Goal relevance and artificial grammar learning
Baruch Eitam1; Yaacov Schul1; Ran R. Hassin*
* The Hebrew University, Jerusalem, Israel

First published on: 03 December 2008
Short article

Goal relevance and artificial grammar learning

Baruch Eitam, Yaacov Schul, and Ran R. Hassin
The Hebrew University, Jerusalem, Israel

This investigation used a newly developed artificial grammar learning (AGL) paradigm in which participants were exposed to sequences of stimuli that varied in two dimensions (colours and letters) that were superimposed on each other. Variation within each dimension was determined by a different grammar. The results of two studies strongly suggest that implicit learning in AGL depends on the goal relevance of the to-be-learned dimension. Specifically, when only one of the two stimulus dimensions was relevant for their task (Experiment 1) participants learned the structure underlying the relevant, but not that of the irrelevant dimension. However, when both dimensions were relevant, both structures were learned (Experiment 2). These findings suggest that implicit learning occurs only in dimensions to which we are attuned. Based on the present results and on those of Eitam, Hassin, and Schul (2008) we suggest that focusing on goal relevance may provide new insights into the mechanisms underlying implicit learning.

Keywords: Implicit learning; Goal-directed cognition; Nonconscious processes; Motivation; Selective attention.

Implicit learning refers to people’s tendency to acquire complex regularities or patterns without intention or awareness (e.g., Frensch, 1998; Reber, 1967, but see Dulany, Carlson, & Dewey, 1984; Shanks & St. John, 1994). Given consciousness’s limited capacity (Kahneman, 1973), implicit learning is an important tool in the toolbox of organisms that live in constantly changing, complex environments. And indeed, 40 years after Reber (1967) introduced the topic to the cognitive sciences, implicit learning is still in central stage—maybe now more than ever before.

Despite decades of research and impressive advancements (for overviews, see Cleeremans, Destrebecqz, & Boyer, 1998; Dienes, 2008; Frensch & Runger, 2003; Pothos, 2007; Reber, 1993), many fundamental questions regarding implicit learning remain unanswered. In this paper we focus on a subtype of implicit...
learning—artificial grammar learning (AGL; Reber, 1967)—and address one of these issues: the role of attention.

Attention and implicit learning

Previous investigations examined the effects of directing attention (to and from stimuli) in other types of implicit learning. They showed that although learning occurs when attention is explicitly directed to the structured stimuli (Cohen, Ivry, & Keele, 1990; Jiang & Chun, 2001; Jimenez & Mendez, 1999; Nissen & Bullemer, 1987; Turk-Browne, Junge, & Scholl, 2005), it may also occur when attention is not explicitly directed towards them (Cock, Berry, & Buchner, 2002; Deroost, Zeischka, & Soetens, 2008; Jiang & Leung, 2005; Schmidtke & Heuer, 1997). Currently, it is unclear when the latter type of learning does occur. We return to this issue in the General Discussion.

In this paper we turn the spotlight on the role of attention in implicit AGL (for a similar exploration, see Tanaka, Kiyokawa, Yamada, Dienes, & Shigemasu, in press). To do so, we introduce a novel and potentially very rich AGL paradigm. The results of two experiments strongly suggest that directing attention to the relevant dimension of the stimuli is a prerequisite for implicit AGL. In the General Discussion section we suggest that the framework adopted here may help us make sense of the seemingly contradicting findings regarding the role of attention in other forms of implicit learning (see above paragraph).

The 2AGL paradigm

In the prototypical AGL experiment (e.g., Reber, 1967) participants are presented with sequences of stimuli that vary on one dimension. Most often, these are letter sequences (e.g., XXRSSVM or XMMXM). Participants are typically instructed to merely memorize the sequences. Unbeknownst to them, the stimuli within sequences are determined by a rather complex finite artificial grammar. The critical question is whether participants are able to implicitly learn the regularities that this grammar dictates.

To assess learning of the grammar, participants perform a “well-formedness” test (Reber, 1967) in which they are presented with novel grammatical and nongrammatical letter sequences, which are to be categorized as grammatical or not. The modal finding is that participants correctly classify novel sequences significantly above chance, thus suggesting that they learned the grammar (or its related regularities). The exact form of the acquired knowledge, the degree to which learning is incidental, and the extent to which participants are aware of the grammar are hotly debated issues that are beyond the scope of the current paper (for a recent summary see Pothos, 2007).

We developed a new paradigm (2AGL) that departs from the original in that participants are simultaneously presented with two different sequences that are superimposed on each other (see A in Figure 1). Specifically, each stimulus
varied in (a) the letters and (b) the colours that appear in the letters’ background. Letters were generated by one artificial grammar, and colours were generated by a different one (see B in Figure 1).

In the 2AGL paradigm participants might be asked to process (e.g., memorize) the colours, the letters, or both. Thus, while both dimensions are in the focus of attention (one cannot look at the colours but not the letters and vice versa), the paradigm allows us to choose whether to render one or both dimensions task relevant. Note further that it is generally assumed that rendering a dimension task relevant leads to the direction of attention towards this dimension (e.g., Jiang & Chun, 2001; Jiang & Leung, 2005; Jimenez & Mendez, 1999). Thus, a given dimension in 2AGL may be in the focus of attention and selectively attended, or in the focus of attention and not selectively attended. In order to refrain from confusing these two meanings of attention, we refer to the dimension that participants are instructed to memorize as the relevant or goal-relevant dimension (in the sense that it is relevant to the focal task goal).

In the current terminology, then, this paper examines whether goal-irrelevant grammars (or their related regularities) can be implicitly learned. To take an example, assume that participants performing the 2AGL task are asked to memorize the colours. Based on past research, we know that they should learn the grammar associated with colours. The question we pose here is whether they will also learn the grammar underlying the letter sequences.

EXPERIMENT 1

Method

Participants

A total of 47 Hebrew University undergraduates (20 females, mean age = 22.8 years, SD = 2.8) participated in the experiment in exchange for course credit or pay. Participants were run individually and were randomly assigned to one of four conditions.

Procedure

Participants first performed the training phase followed by a surprise “well-formedness” test (Reber, 1967). In the test phase participants were requested to classify novel sequences that varied in one dimension only, as either grammatical or not (see below).

Training phase. In each trial of the training phase a sequence of letters and colours appeared in the centre of the screen for 7 seconds (cf. Dienes, Altman, Kwan, & Goode, 1995). The letters were always black, and each of them was encased in a coloured rectangle (see A in Figure 1). The sequences of letters were determined by one grammar and the sequences of colours by another (see Figure 1). The two grammars were similar in their complexity, each containing eight potential paths between nodes and two recursions (see Figure 1).

A total of 32 different sequences were generated from each grammar (see Appendix A). The two sets of unidimensional stimuli were randomly combined, with the restriction of the colour and letter sequences having equal length, to create a set of 32 bidimensional training stimuli. Each training stimulus was repeated three times in the course of training. The order of stimuli was randomly determined for each participant, with the provision that the same stimulus did not appear on consecutive trials.

One group of participants was instructed to memorize the sequences of letters; another, the sequences of colours. Participants were informed that “sometimes, your memory would be probed”, and indeed following (randomly determined) 36 of the 96 trials participants were asked to recall the previous sequence. Specifically, participants were requested to type in the sequence

---

The choice of colouring the background, rather than the letters themselves, was based on the recent suggestion of Turk-Browne, Isola, Junge, and Scholl (2008), who proposed that implicit learning of individual dimensions is impaired when the dimensions are perceived as belonging to the same object (as is the case with coloured letters).
of letters or colours (depending on the condition they were assigned to) into an open-ended response box that appeared immediately after the sequence disappeared.

**Test phase.** Following training, and depending on their experimental condition, participants were informed that the letter/colour sequences adhered to a very complex set of rules. They were then asked to categorize 40 new sequences as either grammatical or not. A total of 20 test stimuli were grammatical, and 20 violated the grammar. The nongrammatical stimuli started and ended with a grammatically correct letter, and each violated the grammar twice (see Appendix B). Moreover, the grammatical test items of one dimension were nongrammatical for the other dimension and vice versa (Dienes et al., 1995). For example, the test stimulus “XXRVM” is grammatical according to the letter grammar but nongrammatical according to the colour grammar.

To foster application of implicit knowledge participants were warned that the rules were very complex, and they were encouraged to use their gut feeling when classifying the novel strings (see, e.g., Dienes et al., 1995). Sequences in the test phase were unidimensional, so that letter sequences appeared with no colour background, and colour sequences appeared with no letter content (see C in Figure 1). The nongrammatical colour sequences were generated from the grammar of the letters, and vice versa.

The test sequences were presented in the centre of the screen until participants responded. No feedback was given, in order to minimize explicit learning during the test phase.

**Design**

The experiment had a 2 (test: dimension of tested grammar, letters vs. colours) × 2 (relevance: whether the tested grammar was relevant during training, relevant vs. irrelevant) between-participants design.

All participants were exposed to exactly the same training sequences. They differed with respect to what they were instructed to memorize (letters or colours) and the stimuli they were tested on (letters or colours). Specifically, about half of the participants were instructed to memorize the letter sequences and half the colour sequences (see Table 1 for exact numbers of participants). Orthogonally, about half of the participants were tested on letter sequences, and half were tested on colour sequences. Accordingly, when considering the training and testing sequences jointly, 24 participants were tested on the grammar of the relevant dimension, and 23 were tested on the grammar of the irrelevant dimension.

**Results**

Our main hypothesis was that a grammar would be learned more successfully when it was goal relevant than when it was not. A two-way analysis of variance (ANOVA) with test (letters vs. colours) and relevance (relevant vs. irrelevant) as between-participants factors was performed. Our primary concern is with the main effect of relevance—namely, would the grammar associated with colours/letters be better learned when it was relevant?

The results presented in Table 1 provide an affirmative answer. The grammar of colours was

<table>
<thead>
<tr>
<th>Type of grammar tested</th>
<th>Relevant</th>
<th>Irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct</td>
<td>SD</td>
</tr>
<tr>
<td>Colour</td>
<td>62.1</td>
<td>10.7</td>
</tr>
<tr>
<td>Letter</td>
<td>71.5</td>
<td>14.6</td>
</tr>
</tbody>
</table>

*Note:* Correct = mean percentage of correct classifications. N = number of participants.
learned more successfully when colour was made relevant through the memorization task (62.1% vs. 47.3%), and the grammar of the letters was learned more successfully when letters were relevant (71.5% vs. 50.7%), leading to a significant main effect of relevance, $F(1, 43) = 30.9, p < .05$, partial eta squared = .42. This effect was not qualified by an interaction with the test factor ($F < 1$).

In addition, the main effect of test type was significant, $F(1, 43) = 4.00, p = .05$, partial eta squared = .09, indicating that the grammar underlying letters was better learned than that of colours. This effect was not hypothesized, and it may reflect factors such as the dimension’s saliency and discriminability, amongst others.

Analyses of the simple effect of relevance, within each type of test stimuli, revealed that it was significant both for colours, $t(21) = 4.06, p < .01$, partial eta squared = .44, and letters, $t(22) = 4.0, p < .01$, partial eta squared = .42. Finally, analyses of the individual cells revealed that while performance of participants who were tested on their relevant grammar was significantly better than chance (both $t$s > 3.7, both $p$s < .01), performance of those who were tested on grammars of the irrelevant dimension did not significantly differ from chance-level performance (both $p$s > .16). To conclude, the results of the Experiment 1 are consistent with the hypothesis that participants only learn the grammar that is relevant to their goals.

**Implicitness of learning**

Following the well-formedness test we assessed participants’ intentions by asking them “to what extent did you try to learn the rules that determined the sequences of stimuli you saw?” (1 = “not at all”–9 = “tried very hard”).

The responses revealed that participants’ conscious intention to learn did not significantly differ as a function of relevance ($M = 4.32, SD = 2.86$, when relevant, and $M = 3.72, SD = 2.37$, when irrelevant), $t(42) < 1$. Thus, the superior performance of participants in the relevant condition cannot be explained by attributing a stronger intention to learn.

Furthermore, participants’ intentions to learn did not significantly correlate with performance, $r(44) = .21, p = .17$.

These two findings lead us to conclude, then, that learning was implicit, in the sense that it was largely unintended.

**Discussion**

Participants in Experiment 1 could, in principle, learn the two grammars, as both the letters and colours were in their focus of attention. Still, only the grammar underlying the relevant dimension was learned. We interpret this to mean that AGL is selective, in that people tend to implicitly learn only information that is relevant to their current goals (which was, in our experiment, memorizing sequences of letters or colours).

Two possible objections to our interpretation of these results may be raised. First, one may argue that the knowledge of the irrelevant grammar was acquired but not applied (see Jiang & Leung, 2005). The results of Experiment 2 suggest that this is not the case, and we return to discussing this objection after presenting them.

Second, it could be argued that given the complexity of the two grammars, learning the relevant grammar exhausted the capacity for implicit learning (Dienes, Broadbent, & Berry, 1991; but see Reber, 1993). Accordingly, one may argue that the irrelevant grammar was not learned because participants simply cannot learn these two grammars simultaneously, regardless of their relevance.

Experiment 2 was designed to investigate the latter objection. To do so, we rendered the two dimensions relevant by asking participants to memorize both letter and colour sequences. We examined whether under these conditions participants can learn both grammars.

**EXPERIMENT 2**

**Method**

**Participants**

A total of 17 Hebrew University undergraduates (12 females, mean age = 22.35 years, $SD = 2.59$)
participated in the experiment in exchange for course credit or pay. Participants were run individually.

Procedure
The procedure was identical to that of Experiment 1 save two changes:

First, we rendered both dimensions goal relevant. This was achieved by instructing participants to memorize both the letters and the colours. Participants were informed that they would occasionally be asked to reproduce either a letter sequence or a colour sequence. Indeed, there were 18 probes of colour sequences and 18 probes of letter sequences, which appeared randomly throughout the 96 learning trials.

Secondly, all participants were tested on their learning of the two grammars. They were given 40 novel sequences of colours and 40 novel sequences of letters (the order was counterbalanced) and were asked to categorize each sequence as grammatical or not. As in Experiment 1, half of the sequences of each type of test stimulus were grammatical.

Results and discussion
The data of one participant who failed to follow instructions were discarded. The results clearly indicate that despite the complexity of the two grammars, participants successfully learned both of them. Participants correctly classified 65.3% ($SD = 11.4$) of the novel letter sequences and 55% ($SD = 8.31$) of the novel colour sequences, $\tau (15) = 5.4$, $p < .01$, partial eta squared = .66, and $\tau (15) = 2.4$, $p < .05$, partial eta squared = .28, respectively.

These results clearly show that participants can simultaneously learn, and then apply, two discrete grammars, thus ruling out the idea that participants in Experiment 1 could not learn both grammars simultaneously. These findings also rule out the other alternative explanation of the results of Experiment 1, according to which both grammars were learned but one of them could not be applied. Thus, the findings of Experiment 2 strongly suggest that participants in Experiment 1 failed to learn the second grammar because it was irrelevant to their goal.

It is worthwhile to note that while we are not the first to examine learning of two grammars in AGL, all previous research introduced the grammars sequentially (Conway & Christiansen, 2006; Dienes et al., 1995). Related forms of implicit learning (e.g., using the Serial Reaction Time [SRT] paradigm) did show simultaneous learning, but these demonstrations were limited to the special case in which one of the dimensions was spatial location, a dimension that may be very unique (cf. Mayr, 1996).

Let us make three final comments. First, comparing the learning scores of the two dimensions revealed that, as in Experiment 1, the grammar underlying letter sequences was better learned, $\tau (15) = 3.3$, $p < .01$, partial eta squared = .42. This finding was not hypothesized and may reflect various factors that are not the focus of the current investigation (see the Discussion section of Experiment 1).

A second observation is that the learning in Experiment 2 seems to be weaker than that in the Experiment 1. Acknowledging the problems associated with comparing results across two experiments, this observation seems to support the position that implicit learning is, to some degree, resource demanding (Dienes et al., 1995; Shanks, Rowland, & Ranger, 2005; but see Jimenez & Mendez, 1999; Mayr, 1996). The weaker learning in Experiment 2 may also reflect the smaller number of probes for each dimension used during the encoding stage (e.g., Dienes et al., 1991; but see Rausei, Makovski, & Jiang, 2007).

Finally, learning in this experiment is considered to be incidental, in the sense that it is not explicitly required. As both dimensions were relevant in Experiment 2, participants’ intention to learn the structures underlying relevant and irrelevant dimensions could not be compared (see Experiment 1). It seems safe to assume, however, that since the cognitive load in Experiment 2 was higher than that in Experiment 1, participants’

---

2 We thank Natacha Deroost for this suggestion.
intention to learn should have been even lower, thus essentially replicating the results of Experiment 1. Therefore, we regard the learning in Experiment 2 as implicit.

**GENERAL DISCUSSION**

Together the results of Experiments 1 and 2 provide compelling evidence for the selectivity of implicit AGL: Learning occurs for goal-relevant dimensions, and it does not occur for goal-irrelevant ones. These results underscore the ability of the mental system to selectively, yet unintentionally and unconsciously, extract relevant structures from our environments.

The idea that implicit learning is goal dependent (see also Frensch & Runger, 2003; Schmidtke & Heuer, 1997) may shed new light on an existing puzzle in the implicit learning literature. The puzzle was succinctly described in the Introduction: Some experiments find that implicit learning (in non-AGL paradigms) occurs without directing participants’ attention to the to-be-learned stimuli (e.g., Cock et al., 2002; Deroost et al., 2008; Jiang & Leung, 2005; Schmidtke & Heuer, 1997), whereas others find the opposite pattern (e.g., Jiang & Chun, 2001; Jimenez & Mendez, 1999; Tanaka et al., in press; Turk-Browne et al., 2005). Our theoretical analysis predicts when the grammar underlying unattended dimensions should be learned. If implicit learning is, as we suggest, goal dependent, then one would expect that learning in unattended dimensions would occur only when the to-be-learned structure is goal relevant.

Indeed, in cases in which implicit learning of the structure underlying unattended (i.e., “irrelevant”) stimuli occurred, the structure was relevant to participants’ performance (Cock et al., 2002; Deroost et al., 2008; Jiang & Leung, 2005; Schmidtke & Heuer, 1997). To take an example Cock et al. (2002) used an SRT task and instructed participants to attend to stimuli of one colour and ignore those of the other colour. Importantly, “irrelevant” and “relevant” stimuli could not appear in the same location. As participants’ task was to respond quickly and accurately to the appearance of the “relevant” stimuli, knowledge of the structure underlying the appearance of the “irrelevant” stimuli could have improved their performance by informing them where targets will not appear (ruling out 25% of the possible locations). Clearly, then, learning this structure was goal relevant. On the other hand, when stimuli that were not selectively attended were irrelevant for performance (and were not otherwise perceived as relevant by the system) their structure was not learned. For example, in the study by Turk-Browne and his colleagues (2005), participants performed a challenging n-back task on one set of shapes and did nothing with another set of shapes. Additionally, the structure of the irrelevant stimuli was not correlated with that of the relevant stimuli and thus could not assist performance. Accordingly, only the structure underlying the relevant set of shapes was learned.

Further research is needed to establish the conditions under which goal relevance is both necessary and sufficient for learning to occur. For example, in Jimenez and Mendez’s (1999) version of the SRT paradigm participants failed to learn a simple relationship across dimensions (i.e., how the shape of the current stimulus predicts the next trial’s location) although such learning would be beneficial to their performance and, thus, is goal relevant. Jimenez and Mendez’s study differed from our own in that the to-be-learned contingency occurred across dimensions while in our study it occurred within dimension. We speculate that learning contingencies of unrelated dimensions may pose a greater challenge to the mental system (see also Hoffmann & Sebald, 2005), which might be overcome if each of the dimensions becomes goal relevant. Another condition in which learning of a secondary (nonfocal) dimension might not occur although it is goal relevant is when the primary dimension lacks any structure (e.g., appears in random order). Such conditions may lead to a “shut-down” of the contingency learning process and thus impair consequent structure learning of both focal and nonfocal dimensions (Deroost et al., 2008, Exp. 3; Junge, Scholl, & Chun, 2007).
The suggestion that implicit learning is goal directed should be conceptually differentiated from a related position of Whittlesea and his colleagues (cf. Whittlesea & Wright, 1997; Wright & Whittlesea, 1998). In a series of studies these authors have convincingly demonstrated that the explicit processing demands of one task (“training phase”) affect how stimuli are represented in the participants’ mind and thus, accidentally, their performance on a subsequent test of structure. Wright and Whittlesea’s conclusion from these findings is that “there are not two forms of learning—‘explicit’ and ‘implicit’—but instead there is only learning, and that people may be aware or unaware of the implications of that learning for the future . . . ” (p. 403). Although the results of the two experiments presented here can be conceived as another example of the effects of processing demands on accidental acquisition of knowledge, the conceptualization we offer here is very different. Although structure learning can proceed incidentally (without intention) it is not at all accidental. Rather, implicit (as explicit) learning is a tool for attaining (current) goals; as such it is tuned by the current goals of the perceiver in the purpose of assisting her current goal pursuits.

To conclude, we suggest that implicit learning within multiple dimensions occurs when these dimensions are goal relevant (note that while instructing participants to attend to stimuli is a useful way of making stimuli goal relevant, it is not the only way in which a stimulus may become goal relevant). This simple idea—that implicit learning is related to participants’ goals—received previous support in studies that showed that non-conscious goal pursuit can improve implicit learning (Eitam, Hassin, & Schul, 2008).

REFERENCES


APPENDIX A

Training stimuli

(Grammars adapted from those used by Dienes, Altman, Kwan, & Goode, 1995).

“This item erroneously contained a violation of the letter grammar.”
APPENDIX B

Test stimuli

<table>
<thead>
<tr>
<th>Letter grammar</th>
<th>Colour grammar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical</td>
<td>Grammatical</td>
</tr>
<tr>
<td>xxrvvm</td>
<td>vvrxml</td>
</tr>
<tr>
<td>vsvsvm</td>
<td>vvrxml</td>
</tr>
<tr>
<td>xxrvsvm</td>
<td>xmsrrm</td>
</tr>
<tr>
<td>vsssvm</td>
<td>vsrsxm</td>
</tr>
<tr>
<td>xmmxmxm</td>
<td>xrrrm</td>
</tr>
<tr>
<td>xxrvsvm</td>
<td>vrrmsm</td>
</tr>
<tr>
<td>vsvsvm</td>
<td>xmsrrm</td>
</tr>
<tr>
<td>xmmxrvsvm</td>
<td>vsrsxm</td>
</tr>
<tr>
<td>xmmxrvsvm</td>
<td>vsrsxm</td>
</tr>
<tr>
<td>xxmsrvsvm</td>
<td>vsrsxm</td>
</tr>
<tr>
<td>xmsxrvsvm</td>
<td>vsrsxm</td>
</tr>
<tr>
<td>xmmxrvsvm</td>
<td>vsrsxm</td>
</tr>
<tr>
<td>xxmsrvsvm</td>
<td>vsrsxm</td>
</tr>
<tr>
<td>vmsrvsvm</td>
<td>xmsrrm</td>
</tr>
<tr>
<td>vsmsrvsvm</td>
<td>xmsrrm</td>
</tr>
<tr>
<td>xmsxmsrvvm</td>
<td>xmsrrm</td>
</tr>
<tr>
<td>xmsxmsrvvm</td>
<td>xmsrrm</td>
</tr>
</tbody>
</table>

(Grammars adapted from those used by Dienes, Altman, Kwan, & Goode, 1995).

bThis item erroneously contained a grammar violation and thus was analysed as nongrammatical.