



PAPER

The best time to acquire new skills: age-related differences in implicit sequence learning across the human lifespan

Karolina Janacsek,¹ József Fiser² and Dezso Nemeth¹

1. Institute of Psychology, University of Szeged, Hungary

2. Department of Cognitive Science, Central European University, Budapest, Hungary

Abstract

Implicit skill learning underlies obtaining not only motor, but also cognitive and social skills through the life of an individual. Yet, the ontogenetic changes in humans' implicit learning abilities have not yet been characterized, and, thus, their role in acquiring new knowledge efficiently during development is unknown. We investigated such learning across the lifespan, between 4 and 85 years of age with an implicit probabilistic sequence learning task, and we found that the difference in implicitly learning high- vs. low-probability events – measured by raw reaction time (RT) – exhibited a rapid decrement around age of 12. Accuracy and z-transformed data showed partially different developmental curves, suggesting a re-evaluation of analysis methods in developmental research. The decrement in raw RT differences supports an extension of the traditional two-stage lifespan skill acquisition model: in addition to a decline above the age 60 reported in earlier studies, sensitivity to raw probabilities and, therefore, acquiring new skills is significantly more effective until early adolescence than later in life. These results suggest that due to developmental changes in early adolescence, implicit skill learning processes undergo a marked shift in weighting raw probabilities vs. more complex interpretations of events, which, with appropriate timing, prove to be an optimal strategy for human skill learning.

Introduction

It is widely accepted that children should be introduced to sports, music or languages early in their life if they are to develop a high proficiency, because late learners seldom become true champions or elite musicians or gain command of a second language similar to that of a native speaker. These observations contradict traditional measures of the ability of factual learning of declarative memories, which showed that humans become increasingly better at many learning tasks up until their late twenties (Craik & Bialystok, 2006). However, an important component of developing new abilities is related to implicit unconscious statistical learning processes (Hikosaka, Nakamura, Sakai & Nakahara, 2002; Keele, Ivry, Mayr, Hazeltine & Heuer, 2003) that underlie the acquisition of not only motor but also cognitive and social skills (Doyon, Bellec, Amsel, Penhune, Monchi, Carrier, Lehericy & Benali, 2009; Hikosaka *et al.*, 2002; Lieberman, 2000; Poldrack, Sabb, Foerde, Tom, Asarnow, Bookheimer & Knowlton, 2005; Ullman, 2001). Thus, to understand complex skill acquisition, the characteristics of both explicit declarative and implicit learning, such as the differences in their efficiency across

the lifespan, must be clarified. In contrast to declarative memory (Tulving & Craik, 2000), the ontogenetic changes in humans' implicit learning abilities have not yet been comprehensively characterized, and, thus, their role in acquiring new knowledge efficiently during development is unknown. The main goal of our study was to examine age differences in implicit learning across the human lifespan using the same task for all groups.

The computational underpinnings and the neural substrates of these different kinds of learning mechanisms are also controversial (Henke, 2010). Explicit learning has been linked more closely to medial temporal lobes of the cortex (Dennis & Cabeza, 2011; Squire & Zola, 1996). By contrast, implicit skill learning often requires fine-tuning of the perceptual-motor system based on experience; therefore, most models of implicit skill learning emphasize the role of the basal ganglia and the cerebellum (Cohen, Pascual-Leone, Press & Robertson, 2005; Dennis & Cabeza, 2011; Doyon *et al.*, 2009; Hikosaka, Nakahara, Rand, Sakai, Lu, Nakamura, Miyachi & Doya, 1999; Hikosaka *et al.*, 2002), whereas the role of the hippocampus remains inconclusive (Albouy, Sterpenich, Balteau, Vandewalle, Desseilles, Dang-Vu, Darsaud, Ruby, Luppi, Degueldre, Peigneux,

Address for correspondence: Dezso Nemeth, Institute of Psychology, University of Szeged, Egyetem u. 2, 6722, Szeged, Hungary; e-mail: nemethd@edpsy.u-szeged.hu

The authors contributed equally to this work. The authors report no conflict of interest and have no financial disclosure.

Luxen & Maquet, 2008; Schendan, Searl, Melrose & Stern, 2003). However, these models focused mostly on motor skill-related learning with less emphasis on more complex skills that could involve learning abstract cognitive dependencies implicitly. The second goal of our study was to relate our behavioral results to the various computational models of explicit and implicit learning.

Two main approaches to implicit learning emerged in developmental neuroscience with a different assessment of how learning abilities change with age: (1) the developmental invariance model and (2) the age-related changes model. Studies supporting the developmental invariance model of implicit learning failed to find significant age-related differences in learning (Meulemans, Van der Linden & Perruchet, 1998; Vinter & Perruchet, 2000). In support of this view, infant studies have shown that adult-like implicit learning mechanisms exist even in very early infancy (Clohessy, Posner & Rothbart, 2001; Saffran, Aslin & Newport, 1996). Developmental invariance models explain this age-independence by linking implicit (or procedural) learning to evolutionarily primitive brain regions, such as the basal ganglia and the cerebellum. These regions are characterized as early-maturation regions and are relatively resistant to neurological impairments (Reber, 1993).

By contrast, the age-related changes models posit that considerable developmental differences can be observed in implicit learning. Several of these studies found that older children and young adults showed stronger learning effects compared to very young participants (Fletcher, Maybery & Bennett, 2000; Kirkham, Slemmer, Richardson & Johnson, 2007; Maybery, Taylor & O'Brien-Malone, 1995; Thomas, Hunt, Vizueta, Sommer, Durston, Yang & Worden, 2004). These models accept the fronto-striatal origin of such learning, but they focus on evidence of continued development of these regions that form the basis of the behavioral changes with age (e.g. Thomas *et al.*, 2004). We compared our empirical results using a new approach to the problem of multiple neural substrates of learning proposed by Daw, Niv and Dayan (2005).

Serial reaction time task and the development of implicit learning

In our study, we used a modified version of the Serial Reaction Time (SRT) Task, which is one of most commonly used methods for measuring implicit skill learning. Serial Reaction Time Task is a four-choice reaction time task containing a hidden repeating sequence that the subject comes to predict and learn implicitly (Nissen & Bullemer, 1987; Poldrack *et al.*, 2005). In an SRT study, Meulemans *et al.* (1998) found that 6- and 10-year-old children showed similar degrees of learning as young adults. In contrast, Thomas *et al.* (2004) found that the learning performance of young adults was better than 7- to 10-year-old children. Studies investigating implicit skill learning at older ages also revealed inconsistent results. For example, several studies have demonstrated

that, for simple repeating patterns (in the SRT task), the extent of implicit sequence learning in elderly adults was comparable to young adults (Frensch & Miner, 1994; D.V. Howard & Howard, 1989, 1992). Moreover, in a recent study, Gaillard, Destrebecqz, Michiels and Cleeremans (2009) found that young (22-year-old), middle-aged (45-year-old), and elderly (71-year-old) participants performed at the same level.

The studies mentioned above used fixed (deterministic) sequences, which can be easily learned, making it less possible to detect age-related differences in learning. Furthermore, they cannot purely determine the acquired sequence-specific knowledge because these tasks (finger-tapping, classical SRT) confound general improvements with sequence-specific learning. Here, we used a modified version of the SRT task, the Alternating Serial Reaction Time (ASRT) task (J.H. Howard & Howard, 1997), which enabled us to measure the 'pure' sequence-specific learning distinguished from general improvements. In the classical SRT task, the structure of a sequence is deterministic with the stimuli following a simple cyclically repeating pattern (e.g. 2134121341213412..., where numbers refer to distinct events within the repeating 21341 pattern). By contrast, in the ASRT task (J.H. Howard & Howard, 1997; Remillard, 2008), repeating events alternate with random ones. Thus, the location of every second stimulus on the screen was determined randomly. If, for instance, the sequence was 12341234..., where the numbers represent locations on the screen, the sequence of stimuli would be 1R2R3R4R1R2R3R4R... in the ASRT paradigm, with R representing a random element. Therefore, the location of every second stimulus on the screen was determined randomly. Because fixed, sequence-specific and random stimuli were alternating, some sequences of three events (called 'triplets') occurred more frequently than others. For example, in the above illustration 1x2, 2x3, 3x4 and 4x1 would occur often, whereas 1x3 or 4x2 would occur infrequently. Following previous studies, we referred to the former as *high-frequency* triplets and the latter as *low-frequency* triplets (Nemeth, Janacsek, Londe, Ullman, Howard & Howard, 2010b; Song, Howard & Howard, 2007). Previous studies have shown that as people practice the ASRT task, they respond more quickly to the high- than low-frequency triplets, revealing probabilistic, sequence-specific learning (J.H. Howard & Howard, 1997; Song *et al.*, 2007). This learning is statistical in nature because it depends on the frequency of the event sequences. Thus, the RT difference between the high- and low-frequency triplets in this ASRT task is a measure of human sensitivity to the relative raw probabilities of events observed implicitly in their environment (Perruchet & Pacton, 2006). In addition, the participants are not generally aware of the alternating structure of the sequences, even after extended practice, or when sensitive recognition tests are used to assess explicit knowledge (D.V. Howard, Howard, Japikse, DiYanni, Thompson & Somberg, 2004; J.H.

Howard & Howard, 1997; Song *et al.*, 2007). Thus, the ASRT task is more implicit than the classical deterministic sequence learning tasks.

Using the ASRT task, recent studies have shown that, although elderly adults can also learn the higher-order structure of these complex sequences, they showed age-related deficits (D.V. Howard *et al.*, 2004; J.H. Howard & Howard, 1997; Nemeth *et al.*, 2010b). Both young and elderly adults were able to learn third-order dependencies (1RR2RR3RR1RR2RR3RR...) although the elderly participants performed at a lower level than the younger participants (Bennett, Howard & Howard, 2007). Whereas several studies investigated implicit learning in children using the ASRT task (Barnes, Howard, Howard, Gilotty, Kenworthy, Gaillard & Vaidya, 2008; Barnes, Howard, Howard, Kenealy & Vaidya, 2010; Nemeth, Janacek, Balogh, Londe, Mingesz, Fazekas, Jámbori, Danyi & Vetró, 2010a), no child–adult comparison of implicit skill learning performance has yet been reported.

In summary, previous studies have addressed the development and aging in implicit skill learning, but no studies have examined age-related differences from childhood to old age with identical methods. Furthermore, in contrast to general skill improvements, using a probabilistic sequence learning task (ASRT) can help us to reveal the age-related differences of the underlying mechanisms of complex skill learning by measuring explicitly the sensitivity to raw probabilities of high- and low-frequency events. Therefore, in this study, we compared the implicit probabilistic sequence learning across the age range of 4–85 years.

Method

Participants

There were 421 participants in the experiment between the ages of 4 and 85 that were clustered into nine age groups between 4–6, 7–8, 9–10, 11–12, 14–17, 18–29, 30–44, 45–59 and 60–85 years of age (Table 1). None of them suffered from any developmental, psychiatric or neurological disorders. All subjects gave signed informed consent (parental consent was obtained for children), and they received no financial compensation for partic-

ipation. All experimental procedures were approved by the Ethics Committee of the University of Szeged.

Implicit probabilistic sequence learning task

We used the ASRT task (J.H. Howard & Howard, 1997; Nemeth *et al.*, 2010b) where a stimulus (a dog's head) appeared in one of the four empty circles arranged in a line on a computer screen. The participants were instructed to respond to different stimulus events by pressing the corresponding response keys (Z, C, B or M) as fast and accurately as possible. The ASRT task consisted of 20 blocks with 85 key presses in each block. The first five responses of each stimulus block served for practice only, and then the eight-element alternating sequence (e.g. 1R2R3R4R) was repeated 10 times within a block. Stimuli were presented 120 ms after the response to the previous stimulus. Between blocks, the participants received feedback on the screen about their overall reaction time (RT) and accuracy. The computer program generated a different repeating ASRT sequence of the four locations for each participant using a permutation rule such that each of the six unique permutations of the four repeating events occurred with equal probability.

To determine the amount of explicit knowledge the subjects acquired about the task, a short questionnaire was administered after the experimental session (Song *et al.*, 2007). This questionnaire included increasingly specific questions, such as 'Have you noticed anything special regarding the task?', 'Have you noticed some regularity in the sequence of stimuli?' The experimenter rated subjects' answers on a 5-point scale where 1 denoted 'Nothing noticed' and 5 denoted 'Total awareness'. None of the participants, young or old, reported noticing the hidden repeating sequence.

Statistical properties of the ASRT task

As mentioned above, the ASRT allows a comparison between responses to high- and low-probability events. For example, if the eight-element sequence is 1R2R3R4R, 1x2, 2x3, 3x4, and 4x1 would occur often (high-frequency triplets) because two consecutive stimuli of the repeating sequence (e.g. 132 consisting of 1R2) as well

Table 1 Demographic data and mean RT and accuracy in the different groups. In all columns, numbers in parentheses show standard deviation

Group	Age	Sex	Education	Mean RT (ms)	Mean accuracy (%)
4–6-year-old (<i>n</i> = 30)	5.31 (0.98)	17 M / 13 F	–	960.06 (214.67)	90.09 (6.34)
7–8-year-old (<i>n</i> = 55)	7.09 (0.56)	31 M / 24 F	1.18 (0.39)	773.24 (159.29)	90.66 (7.03)
9–10-year-old (<i>n</i> = 35)	9.89 (0.58)	14 M / 21 F	3.2 (0.96)	602.84 (121.03)	93.44 (4.21)
11–12-year-old (<i>n</i> = 29)	11.5 (0.5)	21 M / 8 F	4.66 (0.67)	544.15 (95.00)	92.52 (4.23)
14–17-year-old (<i>n</i> = 62)	14.89 (1.06)	46 M / 15 F	8.23 (1.02)	452.52 (67.06)	95.44 (2.92)
18–29-year-old (<i>n</i> = 63)	23.09 (3.67)	40 M / 23 F	15.45 (2.6)	401.79 (50.85)	95.47 (2.45)
30–44-year-old (<i>n</i> = 59)	35 (4.24)	24 M / 35 F	16.64 (3.1)	419.85 (58.68)	95.85 (2.98)
45–59-year-old (<i>n</i> = 36)	50.8 (5.07)	12 M / 24 F	14.18 (3.58)	526.7 (112.99)	97.4 (3.45)
60–85-year-old (<i>n</i> = 52)	69.85 (6.16)	16 M / 36 F	13.39 (3.04)	634.37 (126.54)	96.92 (2.38)

as two consecutive random elements by chance (e.g. the same 132 consisting of R3R) could form these triplets. By contrast, 1x3 or 4x2 would occur less frequently (low-frequency triplets) because they could never be obtained consisting of two consecutive sequence elements. Of the 64 possible triplets, 16 triplets were high-frequency triplets, occurring 62.5% of the time, whereas the remaining 48 triplets were low-frequency triplets, occurring 37.5% of the time. Thus, each low-frequency triplet occurred in approximately 0.8% of the total number of trials, whereas each high-frequency triplet occurred about five times more often, in approximately 4% of the trials. For each keypress response, we defined whether it was in response to a high- or a low-frequency element, depending on whether the element was more or less predictable based on the previous two items in the sequence.

Following the method of previous studies (D.V. Howard *et al.*, 2004; Song *et al.*, 2007), two kinds of low-frequency triplets were excluded from our analyses: repetitions (e.g. 222, 333) and trills (e.g. 212, 343). Repetitions and trills were low frequency for all participants, and in previous studies, the participants often showed pre-existing response tendencies towards them (D.V. Howard *et al.*, 2004; Soetens, Melis & Notebaert, 2004). The elimination of these special triplets from the analyses ensured that the high- versus low-frequency differences found in the study were not confounded by pre-existing response tendencies. After this adjustment, previous studies have found that, following the practice, participants responded more quickly to the high- than to the low-frequency triplets, revealing a sequence learning effect (D.V. Howard *et al.*, 2004; J.H. Howard & Howard, 1997; Nemeth *et al.*, 2010b; Song *et al.*, 2007).

Statistical analysis

We calculated the mean accuracy of all trials and the median reaction time (RT) of correct responses separately for high- and low-frequency triplets. The accuracy and RT measures were analyzed by mixed-model ANOVA with TRIPLET TYPE (high vs. low frequency) as the within-subject factor, and AGE (9 groups) as the between-subject factor. All significant results are reported together with the η_p^2 effect size and Greenhouse-Geisser ϵ correction factors where applicable. Post-hoc analysis was conducted by Fisher's LSD pairwise comparisons.

Results

Overall RTs significantly differed among the age groups (main effect of AGE: $F(8, 412) = 107.11$, $p < .001$, $\eta_p^2 = 0.675$; Table 1). The RT decreased significantly between each group from 4–6 to 18–29 years of age (all $ps < .04$), they were similar between the age groups of 18–29 and 30–44 ($p > .38$) and significantly increased after 44 years of age ($ps < .001$) (Figure 1a). The accuracy monotonically increased over the years (main effect of AGE: $F(8, 412) = 16.94$, $p < .001$, $\eta_p^2 = 0.25$) (Figure 1b).

The comparison of RT to high- vs. low-probability triplets showed a surprising pattern of implicit sequence learning across the age groups. Even though there was significant learning at all ages because the RTs were faster for high-frequency than to low-frequency triplets (main effect of TRIPLET TYPE: $F(1, 412) = 333.7$, $p < .001$, $\eta_p^2 = 0.45$, all $ps < .03$) (Figure 1c), the magnitude of this difference was not uniform. Although the age groups differed significantly from each other in

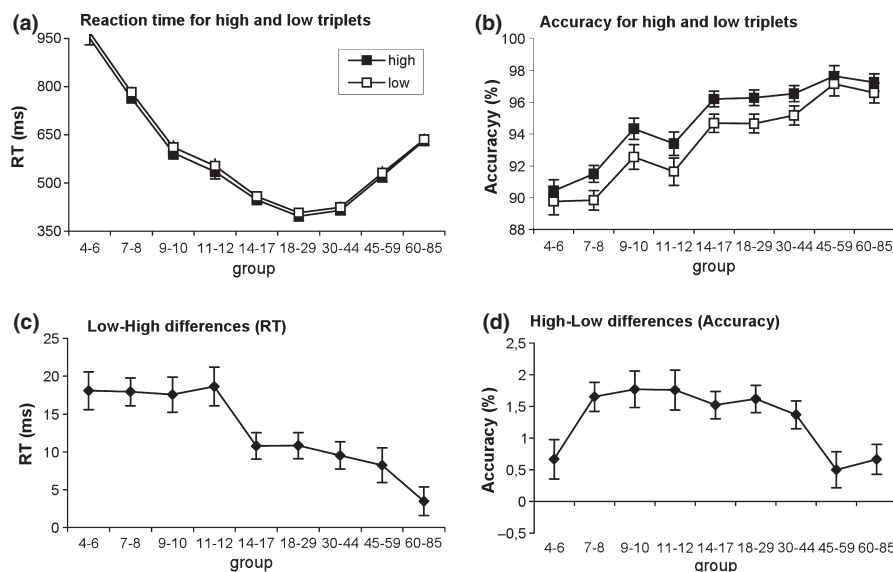


Figure 1 Sequence learning in all groups. Reaction time (a) and accuracy (b) for high- and low-frequency triplets are plotted. Learning measures of RT (c) and accuracy (d) represents the RT/ACC difference between low- and high-frequency triplets. Error bars indicate SEM.

sequence learning (TRIPLET TYPE \times AGE interaction: $F(8, 412) = 6.79, p < .001, \eta_p^2 = 0.12$), the post-hoc test revealed that learning was significantly higher in the 4- to 12-year-old groups than in any other group in the 14–85 range ($ps < .02$). There was no difference in learning between the 14–59 years of age ($ps > .37$), whereas the magnitude of learning decreased significantly in the 60–85-year-old group ($ps < .02$). Thus, learning high-probability events was uniformly effective until the age of 12 where it reduced significantly and remained at a lower level of sensitivity until the age of 60 (Figure 1c).

However, it is a long-standing issue in developmental and aging studies how to compare groups with different baseline speeds. A customary approach to this problem is to analyze the data using z -transformation. Therefore, we calculated the z -scores within each subject (thus, each participant's own mean and SD were used to transform that participant's data (see, for example, Christ, White, Mandernach & Keys, 2001) and conducted an ANOVA based on these z -scores (Figure 2). ANOVA revealed significant sequence-specific learning (main effect of TRIPLET TYPE: $F(1, 412) = 320.12, p < .001, \eta_p^2 = 0.44$), but the extent of learning differed across groups (TRIPLET TYPE \times AGE interaction: $F(8, 412) = 8.91, p < .001, \eta_p^2 = 0.15$). We found that the participants from 9 years of age showed a similar extent of sequence

learning as the adult groups to 44 years of age (all $ps > .25$), but the learning in 4–8 years of age was smaller compared to these adult groups (all $ps < .014$). At older ages, there was a decline in the sequence learning, with both the 45–59 and 60–85-year-old group differing significantly from the groups between 11 and 44 years of age (all $ps < .025$).

The analysis of the response accuracy further enhanced the picture emerging from the results with RTs. We found a significantly greater accuracy for high- than low-frequency triplets (main effect of TRIPLET TYPE: $F(1, 412) = 217.14, p < .001, \eta_p^2 = 0.345$). Although all age groups older than 6 showed significant sequence learning (all $ps < .011$), the age groups differed significantly in the strength of the sequence learning (TRIPLET TYPE \times AGE interaction: $F(8, 412) = 3.73, p < .001, \eta_p^2 = 0.07$). Whereas groups between 7 and 44 years of age showed similar degrees of learning, this was significantly higher than the youngest (4–6) and the two oldest (45–59 and 60–85) groups ($ps < .03$) (Figure 1d).

Figure 3 shows the individual data for sequence learning measured by raw RT, accuracy and z -transformed RT data. The pattern of these data is consistent with the ANOVA results: (1) children between 4 and 12 years of age showed greater sequence learning as measured by raw RT, whereas (2) in sequence-specific learning as measured by accuracy and z -transformed RTs, adults exhibited the highest performance.

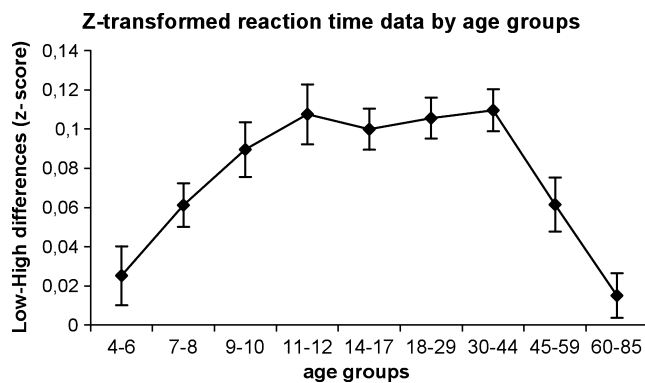


Figure 2 Sequence learning measured the z -transformed RT data in all groups.

Discussion

The goal of the present study was to investigate the differences in implicit skill learning across the human lifespan. This work extends previous studies (Bennett *et al.*, 2007; Gaillard *et al.*, 2009; J.H. Howard & Howard, 1997; Meulemans *et al.*, 1998; Thomas *et al.*, 2004) in two ways: (1) it examined a wide range of ages between 4 and 85 years, and (2) it used a probabilistic task, which enabled us to measure the ‘pure’ sequence-specific learning defined by the sensitivity to raw probabilities of high- and low-frequency events. We found that the 4- to

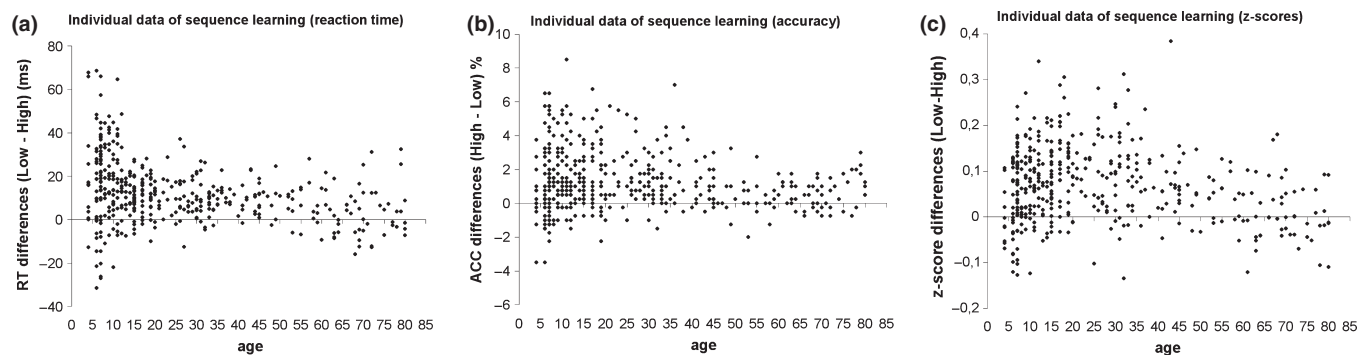


Figure 3 Individual data for sequence learning measured by raw RT (a), accuracy (b), and z -scores (c) in all ages.

12-year-old age groups showed the strongest learning effect measured by the raw RT difference scores. Around the age of 12, we found a striking transition to less pronounced sequence-specific learning, as measured by smaller differences between the responses to high- and low-frequency triplets. Confirming earlier results (D.V. Howard *et al.*, 2004; J.H. Howard & Howard, 1997; Nemeth *et al.*, 2010b), we found that this learning capacity was significantly reduced in the oldest age group. Thus, in contrast to the developmental invariance (Reber, 1993) and the age-related changes approaches (Meulemans *et al.*, 1998; Vinter & Perruchet, 2000), our results demonstrate a gradual decline in learning across the lifespan.

Sequence learning scores based on the accuracy and raw reaction time showed different curves: the former is a bell-shaped curve, whereas the latter is a gradually declining curve (Figure 1c–d). Hence, these two types of learning scores can reflect different underlying mechanisms and brain systems. The accuracy learning score may be more related to attention, mainly voluntary attention, whereas the RT learning score may be related to involuntary attention and intuitive processes (Burgess, Gilbert & Dumontheil, 2007; Prinzmetal, McCool & Park, 2005). The relatively weaker accuracy learning effects in children and older groups may be due to the underdeveloped/deteriorating attentional brain circuits connected to the frontal lobe.

Our study raised a methodological issue that affects developmental studies in general. It is a long-standing issue in the literature how to compare groups with different baseline speeds. It could be argued that the youngest groups in our study have larger raw RT learning scores (greater difference between high- and low-frequency triplets) because they have more room to do so, given their longer RTs. However this argument does not seem to hold in our study, because the oldest group showed equally long RTs as 9–10-year-olds, yet they had a fourfold reduced learning score (Figure 1). Along a similar line of argument, the accuracy measures of the youngest group (4–6-year-olds) and the oldest groups (45–85-year-olds) showed a marked difference (~8%) in general accuracy, giving more room for the young group to produce larger differences between high- and low-frequency triplets, yet the actual sequence learning (i.e. high-low differences in accuracy) showed no difference. In contrast, the two smallest age groups had almost identical general accuracy yet there was a more than twofold increase in sequence learning in the second youngest group (7–8-year-olds). Thus, our data suggest no linear relationship between the general magnitude of reaction time/accuracy and the learning measures.

A second customary approach to the problem of comparing groups with different baseline speeds is to analyze the data using z -transformation. Z -scores of our results show a different picture from the raw RT data analysis: the learning performances from 9 to 44 years of age are similar with weaker performance in the younger

and older age groups. The z -transformation is often used to control general processing speed across different age groups in developmental studies. However, the main function of z -transformation is not to control the processing speed, but to normalize the distribution of responses. Thus, z -transformation has mathematical assumptions about the form of distributions and is therefore not theory independent (Yap, Balota, Sibley & Ratcliff, 2012). Z -scoring fully adjusts for processing speed only if all participants have the same type of distributions. Therefore using z -scores in developmental studies might be misleading. Furthermore, it is unclear how general processing speed and variability contribute to learning and performance in different ages. In recent years, several studies analyzed the variability and noise across a wide age range (Der & Deary, 2006; Hulstsch, MacDonald, Hunter, Levy-Bencheton & Strauss, 2000; Rabbitt, Osman, Moore & Stollery, 2001). For example, Rabbitt *et al.* (2001) found that people's fastest RTs were relatively unaffected by age, but the number of unnecessarily slow responses was higher in older ages, and, thus, the increase in the mean RT was a result of increasing variability, which was an important component of cognitive aging. Moreover, several studies outline that the noise and the performance variability enables adaptive plasticity of motor skills (Slifkin & Newell, 1998, 1999; Turner & Brainard, 2007) and high variability can support effective learning and performance (Sanger, 2010). Thus, based on previous studies and on our analyses, we suggest that both the processing speed (mean reaction time) and variability are inherent aspects of development and aging. We think that the z -transformation eliminates these inherent aspects of learning, making the transformed results difficult to understand and explain.

Based on the reasoning above, we propose that the raw RT difference between the high- and low-frequency triplets in the ASRT task is a measure of human sensitivity to the relative raw probabilities of events observed implicitly in their environment. Thus, our results show a marked decrease in this sensitivity around the age of 12, which is in contrast to the traditional view of a steady improvement of cognitive learning abilities until late in adulthood (Craik & Bialystok, 2006). However, this discrepancy might be explained based on a shift in the structural development of implicit learning. Although the raw probabilities of the sensory environment are important for learning and both infants (Aslin, Saffran & Newport, 1998; Fiser & Aslin, 2002; Saffran *et al.*, 1996; Saffran, Johnson, Aslin & Newport, 1999) and adults (Fiser & Aslin, 2001; Hunt & Aslin, 2001) are highly sensitive to them, there is an ongoing debate on how using these simple probabilities can lead to highly complex knowledge of the world, such as sensory invariances and development of a language (Gomez & Gerken, 1999; Marcus, Vijayan, Rao & Vishton, 1999). Recent studies have proposed that using an internally stored structured model of the world that emerges based

on past experience together with probabilistic learning could help to address this issue and also provide evidence that humans might implement such a strategy during implicit learning (Orban, Fiser, Aslin & Lengyel, 2008; Tenenbaum, Kemp, Griffiths & Goodman, 2011). In this framework, as the internal model develops with experiences becoming more influential, internal interpretations of events become more elaborate and less directly related to their raw probabilities. A recent study argued that from a normative standpoint, existence of multiple learning mechanisms in the brain (cf. model-free vs. model-based learning) with an uncertainty-based arbitration between them would be computationally optimal (Daw *et al.*, 2005). Anchoring this hypothesis biologically, the presumed mechanisms related to these two types of learning were suggested to be related to the prefrontal areas and temporal lobe of the cortex, respectively (Daw *et al.*, 2005). Support for the separated

and complementary nature of the prefrontal- and medial temporal lobe (MTL)-dependent learning based on internal models vs. basal ganglia-dependent model-free learning comes from various studies investigating learning under specific conditions. These studies showed that obstructing the PFC and/or MTL by a demanding secondary task (Foerde, Knowlton & Poldrack, 2006) does not adversely affect implicit learning. Other studies found that inserting a task between the learning sessions (Brown & Robertson, 2007a, 2007b), performing a working memory and an implicit learning task simultaneously (Filoteo, Lauritzen & Maddox, 2010), or a neuropharmacological blockage (Frank, O'Reilly & Curran, 2006) even had a positive effect on performance in an implicit learning task. Importantly, it is known that the cortical areas connected to the internal models related to model-based learning become truly functional late in development, around the age of 12 (Blakemore &

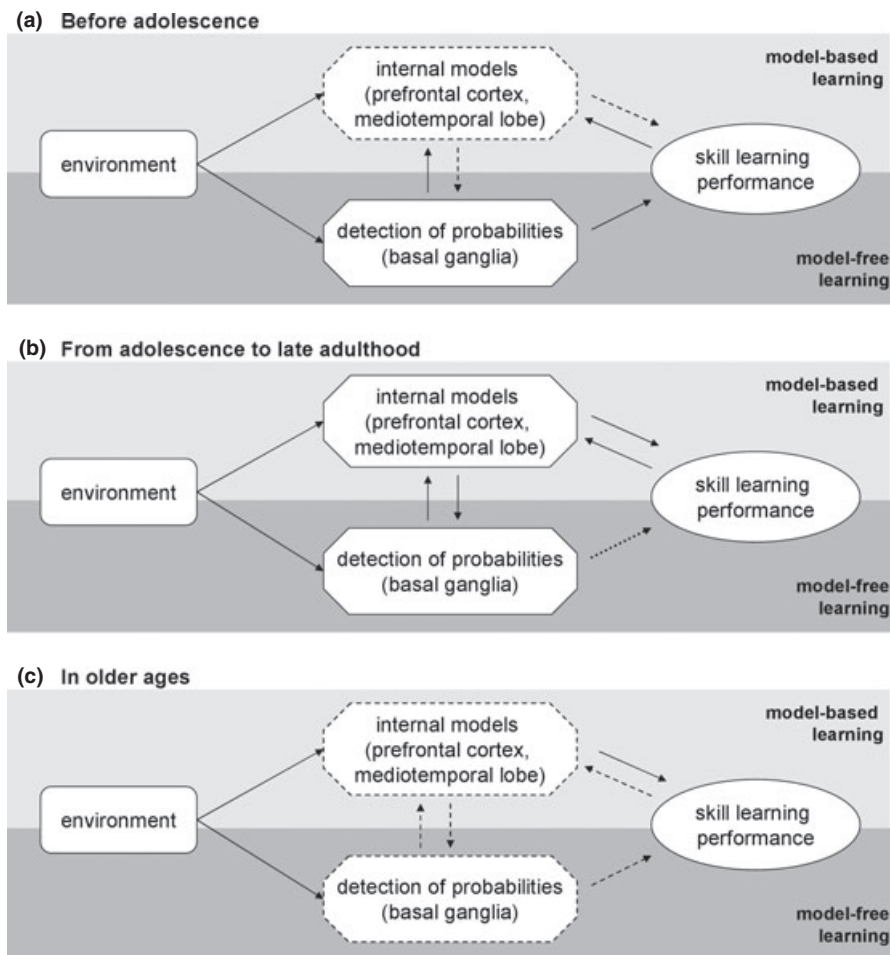


Figure 4 Competition between model-based and model-free neurocognitive subsystems of skill learning across lifespan. (a) Before adolescence, underdeveloped internal models (dashed boundary) have little influence on interpretations of detected raw statistical probabilities of events in the environment (dashed arrows). Skill learning performance is determined by raw probabilities. (b) From adolescence to late adulthood, well-developed internal models (solid boundary) strongly modulate the interpretations of observed statistics of the input. This helps in extracting complex relations but relatively impairs measuring and utilizing raw probabilities in skill learning (dotted arrow). (c) In older ages, skill learning performance decreases. This decline could be caused by the combination of reduced sensitivity to raw statistical probabilities (dashed boundary), increasingly rigid internal models (dashed boundary) and/or weaker connection between these systems (dashed arrows).

Choudhury, 2006; Giedd, Blumenthal, Jeffries, Castellanos, Liu, Zijdenbos, Paus, Evans & Rapoport, 1999), which is about the age when we found the sudden decrement in sensitivity to the relative raw probabilities. We propose that this enhanced functionality signals the shift when the system adapts efficiently to more complex aspects of the world by relying more on internal model-based interpretations, while somewhat neglecting the raw probabilities of the sensory input (Figure 4a–b), and therefore, decreasing the ability to develop and stabilize fundamentally new basic competences. Thus the seemingly paradoxical result of gradually becoming less sensitive to basic statistics, if timed appropriately, could be the optimal strategy for human implicit learning in general.

Our results did not reveal any differences between the young adults and middle-aged groups. Salthouse's (1996) 'simultaneity mechanism' theory of cognitive aging predicts the age-related deficits in probabilistic sequence learning (Curran, 1997; J.H. Howard & Howard, 1997). Feeney, Howard and Howard (2002) found age-related deficits in pattern sensitivity in 'older' (mean age: 49.4) compared with 'younger' (mean age: 41.4) middle-aged groups. These different results could be related to the fact that Feeney used a smaller sample size, a longer version of the ASRT and a different method of analysis.

What are the underlying mechanisms of the decreased performance of the elderly group? In a recent fMRI study, Dennis and Cabeza (2011) showed that older adults recruited the MTL for implicit learning, and this activation was significantly greater, while striatal activity decreased in older people compared with young adults during implicit learning. In a recent study, Rieckmann, Fischer and Bäckman (2010) found similar results: in young adults during the learning session, the activation of the striatum increased, but that of the MTL decreased. By contrast, in older adults, sequence learning positively related to activation increases in both the striatum and MTL. Using multimodal imaging measures, Giorgio, Santelli, Tomassini, Bosnell, Smith, De Stefano and Johansen-Berg (2010) found extensive reductions in the gray matter volume in aging, but reductions were detected earlier in the frontal cortex. Furthermore, a recent diffusion tensor imaging aging study by Bennett, Madden, Vaidya, Howard and Howard (2011) found that the caudate-dorsolateral prefrontal cortex (DLPFC) and hippocampus-DLPFC tract integrity were related to ASRT sequence learning. The caudate-DLPFC tract integrity decreased in the older ages, mediating age-related differences in sequence learning. Within the computational framework proposed by Daw and colleagues (2005), these findings can be interpreted as a deterioration in three mechanisms that contribute to the age-related decline in skill learning: (1) reduced detection of probabilities, (2) rigidity of internal models and/or (3) more restricted connections between internal models and probability detection (Figure 4c). Thus, not only

model-free but also model-based learning might be limited in older ages. Future studies are needed to systematically examine the underlying neural mechanisms of age-related differences in skill learning.

In summary, based on our raw RT results we suggest that acquiring fundamentally new skills that cannot be derived from skills already possessed is most effective before adolescence, and it might be largely based on the fronto-striatal circuitry, such as the basal ganglia and cerebellum, in agreement with earlier skill learning models (Doyon *et al.*, 2009; Hikosaka *et al.*, 1999; Hikosaka *et al.*, 2002) and computational learning models (Daw *et al.*, 2005). Our findings are in good agreement with everyday life experience showing that an early (~ before 12 years) start in learning some sports, musical instruments, second language, etc. often leads to higher levels of competence. These results may have implications for the development of learning and memory, facilitating new skill training and pedagogic methods (e.g. for teaching languages) and may also contribute to the understanding of neurodevelopmental and age-related disorders (e.g. autism, SLI, dyslexia and dementia) and lead to relevant treatment options.

Acknowledgements

Thanks to Darlene V. Howard, James H. Howard, Jr. and Michael Ullman for numerous illuminating discussions. Thanks also to Robert Mingesz, Attila Krajcsi, István Winkler, Gábor Orosz, Ágnes Lukács, Ágnes Kovács and Zsuzsa Londe for their comments on the previous version of the manuscript. This research was supported by the Bolyai Scholarship Program (to DN), by OTKA K 82068 (to DN), and by NIH NEI R01 EY018196 (to JF).

References

- Albouy, G., Sterpenich, V., Balteau, E., Vandewalle, G., Desseilles, M., Dang-Vu, T., Darsaud, A., Ruby, P., Luppi, P.H., Degueldre, C., Peigneux, P., Luxen, A., & Maquet, P. (2008). Both the hippocampus and striatum are involved in consolidation of motor sequence memory. *Neuron*, **58** (2), 261–272.
- Aslin, R.N., Saffran, J.R., & Newport, E.L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, **9** (4), 321–324.
- Barnes, K.A., Howard, J.H. Jr, Howard, D.V., Gilotty, L., Kenworthy, L., Gaillard, W.D., & Vaidya, C.J. (2008). Intact implicit learning of spatial context and temporal sequences in childhood autism spectrum disorder. *Neuropsychology*, **22** (5), 563–570.
- Barnes, K.A., Howard, J.H. Jr, Howard, D.V., Kenealy, L., & Vaidya, C.J. (2010). Two forms of implicit learning in childhood ADHD. *Developmental Neuropsychology*, **35** (5), 494–505.
- Bennett, I.J., Howard, J.H. Jr, & Howard, D.V. (2007). Age-related differences in implicit learning of subtle third-order sequential structure. *Journal of Gerontology: Psychological Sciences*, **62B** (2), 98–103.

- Bennett, I.J., Madden, D., Vaidya, C., Howard, J., & Howard, D. (2011). White matter integrity correlates of implicit sequence learning in healthy aging. *Neurobiology of Aging*, **32** (12), 2317. e1–2317.e12. doi: 10.1016/j.neurobiolaging.2010.03.017
- Blakemore, S.J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, **47** (3-4), 296–312.
- Brown, R.M., & Robertson, E.M. (2007a). Inducing motor skill improvements with a declarative task. *Nature Neuroscience*, **10** (2), 148–149.
- Brown, R.M., & Robertson, E.M. (2007b). Off-line processing: reciprocal interactions between declarative and procedural memories. *Journal of Neuroscience*, **27** (39), 10468–10475.
- Burgess, P.W., Gilbert, S.J., & Dumontheil, I. (2007). Function and localization within rostral prefrontal cortex (area 10). *Philosophical Transactions of the Royal Society B: Biological Sciences*, **362** (1481), 887–899.
- Christ, S.E., White, D.A., Mandernach, T., & Keys, B.A. (2001). Inhibitory control across the life span. *Developmental Neuropsychology*, **20** (3), 653–669.
- Clohessy, A.B., Posner, M.I., & Rothbart, M.K. (2001). Development of the functional visual field. *Acta Psychologica*, **106** (1), 51–68.
- Cohen, D., Pascual-Leone, A., Press, D., & Robertson, E. (2005). Off-line learning of motor skill memory: a double dissociation of goal and movement. *Proceedings of the National Academy of Sciences, USA*, **102** (50), 18237–18241.
- Craik, F.I., & Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends in Cognitive Sciences*, **10** (3), 131–138.
- Curran, T. (1997). Effects of aging on implicit sequence learning: accounting for sequence structure and explicit knowledge. *Psychological Research*, **60** (1-2), 24–41.
- Daw, N.D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, **8** (12), 1704–1711.
- Dennis, N., & Cabeza, R. (2011). Age-related dedifferentiation of learning systems: an fMRI study of implicit and explicit learning. *Neurobiology of Aging*, **32** (12), 2318.e17–2318.e30. doi: 10.1016/j.neurobiolaging.2010.04.004
- Der, G., & Deary, I.J. (2006). Age and sex differences in reaction time in adulthood: results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging*, **21** (1), 62–73.
- Doyon, J., Bellec, P., Amsel, R., Penhune, V., Monchi, O., Carrier, J., Lehericy, S., & Benali, H. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behavioral Brain Research*, **199** (1), 61–75.
- Feeney, J.J., Howard, J.H., & Howard, D.V. (2002). Implicit learning of higher order sequences in middle age. *Psychology and Aging*, **17** (2), 351–355.
- Filoteo, J.V., Lauritzen, S., & Maddox, W.T. (2010). Removing the frontal lobes. *Psychological Science*, **21** (3), 415–423.
- Fiser, J., & Aslin, R.N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, **12** (6), 499–504.
- Fiser, J., & Aslin, R.N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, USA*, **99** (24), 15822–15826. doi: 10.1073/pnas.232472899
- Fletcher, J., Maybery, M.T., & Bennett, S. (2000). Implicit learning differences: a question of developmental level? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **26** (1), 246–252.
- Foerde, K., Knowlton, B.J., & Poldrack, R.A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences, USA*, **103** (31), 11778–11783.
- Frank, M.J., O'Reilly, R.C., & Curran, T. (2006). When memory fails, intuition reigns: midazolam enhances implicit inference in humans. *Psychological Science*, **17** (8), 700–707.
- Frensch, P.A., & Miner, C.S. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory and Cognition*, **22** (1), 95–110.
- Gaillard, V., Destrebecqz, A., Michiels, S., & Cleeremans, A. (2009). Effects of age and practice in sequence learning: a graded account of ageing, learning, and control. *European Journal of Cognitive Psychology*, **21** (2), 255–282.
- Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., & Rapoport, J.L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, **2** (10), 861–863.
- Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S., De Stefano, N., & Johansen-Berg, H. (2010). Age-related changes in grey and white matter structure throughout adulthood. *NeuroImage*, **51** (3), 943–951.
- Gomez, R.L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, **70** (2), 109–135.
- Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience*, **11**, 523–532.
- Hikosaka, O., Nakahara, H., Rand, M.K., Sakai, K., Lu, X., Nakamura, K., Miyachi, S., & Doya, K. (1999). Parallel neural networks for learning sequential procedures. *Trends in Neurosciences*, **22** (10), 464–471.
- Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current Opinion in Neurobiology*, **12** (2), 217–222.
- Howard, D.V., & Howard, J.H. Jr (1989). Age differences in learning serial patterns: direct versus indirect measures. *Psychology and Aging*, **4** (3), 357–364.
- Howard, D.V., & Howard, J.H. (1992). Adult age differences in the rate of learning serial patterns: evidence from direct and indirect tests. *Psychology and Aging*, **7** (2), 232–241.
- Howard, D.V., Howard, J.H. Jr, Japikse, K., DiYanni, C., Thompson, A., & Somberg, R. (2004). Implicit sequence learning: effects of level of structure, adult age, and extended practice. *Psychology and Aging*, **19** (1), 79–92.
- Howard, J.H. Jr, & Howard, D.V. (1997). Age differences in implicit learning of higher-order dependencies in serial patterns. *Psychology and Aging*, **12** (4), 634–656.
- Hultsch, D.F., MacDonald, S.W.S., Hunter, M.A., Levy-Bencheton, J., & Strauss, E. (2000). Intraindividual variability in cognitive performance in older adults: comparison of adults with mild dementia, adults with arthritis, and healthy adults. *Neuropsychology*, **14** (4), 588–598.
- Hunt, R.H., & Aslin, R.N. (2001). Statistical learning in a serial reaction time task: access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General*, **130** (4), 658–680.

- Keele, S.W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review*, **110** (2), 316–339.
- Kirkham, N., Slemmer, J., Richardson, D., & Johnson, S. (2007). Location, location, location: development of spatio-temporal sequence learning in infancy. *Child Development*, **78** (5), 1559–1571.
- Lieberman, M.D. (2000). Intuition: a social cognitive neuroscience approach. *Psychological Bulletin*, **126** (1), 109–137.
- Marcus, G.F., Vijayan, S., Rao, S.B., & Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science*, **283**, 77–80.
- Maybery, M., Taylor, M., & O'Brien-Malone, A. (1995). Implicit learning: sensitive to age but not IQ. *Australian Journal of Psychology*, **47** (1), 8–17.
- Meulemans, T., Van der Linden, M., & Perruchet, P. (1998). Implicit sequence learning in children. *Journal of Experimental Child Psychology*, **69** (3), 199–221.
- Nemeth, D., Janacek, K., Balogh, V., Londe, Z., Mingesz, R., Fazekas, M., Jámbori, S., Danyi, I., & Vetró, A. (2010a). Learning in autism: implicitly superb. *PLoS ONE*, **5** (7), e11731.
- Nemeth, D., Janacek, K., Londe, Z., Ullman, M.T., Howard, D., & Howard, J. (2010b). Sleep has no critical role in implicit motor sequence learning in young and old adults. *Experimental Brain Research*, **201** (2), 351–358. doi: 10.1007/s00221-009-2024-x
- Nissen, M.J., & Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cognitive Psychology*, **19**, 1–32.
- Orban, G., Fiser, J., Aslin, R.N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences, USA*, **105** (7), 2745–2750.
- Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, **10** (5), 233–238.
- Poldrack, R.A., Sabb, F.W., Foerde, K., Tom, S.M., Asarnow, R.F., Bookheimer, S.Y., & Knowlton, B.J. (2005). The neural correlates of motor skill automaticity. *Journal of Neuroscience*, **25**, 5356–5364.
- Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General*, **134** (1), 73–92.
- Rabbitt, P., Osman, P., Moore, B., & Stollery, B. (2001). There are stable individual differences in performance variability, both from moment to moment and from day to day. *Quarterly Journal of Experimental Psychology Section A*, **54** (4), 981–1003.
- Reber, A.R. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious (Vol. 19)*. New York: Oxford University Press.
- Remillard, G. (2008). Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies. *Quarterly Journal of Experimental Psychology*, **61** (3), 400–424.
- Rieckmann, A., Fischer, H., & Bäckman, L. (2010). Activation in striatum and medial temporal lobe during sequence learning in younger and older adults: relations to performance. *NeuroImage*, **50** (3), 1303–1312.
- Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, **274**, 1926–1928.
- Saffran, J.R., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, **70** (1), 27–52. doi: S0010-0277(98)00075-4
- Salthouse, T.A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, **103** (3), 403–428.
- Sanger, T.D. (2010). Controlling variability. *Journal of Motor Behavior*, **42** (6), 401–407.
- Schendan, H., Searl, M., Melrose, R., & Stern, C. (2003). An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron*, **37** (6), 1013–1025.
- Slifkin, A.B., & Newell, K.M. (1998). Is variability in human performance a reflection of system noise? *Current Directions in Psychological Science*, **7** (6), 170–177.
- Slifkin, A.B., & Newell, K.M. (1999). Noise, information transmission, and force variability. *Journal of Experimental Psychology: Human Perception and Performance*, **25** (3), 837–851.
- Soetens, E., Melis, A., & Notebaert, W. (2004). Sequence learning and sequential effects. *Psychological Research*, **69** (1), 124–137.
- Song, S., Howard, J.H. Jr., & Howard, D.V. (2007). Sleep does not benefit probabilistic motor sequence learning. *Journal of Neuroscience*, **27** (46), 12475–12483.
- Squire, L.R., & Zola, S.M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences, USA*, **93**, 13515–13522.
- Tenenbaum, J.B., Kemp, C., Griffiths, T.L., & Goodman, N.D. (2011). How to grow a mind: statistics, structure, and abstraction. *Science*, **331** (6022), 1279–1285.
- Thomas, K.M., Hunt, R.H., Vizueta, N., Sommer, T., Durston, S., Yang, Y., & Worden, M.S. (2004). Evidence of developmental differences in implicit sequence learning: an fMRI study of children and adults. *Journal of Cognitive Neuroscience*, **16** (8), 1339–1351.
- Tulving, E., & Craik, F.I.M. (Eds.) (2000). *The Oxford handbook of memory*. Oxford: Oxford University Press.
- Turner, E.C., & Brainard, M.S. (2007). Performance variability enables adaptive plasticity of crystallized adult birdsong. *Nature*, **450** (7173), 1240–1244.
- Ullman, M.T. (2001). A neurocognitive perspective on language: the declarative/procedural model. *Nature Reviews Neuroscience*, **2**, 717–726.
- Vinter, A., & Perruchet, P. (2000). Implicit learning in children is not related to age: evidence from drawing behavior. *Child Development*, **71** (5), 1223–1240.
- Yap, M.J., Balota, D.A., Sibley, D.E., & Ratcliff, R. (2012). Individual differences in visual word recognition: insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance*, **38** (1), 53–79. doi: 10.1037/a0024177

Received: 26 July 2010

Accepted: 13 January 2012