Koriat, 2000, 2012, 2015) and models of evidence accumulation are often found in research on error-monitoring (Yeung & Summerfield, 2012). Implicit sequence learning paradigms can augment this research by providing additional opportunities (to the predominant priming paradigms) to manipulate the first-order signal strength, external cues, as well as the role of prior expectations and how these expectations develop over the course of learning.

Acknowledgements

We would like to thank Dr. Peter Frensch, Dr. Robert Gaschler, Dr. Michael Rose, and Dr. Christoph Stahl for constructive remarks which helped to improve an earlier version of the manuscript. This work was supported by the grant HA-5447/10-1 from the German Research Foundation.

References

Abrahamse, E. L., Jiménez, L., Verwey, W. B., & Clegg, B. A. (2010). Representing serial action and perception. *Psychonomical Bulletin and Review, 15*(5), 603-623. doi: 10.3758/PBR.17.5.603

Abrahamse, E. L., van der Lubbe, R. H., & Verwey, W. B. (2008). Asymmetrical learning between a tactile and visual serial RT task. *Quarterly Journal of Experimental Psychology, 61*(2), 210-217. doi:10.1080/17470210701566739

Baars, B. J. (1997). In the theatre of consciousness: Global Workspace Theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies, 4*(4), 292-309.

Baars, B. J. (2005). Global workspace theory of consciousness: towards a cognitive neuroscience of human experience? *Progress in Brain Research, 150*, 45-53. doi:10.1016/S0079-6123(05)50004-9

Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. Trends in Cognitive Sciences, 7(4), 166-172. doi:10.1016/S1364-6613(03)00056-1

Baars, B. J., Franklin, S., & Ramsøy, T. Z. (2013). Global workspace dynamics: cortical “binding and propagation” enables conscious contents. *Frontiers in Psychology, 4*. doi:10.3389/fpsyg.2013.00200

Barrett, A. B., Dienes, Z., & Seth, A. K. (2013). Measures of metacognition in signal-detection theoretic models. *Psychological Methods, 18*(4), 535-552. doi:10.1037/a003326

Bartels, A., & Zeki, S. (2000). The architecture of the colour centre in the human visual brain: new results and a review. *The European Journal of Neuroscience, 12*(1), 172-193. doi:10.1046/j.1460-9568.2000.00905.x

Block, N. (1995). On a confusion about the function of consciousness*. Behavioral and Brain Sciences, 18*(2), 227-287. doi:10.1017/S0140525X00038188

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences, 30*(5-6), 481-499. doi:10.1017/S0140525X07002786

Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience, 7*(4), 356-366. doi:10.3758/CABN.7.4.356

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 642-652. doi:10.1037//0033-295X.108.3.624

Chalmers, D. J. (1995). Facing up to the problems of consciousness. *Journal of Consciousness Studies, 2*(3), 200-219. doi:10.1093/acprof:oso/9780195311105.003.0001

Changeux, J. P., & Dehaene, S. (1989). Neuronal models of cognitive functions. *Cognition, 33*(1-2), 63-109. doi.org/10.1016/0010-0277(89)90006-1

Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes, 13*(2&3), 221-268. doi:10.1080/016909698386528

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences, 36*(3), 181-204. doi:10.1017/S0140525X12000477

Clark, A. & Karmiloff-Smith, A. (1993). The cognizer's innards: A psychological and philosophical perspective on the development of thought. *Mind and Language, 8*(4), 487-519. doi:10.1111/j.1468-0017.1993.tb00299.x

Cleeremans, A. (2008). Consciousness: the radical plasticity thesis. In R. Banerjee & B. K. Chakrabarti (Eds.), *Models of Brain and Mind. Physical, Computational and Psychological Approaches* (pp. 19-33). Amsterdam: Elsevier.

Cleeremans, A. (2011). The radical plasticity thesis: How the brain learns to be conscious. *Frontiers in Psychology, 2*. doi:10.3389/fpsyg.2011.00086

Cleeremans, A. (2014). Connecting conscious and unconscious processing. *Cognitive Science, 38*(6), 1286-1315. doi:10.1111/cogs.12149

Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness: An empirical, computational and philosophical consensus in the making?* (pp. 1–40). Hove: Psychology Press.

Cleeremans, A. & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120*(3), 235-253. doi:10.1037//0096-3445.120.3.235

Cleeremans, A., Timmermans, B., & Pasquali, A. (2007). Consciousness and metarepresentation: A computational sketch. *Neural Networks, 20*(9), 1032-1039. doi:10.1016/j.neunet.2007.09.011

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(1), 17-30. doi:10.1037/0278-7393.16.1.17

Colby, C. L., & Goldberg, M. E. (1999). Space and attention in partietal cortex. *Annual Review of Neuroscience, 22*, 319-349. doi10.1146/annurev.neuro.22.1.319

Collins, A. G., & Frank, M. J. (2013). Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychological Review, 120*(1), 190-229. doi:10.1037/a0030852

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: pitting abstract against stimulus-specific representations. *Psychological Science, 17*(10), 905-912. doi:10.1111/j.1467-9280.2006.01801.x

Coomans, D., Deroost, N., Vandenbossche, J., van den Bussche, E., & Soetens, E. (2012). Visuospatial perceptual sequence learning and eye movements. *Experimental Psychology, 59*(5), 279-185. doi:10.1027/1618-3169/a000155

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science, 19*(1), 51-57. doi:10.1177/0963721409359277

Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, C. (1999). Action for perception: A motor-visual attentional effect. *Journal of Experimental Psychology: Human Perception and Performance, 25*(6), 1673-1692. doi:10.1037/0096-1523.25.6.1673

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*(2), 200-227. doi:10.1016/j.neuron.2011.03.018

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79*(1-2), 1-37. doi:10.1016/S0010-0277(00)00123-2

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences, 10*(5), 204-211. doi:10.1016/j.tics.2006.03.007

Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness and could machines have it? *Science, 358*(6362), 48-492. doi:10.1126/science.aan8871

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology, 5*(10):e260. doi:10.1371/journal.pbio.0050260

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., & Slachevsky, A. (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain, 132*(9), 2531-2540. doi:10.1093/brain/awp111

Deroost, N., & Soetens, E. (2006). The role of response selection in sequence learning. *Quarterly Journal of Experimental Psychology, 59*(3), 449-456. Doi: 10.1080/17470210500462684

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review, 8*(2), 343-350. doi:10.3758/BF03196171

Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review, 4*(1), 3-23. doi:10.3758/BF03210769

Dienes, Z. & Longuet-Higgins, C. (2010). Can musical transformations be implicitly learned? *Cognitive Science, 28*(4), 531-558. doi:10.1207/s15516709cog2804_2

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences, 22*(5), 735-808. doi:org/10.1017/S0140525X99002186

Dienes, Z. & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge from judgement knowledge. Psychological Research, 69(5-6), 338-351. doi:10.1007/s00426-004-0208-3

Dreisbach, G., & Haider, H. (2009). How task representations guide attention: further evidence for the shielding function of task sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(2), 477-486. doi:10.1037/a0014647

Ebbinghaus, H. (1885). *Ueber das Gedaechtnis: Untersuchungen zur experimentellen Psychologie*. Leipzig: Duncker & Humbolt

Eberhardt, K., Esser, S., & Haider, H. (2017). Abstract feature codes: The building blocks of the implicit learning system. *Journal of Experimental Psychology: Human Perception and Performance, 43*(7), 1275-1290. doi:10.1037/xhp0000380

Eisenberg, M., Shmuelof, L., Vaadia, E., & Zohary, E. (2011). The representation of visual and motor aspects of reaching movements in the human motor cortex. *The Journal of Neuroscience, 31*(34), 12377-12384. doi:10.1523/JNEUROSCI.0824-11.2011

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 229-240. doi:10.1037//0096-1523.27.1.229

Esser, S. & Haider, H. (2017). The emergence of explicit knowledge in a serial reaction time task: The role of experienced fluency and strength of representation. *Frontiers in Psychology*, doi:10.3389/fpsyg.2017.00502

Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychological Review, 124*(1), 91-114. doi:10.1037/rev0000045

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience, 8*. doi:10.3389/fnhum.2014.00443

Francis, A. P., Schmidt, G. L., Carr, T. H., & Clegg, B. A. (2009). Incidental learning of abstract rules for non-dominant word orders. *Psychological Research, 73*(1), 60-74. doi:10.1007/s00426-008-013

Frensch, P. A., Haider, H., Rünger, D., Neugebauer, U., Voigt, S., & Werg, D. (2003). The route from implicit learning to awareness of what has been learned. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 335-366). New York, NY: John Benjamins Publishing Company.

Frensch, P. A., Lin, J., & Buchner, A. (1998). Learning versus behavioral expression of the learned: The effects of a secondary tone-counting task on implicit learning in the serial reaction task. *Psychological Research, 61*(2), 83-98. doi:10.1007/s004260050015

Frensch, P. A., Wenke, D., & Rünger, D. (1999). A secondary tone-counting task suppresses performance in the Serial Reaction Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(1), 260-274. doi:10.1037/0278-7393.25.1.260

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews: Neuroscience, 11*(2), 127-138. doi:10.1038/nrn2787

Galvin S.J., Podd J.V., Drga V., Whitmore J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review, 10*(4), 843-76. doi:10.3758/BF03196546

Gaschler, R., Frensch, P. A., Cohen, A., & Wenke, D. (2012). Implicit sequence learning based on instructed task set. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38(5)*, 1389-1407. doi:10.1037/a0028071

Gaschler, R., Marewski, J. N., Wenke, D., & Frensch, P. A. (2014). Transferring control demands across incidental learning tasks – Stronger sequence usage in serial reaction task after shortcut option in letter string checking. *Frontiers in Psychology, 5:1388. doi:10.3389/fpsyg.2014.01388*

*Gaschler, R., & Nattkemper, D. (2012). Instructed task demands and utilization of action effect anticipation. Frontiers in Psychology, 3:578. doi:*10.3389/fpsyg.2012.00578

Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of Cognitive Neuroscience, 7*(4), 497-510. doi:10.1162/jocn.1995.7.4.497

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. Psychological Review, 77(2), 73-99. doi:10.1037/h0028689

Gomez, R. L. (1997). Transfer and Complexity in Artificial Grammar Learning. *Cognitive Psychology, 33*(2), 154-207. doi:10.1006/cogp.1997.0654

Gomez, R. L., Gerken, L., & Schvaneveldt, R. W. (2000). The basis of transfer in artificial grammar learning. *Memory & Cognition, 28*(2), 253-263. doi:10.3758/BF03213804

Goodale, M. A. (2011). Transforming vision into action. Vision research, 51(13), 1567-1687. doi:10.1016/j.visres.2010.07.027

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*(1), 20-25. doi:10.1016/0166-2236(92)90344-8

Goschke, T., & Bolte, A. (2007). Implicit learning of semantic category sequences: Response-independent acquisition of abstract sequential regularities. Journal of Experimental Psychology-Learning Memory and Cognition, 33(2), 394-406. doi: 10.1037/0278-7393.33.2.394

Goschke, T., & Bolte, A. (2012). On the modularity of implicit sequence learning: Independent acquisition of spatial, symbolic, and manual sequences. *Cognitive Psychology, 65*(2), 284-320. doi:10.1016/j.cogpsych.2012.04.002

Haider, H., Eberhardt, K., Esser, S., & Rose, M. (2014). Implicit visual learning: How the task set modulates learning by determining the stimulus-response binding. *Consciousness and Cognition, 26(1),* 145-161. doi:10.1016/j.concog.2014.03.005

Haider, H., Eberhardt, K., Kunde, A., & Rose, M. (2012). Implicit visual learning and the expression of learning. *Consciousness and Cognition, 22*(1), 82-98. doi:10.1016/j.concog.2012.11.003

Haider, H., Eichler, A., & Lange, T. (2011). An old problem: How can we distinguish between conscious and unconscious knowledge acquired in an implicit learning task? *Consciousness* and Cognition, 20(3), 658-672. doi:10.1016/j.concog.2010.10.021

Haider, H., Esser, S., & Eberhardt, K. (2018). Feature codes in implicit sequence learning: perceived stimulus locations transfer to motor response locations. *Psychological Research*. doi:10.1007/s00426-018-0980-0

Haider, H., & Frensch, P. A. (2005). The generation of conscious awareness in an incidental learning situation. *Psychological Research, 69*(5-6), 399–411. doi:10.1007/s00426-004-0209-2

Haider, H., & Frensch, P. A. (2009). Conflicts between expected and actually performed behavior lead to verbal report of incidentally acquired sequential knowledge. *Psychological Research, 73*(6), 817–834. doi:10.1007/s00426-008-0199-6

Haider, H., Frensch, P. A., & Joram, D. (2005). Are strategy shifts caused by data-driven processes or by voluntary processes? *Consciousness and Cognition, 14*(3), 495-519. doi:10.1016/j.concog.2004.12.002

Hardwick, R. M., Rottschy, C., Miall, R. C., & Eickhoff, S. B. (2013). A quantitative meta-analysis and review of motor learning in the human brain. *Neuroimage, 67*(2), 283-297. doi:10.1016/j.neuroimage.2012.11.020

Hazeltine, E., Grafton, S. T., & Ivry, R. (1997). Attention and stimulus characteristics determine the locus of motor-sequence encoding. A PET study. *Brain, 120*(1), 123-140. doi:10.1093/brain/120.1.123

Herwig, A., & Waszak, F. (2009). Intention and attention in ideomotor learning. *The Quarterly Journal of Experimental Psychology, 62*(2), 219-227. doi:10.1080/17470210802373290

Herwig, A., & Waszak, F. (2012). Action-effect bindings and ideomotor learning in intention- and stimulus-based actions. *Frontiers in Psychology*, 3. doi:10.3389/fpsyg.2012.00444

Higuchi, Y., & Saiki, J. (2017). Implicit learning of spatial configuration occurs without eye movement. *Japanese Psychological Research, 59*(2), 122-132. doi:10.1007/s00426-012-0435-y

Hoffmann, J., Sebald, A., & Stöcker, C. (2001). Irrelevant response effects improve serial learning in serial reaction time tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(2), 470-482. doi:10.1037/0278-7393.27.2.470

Hommel, B. (1993). Inverting the Simon effect by intention. *Psychological Research, 55*(4), 270-279. doi:10.1007/BF00419687

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*(11), 494-500. doi:10.1016/j.tics.2004.08.007

Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychological Research, 73*(4), 512-526. doi:10.1007/s00426-009-0234-2

Hommel, B. (2015). The theory of event coding (TEC) as embodied-cognition framework. *Frontiers in Psychology, 6:1318. doi:*10.3389/fpsyg.2015.01318

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral & Brain Sciences, 24*(5), 849-878. doi:10.1007/s00426-009-0234-2

Howard, J. H., Howard, D. V., Dennis, N. A., & Kelly, A. J. (2008). Implicit learning of predictive relationships in three-element visual sequences by young and old adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1139-1157. doi:10.1037/a0012797

Howard, J. H., Mutter, S. A., & Howard, D. V. (1992). Serial pattern learning by event observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5)*,* 1029-1039. doi:10.1037/0278-7393.18.5.1029

Hoyndorf, A. & Haider, H. (2009). The “Not Letting Go” phenomenon: Accuracy instructions can impair behavioral and metacognitive effects of implicit learning processes. *Psychological Research, 73(5),* 695-706. doi:10.1007/s00426-008-0180-4

James, W. (1890). *The principles of psychology* (Vol. 2). New York: Dover Publications.

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning*. The Quarterly Journal of Experimental Psychology, 54*(4), 1105-1124. doi:10.1080/713756001

Jiang, S., Zhu, L., Guo, X., Ma, W., Yang, Z., Dienes, Z. (2012). Unconscious structural knowledge of tonal symmetry: Tang poetry redefines limits of implicit learning. *Consciousness and Cognition, 21*(1), 476-486. doi:10.1016/j.concog.2011.12.009

Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of experimental psychology: Learning, memory, and cognition, 25*(1), 236-259. doi:10.1037//0278-7393.25.1.236

Jiménez, L., Vaquero, J. M., & Lupiañez, J. (2006). Qualitative differences between implicit and explicit sequence learning. *Journal of experimental psychology: Learning, memory, and cognition, 32*(3), 475-490. doi:10.1037/0278-7393.32.3.475

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132*(1), 47-70. doi:10.1037/0096-3445.132.1.47

Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review, 110*(2), 352-339. doi:10.1037/0033-295X.110.2.316

Keele, S. W., & Summers, J. J. (1976). The structure of motor programs. In: G. E. Stelmach (Ed.), *Motor Control: Issues and Trends* (pp. 109-142). New York: Academic Press.

Kesner, R. P., & Rogers, J. (2004). An analysis of independence and interactions of brain substrates that subserve multiple attributes, memory systems, and underlying processes. *Neurobiology of Learning and Memory, 82*(3), 199-215. doi:10.1016/j.nlm.2004.05.007

Kiefer, M. (2012). Executive control over unconscious cognition: Attentional sensitization of unconscious information processing. *Frontiers in Human Neuroscience, 6*:61. doi: 10.3389/fnhum.2012.00061

Kiefer, M., Adams, S. C., & Zovko, M. (2012). Attentional sensitization of unconscious visual processing: Top-down influences on masked priming. *Advances in Cognitive Psychology*, *8*(1), 50–61. doi:10.2478/v10053-008-0102-4

Kim, D., Johnson, B. J., Gillespie, R. B., & Seidler, R. D. (2014). The effect of haptic cues on motor and perceptual based implicit learning. *Frontiers in Human Neuroscience, 8*:130. doi:10.3389/fnhum.2014.00130

Kinsbourne, M. (1996). What qualifies a representation for a role in consciousness? In J. D. Cohen & J. W. Schooler (eds.), *Scientific Approaches to the Study of Consciousness* (pp. 335-355). Hillsdale, NJ: Erlbaum.

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition, 9*(2), 149-171. doi:10.1006/ccog.2000.0433

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289-325). New York, NY: Cambridge University Press.

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review, 119*(1), 80-114. doi:10.1037/a0025648

Koriat, A. (2015). Knowing by doing: When metacognitive monitoring follows metacognitive control. In S. D. Lindsay, C. M. Kelley, A. P. Yonelinas, & H. L. Roediger III (Eds.), *Remembering: Attributions, Processes, and Control in Human Memory: Essays in honor of Larry Jacoby* (pp. 185-197). New York: Psychology Press.

Kouider, S., de Gardelle, V., Sackur, & J., Dupoux. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences, 14*(7), 301-307. doi:10.1016/j.tics.2010.04.006

Kouider, S., & Dupoux, E. (2004). Partial Awareness Creates the “Illusion” of Subliminal Semantic Priming. *Psychological Science, 15*(2), 75-81. doi:10.1111/j.0963-7214.2004.01502001.x

Kouider, S. & Faivre, N. (2017). Conscious and unconscious perception. In S. Schneider & M. Velmans (Eds.), *The Blackwell Companion in Consciousness, 2nd Edition* (pp. 855-864). New Jersey: Wiley-Blackwell

Kunde, W., Hoffmann, J., & Zellmann, P. (2002). The impact of anticipated action effects on action planning. *Acta Psychologica, 109*(2), 137-155. doi:10.1016/S0001-6918(01)00053-1

Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition, 88*(2), 223-242. doi:10.1016/S0010-0277(03)00023-4

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences, 10*(11), 494-501. doi:10.1016/j.tics.2006.09.001

Lau, H. C. (2008). A higher order Bayesian decision theory of consciousness. *Progress in Brain Research, 168*, 35-48. doi:10.1016/S0079-6123(07)68004-2

Lau, H. C., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences, 15*(8), 365-373. doi:10.1016/j.tics.2011.05.009

Ling, X., Li F., Qiao, F., Guo, X., & Dienes, Z. (2016). Fluency expresses implicit knowledge of tonal symmetry. *Frontiers in Psychology*, 7:57. doi: 10.3389/fpsyg.2016.00057

Longman, C. S., Lavric, A., & Monsell, S. (2016). The coupling between spatial attention and other components of task-set: A task-switching investigation. *Quarterly journal of experimental psychology, 69*(11), 2248-2275. doi:10.1080/17470218.2015.1115112

Newell, B. R., & Shanks, D. R. (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences, 37*(1), 1-61. doi:10.1017/S0140525X12003214

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1-32. doi:10.1016/0010-0285(87)90002-8

Maniscalco,B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition, 21*(1), 422-430. doi:10.1016/j.concog.2011.09.021

Maniscalco, B., & Lau, H. C. (2016). The signal processing architecture underlying subjective reports of sensory awareness. *Neuroscience of Consciousness, 2016*(1). doi:10.1093/nc/niw002

Marcus, D. J., Karatekin, C., & Markiewicz, S. (2006). Oculomotor evidence of sequence learning on the serial reaction time task. *Memory & Cognition, 34*(2), 420-432. doi:10.3758/BF03193419

Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22*(2), 350-364. doi:10.1037/0278-7393.22.2.350

Memelink, J., & Hommel, B. (2013). Intentional weighting: A basic principle in cognitive control. *Psychological Research, 77*(3), 249-259. doi:10.1007/s00426-012-0435-y

Musz, E., Weber, M. J., & Thompson-Schill, S. L. (2015). Visual statistical learning is not reliably modulated by selective attention to isolated events. *Attention, Perception, & Psychophysics, 77*(1), 78-96. doi:10.3758/s13414-014-0757-5

Overgaard, M. (2003). On the theoretical and methodological foundations for a science of consciousness. *Journal of Anthropological Psychology, 13*, 6-31.

Overgaard, M. (2017). The Status and Future of Consciousness Research. *Frontiers in Psychology, 8*:1719. doi:10.3389/fpsyg.2017.01719

Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development, 76*(2), 324-339. doi:10.1111/j.1467-8624.2005.00848.x

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. Journal of Experimental Psychology: General, 130(3), 401-426. doi:10.1037/0096-3445.130.3.401

Pasquali, A., Timmermans, B., & Cleeremans, A. (2010). Know thyself: Metacognitive networks and measures of consciousness. *Cognition, 117*(2), 182-190. doi:10.1016/j.cognition.2010.08.010

Perruchet, P. & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *TRENDS in Cognitive Sciences, 10*(5), 233-238. doi: 10.1016/j.tics.2006.03.006

Perruchet, P. & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences, 25*(3), 297-388. doi:10.1017/S0140525X02000067

Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience, 10(2)*, 257-261. doi:10.1038/nn1840

Persuh, M., LaRock, E., & Berger, J. (2018). Working memory and consciousness: The current state of play. *Frontiers in Human Neuroscience, 12*:27, doi:10.3389/fnhum.2018.00078

Peters, M. A., & Lau, H. (2015). Human observers have optimal introspective access to perceptual processes even for visually masked stimuli. *eLife, 3*. doi:10.7554/eLife.09651

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological Review, 117*(3), 864-901. doi: 10.1037/a0019737

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*(6), 855-863. doi:10.1016/S0022-5371(67)80149-X

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*(3), 219-235.

Remillard, G. (2009). Pure perceptual-based sequence learning: A role for visuospatial attention. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 35*(2), 528-541. doi:10.1037/a0014646

Remillard, G., & Clark, J. M. (2001). Implicit learning of first-, second-, and third-order transition probabilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(2), 483-498. doi:10.1037/0278-7393.27.2.483

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. Prokasky (Eds.), *Classical Conditioning II: Current theory and research* (pp-64-99). New York: Aplleton-Century Crofts.

Rose, M., Haider, H., & Büchel, C. (2010). The emergence of explicit memory during learning. *Cerebral Cortex, 20*(12), 2787-2797. doi:10.1093/cercor/bhq02

Rose, M., Haider, H., Weiller, C., & Büchel, C. (2002). The role of medial temporal lobe structures in implicit learning: An event-related fMRI study. *Neuron, 36* (6), 1221-1231. doi:10.1016/S0896-6273(02)01105-4

Rosenthal, D. (1997). A theory of consciousness. In N. Block, O. Flanagan, G. Güzeldere (Eds.), *The nature of consciousness: Philosophical debates* (pp. 729-753). Cambridge, MA: MIT Press.

Rosenthal, D. (2012). Higher-order awareness, misrepresentation and function. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 367*(1594), 1424-1438. doi:10.1098/rstb.2011.0353

Rothkirch, M., & Hesselmann, G. (2017). What we talk about when we talk about unconscious processing – a plea for best practice. *Frontiers in Psychology, 8*:835. doi:10.3389/fpsyg.2017.00835

Rünger, D. (2012). How sequence learning creates explicit knowledge: The role of response–stimulus interval. *Psychological Research, 76(5),* 579–590. doi: 10.1007/s00426-011-0367-y

Rünger, D., & Frensch, P. A. (2008). How incidental sequence learning creates reportable knowledge: The role of unexpected events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1011-1026. doi:10.1037/a0012942

Sakai, K. (2008). Task set and prefrontal cortex. *Annual review of neuroscience, 31*(1), 219-245. doi:10.1146/annurev.neuro.31.060407.125642

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: Is one measure better than the other? *Consciousness and Cognition, 19*(4), 1069-1078. doi:10.1016/j.concog.2009.12.013

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1-66. doi:10.1037/0033-295X.84.1.1

Schwager, S., & Hagendorf, H. (2009). Goal-directed access to mental objects in working memory: The role of task-specific feature retrieval. *Memory & Cognition, 37*(8), 1103-1119. doi:10.3758/MC.37.8.1103

Schwager, S., Rünger, D., Gaschler, R., & Frensch, P. A. (2012). Data-driven sequence learning or search: What are the prerequisites for the generation of explicit sequence knowledge? *Advances in Cognitive Psychology, 8*(2), 132-143. doi:10.2478/v10053-008-0110-4

Schwarb, H., & Schumacher, E. H. (2010). Implicit sequence learning is represented by stimulus-response rules. *Memory & Cognition, 38*(6), 677-688. doi:10.3758/MC.38.6.67

Scott, R. B., & Dienes, Z. (2010). Prior familiarity with components enhances unconscious learning of relations. *Consciousness & Cognition, 19*(1), 413-418. doi: 10.1016/j.concog.2009.12.012

Scott, R. B., Dienes, Z., Barrett, A. B., Bor, D., & Seth, A. K. (2014). *Blind Insight*: Metacognitive Discrimination Despite Chance Task Performance. *Psychological Science*, *25*(12), 2199–2208. doi:10.1177/0956797614553944

Shanahan, M., & Baars, B. (2005). Applying the global workspace theory to the frame problem. *Cognition, 98*(2), 157-176. doi:10.1016/j.cognition.2004.11.007

Shanks, D. R., & St John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17*(3), 367-447. doi:10.1017/S0140525X00035032

Sherman, M. T., Seth, A. K., Barrett, A. B., & Kanai, R. (2015). Prior expectations facilitate metacognition for perceptual decision. *Consciousness & Cognition, 35*, 53-65. doi:10.1016/j.concog.2015.04.015

Sheremata, S. L., & Shomstein, S. (2017). Task set induces dynamic reallocation of resoureces in visual short-term memory. *Psychonomic Bulletin & Review, 24*(4), 1113-1120. doi:10.3758/s13423-016-1181-8

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127-190. doi:10.1037/0033-295X.84.2.127

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin, 136*(6), 943-974. doi:10.1037/a0020541

Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia, 45*(12), 2802-2810. doi:10.1016/j.neuropsychologia.2007.05.002

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51*(3), 300-304. doi:10.1037/h0020586

Simons, D. J., Franconeri, S. L., & Reimer, R. L. (2000). Change blindness in the absence of visual disruption. *Perception, 29*(10), 1143-1154. doi:10.1068/p3104

Stöcker, C., Sebald, A., & Hoffmann, J. (2003). The influence of response-effect compatibility in a serial reaction time task. *The Quarterly Journal of Experimental Psychology, 56*(4), 685-703. doi:10.1080/02724980244000585

Tanaka, K., & Watanabe, K. (2017). Explicit instruction of rules interferes with visuomotor skill transfer. *Experimental Brain Research, 235*(6), 1-12. doi:0.1007/s00221-017-4933-4

Tubau, E., López-Moliner, J., & Hommel, B. (2007). Modes of executive control in sequence learning: from stimulus-based to plan-based control. *Journal of Experimental Psychology: General, 136*(1), 43-63. doi:1037/0096-3445.136.1.43

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience, 21*(10), 1934-1945. doi:10.1162/jocn.2009.21131

Tzvi, E., Münte, T. F., & Krämer, U. M. (2014). Delineating the cortico-striatal-cerebellar network in implicit motor sequence learning. *Neuroimage, 94*, 222-230. doi:10.1016/j.neuroimage.2014.03.004.

van Gaal, S., de Lange, F. P., Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making*. Frontiers in human neuroscience, 6*:121. doi:10.3389/fnhum.2012.00121

van Gaal, S., Naccache, L., Meuwese, J.D.I., van Loon, A.M., Leighton, A., Cohen, L., Dehaene, S. (2014) Can the meaning of multiple unconscious words be integrated unconsciously? *Philosophical Transactions of the Royal Society: B, 369(161)* doi:10.1098/rstb.2013.0212

Weiermann, B., & Meier, B. (2012). Implicit task sequence learning with auditory stimuli. *Journal of Cognitive Psychology, 24*(4), 468-475. doi:10.1080/20445911.2011.653339

Wenke, D., & Frensch, P. A. (2005). The Influence of Task Instruction on Action Coding: Constraint Setting or Direct Coding*? Journal of Experimental Psychology: Human Perception and Performance, 31*(4), 803-819. doi:10.1037/0096-1523.31.4.803

Wessel, J., Haider, H., & Rose, M. (2012). The transition from implicit to explicit representations in incidental learning situations: more evidence from high-frequency EEG coupling. *Experimental Brain Research, 217*(1), 153-162. doi:10.1007/s00221-011-2982-7

Whittlesea, B. W. A. (2002). Two routes to remembering (and another to remembering not*). Journal of Experimental Psychology: General, 131*(3), 325-348. doi:10.1037//0096-3445.131.3.325

Whittlesea, B. W., & Williams, L. D. (2000). The source of feelings of familiarity: the discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 547-565. doi: 10.1037//0278-7393.26.3.547

Wilbert, J. & Haider, H. (2012). The subjective experience of committed errors and the Discrepancy-Attribution hypothesis. *Acta Psychologica, 139*(2), 370-381. doi:10.1016/j.actpsy.2011.11.010

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105*(3), 558-584. doi:10.1037/0033-295X.105.3.558

Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural and declarative knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(6)*,* 1047-1060. doi:10.1037/0278-7393.15.6.1047

Willingham, D. B., Wells, L. A., Farrell, J. M., & Stemwedel, M. E. (2000). Implicit motor sequence learning is represented in response locations. Memory & Cognition, 28(3), 366-375. doi:10.3758/BF03198552

Wise, S. P., & Willingham, D. T. (2009). Motor Skill Learning. In: L. R. Squire (ed.) *Encyclopedia of Neuroscience*, *vol. 5*, pp. 1057-1066. Oxford: Academic Press.

Wykowska, A., Schubö, A., & Hommel, B. (2009). How you move is what you see: Action planning biases selection in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 35*(6), 1755-1769. doi:10.1037/a0016798

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences, 367*(1594), 1310-1321. doi:10.1098/rstb.2011.0416

Zach, N., Inbar, D., Grinvald, Y., Bergman, H., & Vaadia, E. (2008). Emergence of novel representations in the primary motor cortex and premotor neurons during associative learning. *The Journal of Neuroscience, 28*(38), 9545-9556. doi:10.1523/JNEUROSCI.1965-08.2008

Ziessler, M. (1998). Response-effect learning as a major component of implicit serial learning*. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(4), 962-978. doi:10.1037/0278-7393.24.4.962

Ziessler, M., & Nattkemper, D. (2001). Learning of event sequences is based on response-effect learning: further evidence from a serial reaction time task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(3), 595-613. doi: 10.1037/0278-7393.27.3.595

Zimmer, H. D. (2008). Visual and spatial working memory: From boxes to networks*. Neuroscience and Biobehavioral Reviews,* 32(8), 1373-1395. doi:10.1016/j.neubiorev.2008.05.01