EÖTVÖS LÓRÁND UNIVERSITY

FACULTY OF EDUCATION AND PSYCHOLOGY

Theses of the Doctoral Dissertation

Emese Szegedi-Hallgató

Methodological and theoretical considerations in implicit learning research

Doctoral School of Psychology

Head of the School: Prof. Zsolt Demetrovics (PhD, DSc), professor, ELTE

Cognitive Psychology Program

Head of the Program: Prof. Ildikó Király (PhD), professor, ELTE

Supervisor

Prof. Dezső Németh (PhD, DSc), professor, ELTE

Budapest, 2019

GENERAL INTRODUCTION

Skill acquisition, habit formation, and development of behavioral automatisms are all results of learning processes, sharing a unique combination of features that makes them different from other kinds of learning. According to one point of view, these learning types are forms of non-declarative learning, underscoring thus that learning is not dependent on the mediotemporal brain structures (Squire & Zola, 1996). Another viewpoint emphasizes the fact that learners are usually not fully aware of the information that had been acquired, and it is only their improving performance that implies learning, thus emphasizing conscious awareness (or the lack of it) as a defining criterion. Learning that occurs without awareness is called implicit - in contrast with explicit learning where conscious awareness accompanies learning (Graf & Schacter, 1985). A third approach relies on three variables: the speed of encoding (rapid vs. slow); whether a single item is encoded or associations among multiple items; and the compositionality (vs. rigidity) of the resulting memory (Henke, 2010). According to this view, skill acquisition and habit formation is a form of slow encoding of rigid associations (as is classical conditioning and semantic memory). And finally, there is a separate research tradition, namely the investigation of statistical learning abilities originating from Saffran, Aslin, & Newport (1996) that also deals with the unsupervised, incidental learning of an inherent structure present in the to-be-learned material; Perruchet & Pacton (2006) went as far as suggesting that implicit learning and statistical learning is actually the same phenomenon (see also Christiansen, 2018). In a similar vein, Reber (2013) proposed that implicit memory manifests as an improvement from experience based on mechanisms of cortical plasticity; the extraction of the underlying statistical structure is incremental, and it allows for a distributed representation of information.

Despite the similarities between these research traditions, and the substantial overlap of their proposed constructs, their notions are not synonyms. For example, the term *implicit learning* is broader than the term *skill learning*, as other types of implicit learning phenomena also exist, e.g. priming, classical conditioning and habituation/sensitization (Squire & Zola, 1996). On the other hand, skill learning does not only rely on implicit processes but also on explicit learning (Ghilardi, Moisello, Silvestri, Ghez, & Krakauer, 2009; Taylor, Krakauer, & Ivry, 2014). Third, although statistical learning is thought to be an implicit learning process (e.g. Kim, Seitz, Feenstra, & Shams, 2009; Perruchet & Pacton, 2006; Turk-Browne, Scholl, Chun, & Johnson, 2008), there is evidence that explicit knowledge can also emerge after the encounter with statistically structured stimuli (e.g. Perruchet, Bigand, & Benoit-Gonin, 1997; Rünger & Frensch, 2008; Goujon, Didierjean, & Poulet, 2014). The narrow field I was interested in (which is summarized in this work) is the *implicit statistical learning*, not implicit learning or skill learning in general.

Tasks of implicit statistical learning

Typical tests of implicit (statistical) learning include the Artificial Grammar Learning (AGL) Task (e.g. Danner, Hagemann, & Funke, 2017); the Weather Prediction (WP) Task (or more generally the Probabilistic Classification tasks) (e.g. Knowlton, Squire, & Gluck, 1994); the Sugar Factory task (or more generally the Dynamic Systems Control tasks) (Berry & Broadbent, 1984); the Contextual Cueing

(CC) paradigm (Chun, 2000); and the Serial Reaction Time task (Nissen & Bullemer, 1987) - or more generally the Sequence Learning tasks.

In all of the research projects presented in the Dissertation we used a sequence learning task, namely the ASRT (Alternating Serial Reaction Time) task. During the task visual stimuli are presented on a computer screen in one of four possible locations, and the subject's task is to react as fast and as accurately as possible to the location of the stimuli by pressing the corresponding response button (usually aligned similarly as the stimuli to allow for a simple 1:1 stimulus-response mapping). The stream of stimuli is not entirely random: a pre-defined four-element long pattern (P) is embedded in a stream of random (R) trials so that P and R trials alternate (hence the name of the task). The structure that results from their alternation is a second-order probabilistic sequence. In this particular case, after encountering any two consecutive trials, a prediction could be made about what to expect next. The term *probabilistic* refers to the fact that sometimes the following trial is „unexpected", not very probable. Learning can be derived from comparing performance on probable vs. improbable trials. The authors introducing the task reported that not a single subject became aware of the hidden pattern (J. H. Howard & Howard, 1997), and our experience with the task corroborates their notion. In summary then, the ASRT task is an implicit visuomotor statistical learning task measuring the ability to acquire second-order probabilistic information.

The problem

The aformentioned tasks of implicit statisitcal learning differ in many ways; e.g. whether the regularity is present temporally or spatially, whether the exposure of the regularity is passive or requires some activity from the participant, etc. Nevertheless, they all rely on the detection of statistical regularities that is covertly present in the task (Arciuli & Conway, 2018). It is thus somewhat surprising that learning scores gained from different tests do not correlate with each other (Gebauer & Mackintosh, 2007; Sævland & Norman, 2016; Siegelman & Frost, 2015) or even if they do, the correlation is weak (Kalra, Gabrieli, & Finn, 2019).

The lack of correlation between the different measures of implicit statistical learning is alarming, and it is important to find the reasons behind it. First, it is possible that there is truly no relationship between these measures and hence research is (rightfully) unable to find one. Theoretically, this scenario would question the domain-generality (opposed to domain-specificity) and/or the unitary nature (opposed to multicomponentiality) of implicit statistical learning. In other words, it would mean that there is no such thing as „the implicit statistical learning", only different types of it. Practically, it would highlight the need to find the factors that differentiate between different types of implicit statistical learning, and this knowledge – in turn – be utilized in designing new tasks and/or help us chose from the existing tasks so that we could use the most relevant measures for our purposes.

In a second scenario, there *is* a positive relationship between these different measures, but – for some reason – researchers have been unable to find it. The reason behind this could be methodological and/or related to the psychometric properties of the tasks. In spite of bearing the hope that we could somehow overcome these obstacles in the future, this scenario would also mean that our knowledge about the nature of implicit statistical learning is seriously biased (possibly wrong in many aspects). If our tests are so weak that they barely correlate with each other, how could we interpret the lack of correlation with other kinds of tests?

Possible Causes of Low Correlation

Modality Specificity - Accumulating evidence suggests that there are qualitative differences in patterns of implicit statistical learning in the auditory, visual and tactile modalities, which corroborate the notion of modality specificity of implicit statistical learning (Emberson, Conway, & Christiansen, 2011; Li, Zhao, Shi, Lu, & Conway, 2018; Walk & Conway, 2016). The explanation might be that encoding of information follows constraints that are determined by the specific properties of the input in the respective brain cortices (despite similar sets of computational principles) (Conway & Christiansen, 2005).

Independency from other cognitive abilities - Arciuli (2017) suggested that implicit statistical learning is a multicomponent ability (being comprised of certain types of attention, processing speed, and memory, etc.); and that performance on different tasks might depend on the way they draw on particular underlying components (Arciuli, 2017; Arciuli & Conway, 2018). Additionally, even if implicit and explicit processes dissociate, it does not exclude the possibility of interplay between these memory systems; and although the evidence is not unequivocal, some results do point towards this possibility (Boyd & Winstein, 2003; Arnaud Destrebecqz et al., 2005; Dew & Cabeza, 2011; Lagarde, Li, Thon, Magill, & Erbani, 2002; Sun, Zhang, Slusarz, & Mathews, 2007; but see Sanchez & Reber, 2013; and Curran & Keele, 1993).

Type of statistics – It has been recognized that humans are capable of learning at least two types of statistics: joint probaibilities (i.e. distributional statistics of chunks of information), and conditional probabilities (i.e. the predictability of a target event given its antecedents) (J. H. Howard, Howard, Dennis, & Kelly, 2008; Thiessen, Kronstein, & Hufnagle, 2013; Thiessen, 2017) and it has been suggested that they are the results of independent processes (Thiessen, 2017). However, the relative contribution of different types of statistics in a specific learning task is rarely discussed (but see J. H. Howard et al., 2008). Additionally, the complexity of the embedded statistical structure might also contribute to differences observed with different statistical learning tasks. For example, in sequential tasks, when the previous element predicts the next element, it is called a first-order sequential structure; when the N-2th trial has predictive power on the current target, the sequence has a second-order structure, and so on. It has been shown that humans are capable of learning up to fourth-order statistical regularities (Remillard, 2008, 2011), or even fifth- and sixth-order regularities (Remillard, 2010). At the same time it has been shown that learning of higher-order information can be selectively impaired (in dyslexia: W. Du & Kelly, 2013; J. H. Howard et al., 2006; in Parkinson's disease: Smith & McDowall, 2004; in Schizophrenia: Schwartz, Howard, Howard, Hovaguimian, & Deutsch, 2003; with age: J. H. Howard, Howard, Dennis, & Yankovich, 2007; D. V. Howard et al., 2004; Feeney, Howard, & Howard, 2002; J. H. Howard & Howard, 1997; Urry, Burns, & Baetu, 2018). It is a matter of question, though, whether lower- and higher-order sequence learning should be thought of as worse or better performance on the same measure, or as different abilities.

Low reliability - Other things being equal, the correlation between two variables will be lower when the reliability of the measures are lower (i.e. measurement error is high). Since reliability is the correlation of a test with itself, it is easy to see that a measure that does not correlate with itself can not correlate with other variables either (Goodwin & Leech, 2006).

Low individual variability - It is hypothesized that implicit learning is evolutionarily older than explicit learning, implying that it is also more robust and results in less inter- and intra-species variability (Reber & Allen, 2000). It has been assumed that individual differences in implicit cognition are minimal relative to individual differences in explicit cognition (Reber, 1993). In line with this assumption, the individual differences in implicit cognition remained largely unexplored (Reber & Allen, 2000; but see Kaufman et al., 2010; and Kalra, Gabrieli, & Finn, 2019).

Issues related to reaction-time based measures - Difference scores derived from reaction times are thought to be unstable (Kaufman et al., 2010), and it was suggested that accuracy (Urry, Burns, & Baetu, 2015; Urry et al., 2018) or reaction time ratio measures (Kaufman et al., 2010) provide better measures of learning, and are less prone to result in floor effects (Urry et al., 2015). The fact that difference scores based on reaction times and difference scores based on accuracy do not show correlation (Hedge, Powell, Bompas, Vivian-Griffiths, & Sumner, 2018) also implies that the choice between the two types of measures should not be based on convenience or traditions only, but should be a matter of theoretical consideration.

Second, there is an often-overlooked factor that might influence serial reaction time tasks, namely that different series of responses are not equally easy to be performed. This is sometimes referred to as „pre-existing sequential effects" and „preexisting biases" in the context of serial reaction time tasks (Song, Howard, & Howard, 2007a) or, more generally, „sequential effects" in the context of the broader category of forced-choice reaction time tasks (e.g. Remington, 1969). Complementary to these cognitive effects, there are also biomechanical constraints of the body that also affect serial reaction times (Y. Du & Clark, 2017). Apart from manifesting as an artifact, and thus influencing our interpretations of the results, these biases might also mask the individual variability of implicit learning (given that they are robust and similar in direction for every participant).

Questions and aims of the studies

Taken together, there are a ton of questions regarding implicit statistical learning that needs to be clarified. The nature of the resulting statistical knowledge should be assessed for each (possible) subtype of statistical learning – considering modality, the type of statistics embedded in the task, etc. so that we could get to a conclusion about the theoretical questions (what factors matter and how). Also, psychometric properties of the tasks used should be routinely reported, along with the observed individual variability in a particular experiment and the assessment of possible artifacts biasing the results. Only this way could we be sure that the theory that we build is not the by-product of questionable methodology.

Admittedly, this is a very ambitious goal requiring lots of investment. In the present Dissertation, I present four studies covering only a tiny slice of these goals: to increase our knowledge about the nature of implicit statistical learning that could be measured with the ASRT (J. H. Howard & Howard, 1997) task, to learn about the psychometric properties of the task, and to improve the analysis methods to overcome its flaws.

STUDY 1

In **Study 1** the main question was whether perceptual information is learned in a temporally structured visuomotor sequence such as the ASRT (in addition to motor sequencing), and if it is learned, then is perceptual learning comparable to motor learning in the paradigm. In order to assess this question, we modified the ASRT task so that stimuli always appeared in the center of the screen (and their identity was differentiated based on perceptual features rather than the location of appearance), this way minimizing the movements of the eye.

Methods

Participants

Thirty-four healthy right-handed individuals took part in the experiment. Half of the participants were randomly assigned to the perceptual condition (mean age M = 21.76 years, SD = 2.02; 7 male/10 female), and the other half were assigned to the motor condition (mean age M = 21.76 years, SD = 1.64; 8 male/9 female). Participants did not suffer from any developmental, psychiatric, or neurological disorders. All participants provided signed informed consent agreements and received no financial compensation for their participation.

Tasks and procedure

We used a modified version of the ASRT task (J. H. Howard & Howard, 1997), the socalled AS-RT-Race. We created a story about a car race for the task. The stimuli were the left, right, up, and down arrows (5 cm long and 3 cm wide), which appeared on the center of the screen. When the stimulus appeared on the screen, it represented the car's direction. For example, when the participants saw an up arrow, they had to press the up button on the keyboard to move the car forward, the left button to turn left, and so on. All participants pressed the keys with their dominant hand.

After the starting block of 85 random presses, they were told that there was a car crash and the steering wheel failed. The car now kept going to the left if they wanted to go straight, but by turning the steering wheel right they could correct this malfunction, and could continue to go straight. Thus participants had to mentally rotate the arrows (the steering wheel) by 90° to the right, and press the button corresponding to this rotated arrow.

In the learning phase, five practice blocks were presented (these were excluded from the analysis), followed by 20 learning blocks with 85 key presses in each block. The stimulus remained on the screen until the participant pressed the correct button. The next stimulus appeared after a 120-ms delay (response to stimulus interval) after the participant's correct response. Participants were told to respond as fast and as accurately as they could.

After the learning phase, the participants were told that the car had been taken to a service station and the steering wheel had been fixed. They were told to use the answer keys corresponding to the arrows that appeared on the screen (up button for up arrow, left button for left arrow, etc.). In the Transfer Phase, half of the participants were assigned to the perceptual condition and the other half to the motor condition (**Fig. 1a**). In the perceptual condition, participants responded to the sequence seen during the learning phase (e.g. 2–R–3–R–1–R–4–R, **Fig. 1b**), and the appropriate key presses represented a new sequence (also 2–R–3–R–1–R–4–R), which they had not practiced before. In contrast, participants in the motor condition had to respond by key presses practiced before (e.g. 3–R–4–R–2–R–1–R, **Fig. 1b**) but the corresponding stimuli on the screen followed another sequence (also 3–R–4–R–2–R–1–R), which they had not seen before. Thus, in the perceptual condition, the perceptual sequence was the same but the motor sequence (key presses) changed compared with the previously practiced sequence. However, in the motor condition, key presses followed the previously learned sequence and the perceptual information (the sequence of the stimuli displayed on the screen) changed. By comparing the participant's performance between the two conditions, we could determine whether the perceptual and the motor component had the same or different effects on learning.

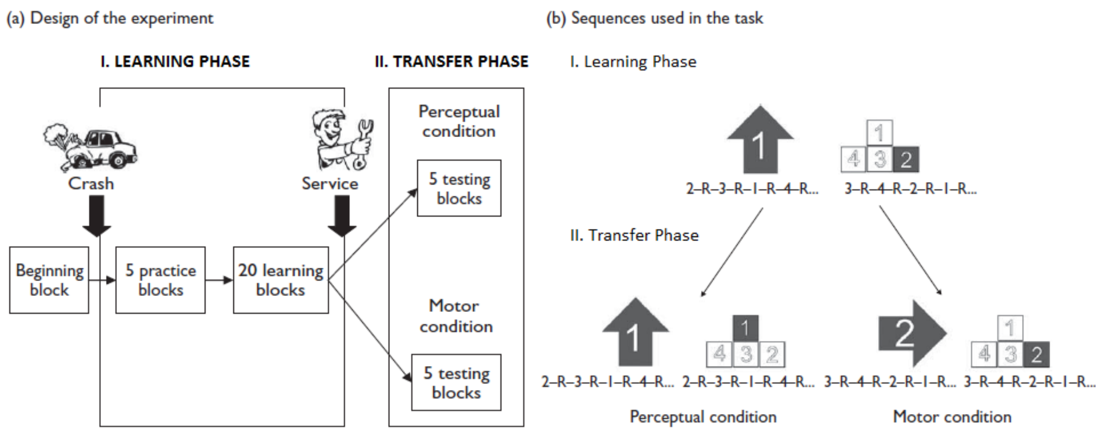

Figure 1. a) Schematic diagram of the experiment. b) In the perceptual condition, the perceptual sequence was the same and the motor sequence (button presses) changed compared with the sequences in the Learning Phase. In the motor condition, key presses followed the learned sequence and the preceptual information changed.

Results

Learning phase

The 2 (triplet: high and low) 4 (epochs: 1–4) repeated measures analysis of variance with condition (perceptual vs. motor) as the between-subject factor revealed sequence-specific learning, indicated by a significant main effect of the triplet: $F(1,23) = 124$, mean square error $MSE = 56.65$, $p < 0.001$, $\eta_p^2 = 0.63$, as well as general motor skill learning shown by the significant main effect of the epoch: $F(4,20) = 8.85$, $MSE = 32.53$, $p < 0.001$, $\eta_p^2 = 0.72$, thereby suggesting that the more the participants practiced, the faster their responses became (**Fig. 2a** and **Fig. 2b**). The two groups (perceptual and motor conditions) did not differ either in sequencespecific or in general motor skill learning ($p > 0.31$).

Transfer phase

To compare the perceptual and motor conditions in the transfer phase, a 2 (triplet: high and low) 2 (epochs: 4–5) repeated-measures analysis of variance was conducted with condition (perceptual vs. motor) as the betweensubject factor. The main effect of the triplet was significant, $F(1,32) = 69.72$, $MSE = 139.36$, $p < 0.001$, $\eta_p^2 = 0.69$, such that participants responded faster for high-frequency than for low-frequency triplets (**Fig. 2c**). The main effect of the epoch was also significant, $F(1,32) = 115.4$, $MSE = 1448.27$, $p < 0.001$, $\eta_p^2 = 0.78$, whereby participants were faster in the Transfer Phase (455 ms) than in the learning phase (525 ms). Interestingly, the triplet epoch interaction was also significant, $F(1,32) = 5.75$, $MSE = 117.79$, $p = 0.02$, $\eta_p^2 = 0.15$, thereby suggesting that the sequence-specific knowledge decreased between the learning and the transfer phases (the RT difference between the high-frequency and low-frequency triplets was 21 ms in epoch 4 and 12 ms in epoch 5). However, despite this decrease, participants still showed a significant triplet type effect in epoch 5, indicated by a one-sample t-test: $t(33) = 4.52$, $p < 0.001$. In addition, there was no difference between the conditions either in sequence-specific ($p = 0.38$) or in general motor skill ($p = 0.10$).

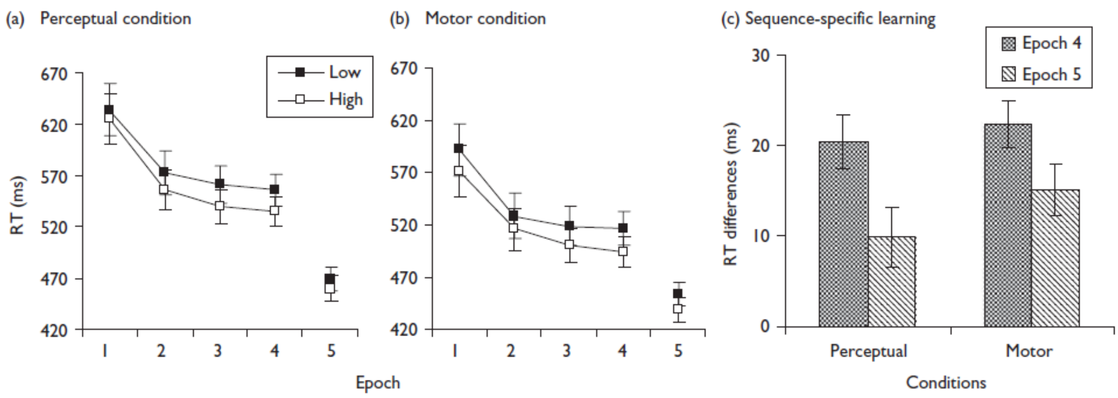

Figure 2. Results of the Learning Phase (Epochs 1-4) and Transfer Phase (Epoch 5) for perceptual (a) and motor (b) conditions. Filled squares represent low-frequency triplets; open squares represent high-frequency triplets. Comparing the sequence-specific knowledge (the reaction time (RT) differences between high-frequency and low-frequency triplets) of perceptual and motor conditions (c). Error bars indicate standard error of mean (SEM).

Discussion

Our results showed that under this research paradigm, both motor and perceptual transfer was significant. These results support the different methods of Song et al. (Song, Howard, & Howard, 2008), which showed perceptual learning with probabilistic sequence learning tasks. In contrast, our results partly differ from that of Willingham et al. (Willingham, Wells, Farrell, & Stemwedel, 2000), which did not find perceptual learning to be an important element of learning.

STUDY 2

In **Study 2** we extended the findings of **Study 1** with assessing consolidation of these different learning types with the inclusion of off-line periods either including sleep or not.

Methods

Participants

There were 102 individuals (students attending the University of Szeged) in the experiment (mean age M = 22.34, SD = 3.82; 44 males, 58 females). None of them suffered from any developmental, psychiatric or neurological disorders. Participants were randomly assigned to the perceptual group or to the motor group. The perceptual and motor groups were further divided by the length of delay (12- or 24-h delay) and by the daytime (morning-first, AM-PM/ AM-AM and evening-first, PM-AM/PM-PM). The eight experimental groups did not differ in their sleep quality, $F(7,89) = 0.98$, $p = .45$. measured by the Pittsburgh Sleep Quality Index (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989) (Due to data collection scheduling problems five out of 102 participants failed to administer this test). All individuals provided signed informed consent, and received no financial compensation for their participation.

Tasks and Procedure

In our study, we used the ASRT-Race task (Nemeth, Hallgató, Janacsek, Sándor, & Londe, 2009) to examine the possible difference in the magnitude of motor and perceptual learning after a 12-h and a 24-h retention period. In addition, we also aimed at exploring the role of sleep in offline consolidation of these two factors of skill learning. Therefore a 12-h delay was administered between the Learning Phase and Transfer Phase of the experiment, during which participants either had a sleep (night group) or they were awake (day group). In order to avoid a time-of-day effect we also administered a 24-h delay condition.

Results

Learning Phase

To be able to investigate the effect of transfer after 12- and 24-h delay, the learning in Session 1 must be similar in the groups. From this point of view, the end of the Learning Phase is crucial (Nemeth and Janacsek, 2011; Nemeth et al., 2010b; Press et al., 2005; Song et al., 2007). Therefore, we analyzed the SLE of the last five blocks of the Learning Phase for every group. Univariate analysis of variance (ANOVA) was conducted with CONDITION (perceptual vs motor), DAYTIME (morning-first vs evening-first groups) and DELAY (12- and 24-h) as betweensubject factors. ANOVA revealed significant sequence learning, $F(1,94) = 32.31$, $p < 0.001$] which is inferred from the test whether the overall mean is different from zero (Mean SLE = 11.16 msec). There were no other significant main effects or interactions involving CONDITION, DAYTIME and DELAY (all $p > 0.32$), thus these between-subject factors had no significant effect on sequence learning.

Transfer of SLE from the Learning Phase to the Transfer Phase

To determine whether the performance in the Transfer Phase declined, improved, or was constant in relationship to the end of the Learning Phase, we subtracted the SLE-score of the last five blocks of the Learning Phase from the SLE-score of the Transfer Phase (Transfer-SLE). As the groups were similar in SLE at the end of the Learning Phase, any difference among groups in Transfer-SLE could be attributed to the differential effects of consolidation. We conducted a univariate ANOVA for this Transfer-SLE-score with CONDITION (perceptual vs motor), DAYTIME (morning-first vs eveningfirst groups) and DELAY (12- and 24-h) as between-subject factors. ANOVA revealed a main effect of CONDITION, $F(1,94) = 4.92$, $p = .029$, the motor group showing larger SLE than the perceptual group (**Fig. 3**). ANOVA showed no significant main effect or interaction with DAYTIME (all $p > 0.45$), suggesting that the AM-PM, PM-AM, AM-AM and PM-PM groups did not differ in their SLE. In addition, main effect and interactions with DELAY were not significant either (all $p > 0.25$), suggesting that 12- and 24-h delay groups performed at a similar level.

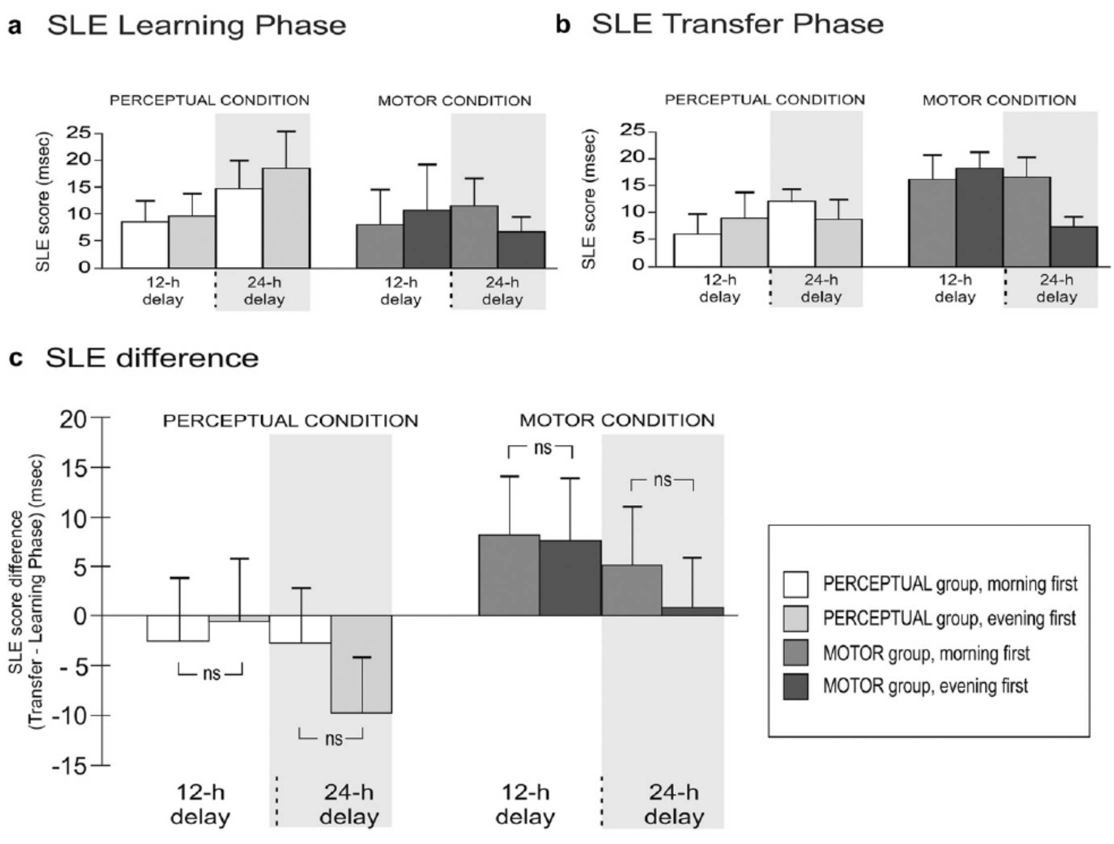

Figure 3. (a) SLE-score of each experimental group in the last five blocks of the Learning Phase. (b) SLE-score of each experimental group in the Transfer Phase (Session 2). (c) Difference between SLE-scores of the five blocks of Transfer phase and the last five blocks of Learning phase (Transfer-SLE-score). The perceptual groups showed weaker transfer effect than the motor groups both after 12 and 24 h. Error bars indicate Standard Error of Mean.

Thus, the only significant effect in the ANOVA was the main effect of CONDITION, suggesting differential consolidation of perceptual and motor groups with better consolidation for the motor group, irrespective of the delay or daytime. Despite this difference in consolidation, SLE in the Transfer Phase was significantly different from zero for both the perceptual and motor groups (one-sample t-tests for SLEscores: $t(49) = 5.25$, $p < .001$ and $t(51) = 8.72$, $p < .001$ respectively). Thus, in spite of the weaker consolidation in the perceptual group, they still showed significant SLE in the Transfer Phase.

Discussion

Despite the 12-h or the 24-h offline period we found a significant perceptual and motor learning effect in the Transfer Phase, however the transfer of motor knowledge was more robust, irrespective of whether sleep occurred in the consolidation period or not. These results have important implications for the perceptual/motor and also for the sleep debate in skill learning in the following ways: (1) Previous experiments in this field included only one session which can reveal short-term performance changes in behavior. (2) Sleep has no contribution to this type of learning. (3) The retention period itself (regardless of sleep) has a modifying effect on the consolidation of perceptual/motor knowledge and the underlying brain networks.

STUDY 3

In **Study 3** we assessed the question of interference between similar (but different) sequences learned in succession; whether the sequence learned in the first place could be „overwritten” with a second sequence, whether there are costs associated with the proactive interference caused by the first sequence; and whether it really gets „overwritten” (rewired) or the knowledge for both sequences is accessible later. Finally, an important question related to the effect of explicit (top-down) knowledge about the rule (but not about the statistical structure) embedded in the sequence, and whether this knowledge – or the differences in participants mindsets owing to this knowledge – results in differences in implicit statistical learning measured on trials on which the explicit knowledge could not be utilized.

Methods

Participants

Eighty-four healthy young adults took part in the experiment. Participants were recruited at University of Szeged and were randomly assigned to one of three groups: the Implicit-Implicit group ($N = 28$; 20 females; Age: $M = 20.46$ years, $SD = 2.10$), the Implicit-Explicit group ($N = 28$; 17 females; Age: $M = 22.14$ years, $SD = 1.96$), and the Explicit-Explicit group ($N = 27$; 18 females; Age: $M = 22.54$ years, $SD = 3.33$). One participant was excluded from the analysis because errorneously the same sequence was administered to him on each day of the study. Participants did not suffer from any psychiatric or neurological disorders. Prior to their inclusion in the study, participants provided informed consent to the procedure as approved by the research ethics committee of University of Szeged, Szeged, Hungary. The study was conducted in accordance with the Declaration of Helsinki and participants received course credits for taking part in the experiment.

Tasks and Procedure

We used the original *Implicit variant* of the ASRT task (J. H. Howard & Howard, 1997), and its *Explicit (cued) variant* (Nemeth, Janacsek, & Fiser, 2013; Song et al., 2007a, 2009). The latter version differed from the implicit task in three respects: firstly, random and pattern stimuli appeared in different colors (pattern elements appeared green while random elements appeared blue). Secondly, participants were instructed to pay attention primarily to the four-element long pattern (the green trials) to be able to report it after each block. Finally, the feedback after the blocks did not contain information about RTs and erroneous buttonpresses on random trials as the instruction highlighted performance on the pattern trials. Importantly, responses given to pattern trials were never analysed, only performance on random trials that could not be explicitly anticipated. This way the data gathered with the explicit and implicit variants of the task were comparable, and results reflect the implicit statistical learning rather than the knowledge about the pattern.

After learning an ASRT sequence (referred to as Sequence A) on the first day of the study (in the *Learning Phase*), participants were given a different ASRT sequence (referred to as Sequence B) on the second day (in the *Rewiring Phase*). The two sequences shared some of their transitional, meaning that at some points in Sequence B the most probable upcoming stimulus was the same as in Sequence A (*unchanged* sequence parts). Other transitional probabilities changed: the most probable continuation of the previous trials was different from that on the previous day (*changed* sequence parts). This way we could compare learning with and without interference from the previous learning episode. The *Implicit-Implicit* group performed the implicit version of the task in both the *Learning Phase* and in the *Rewiring Phase*; the *Implicit-Explicit*group performed the implicit version in the *Learning Phase* and the explicit version in the *Rewiring Phase;*finally, the *Explicit-Explicit* group performed the explicit version in both phases.

On the third day of the study – in the *Follow-up Phase* – the magnitude of statistical knowledge for both sequences was assessed to investigate possible retroactive interference effects after a 24-hour consolidation period.

Results

We measured statistical learning as (**a**) difference in reaction times (RTs) given to anticipated (probable) stimuli in contrast to unexpected (less probable) stimuli, termed as Statistical Learning Effect (SLE), and (**b**) by determining whether erroneous responses reflect anticipations of the most probable stimuli in cases when less probable trials came up. Statistical learning measured by the SLE score was evident in both the Learning and Rewiring Phase (see 95% confidence intervals, CIs, on **Fig. 4**). In the Learning Phase, there could not possibly be any interference effects as only Sequence A had been introduced yet, we nevertheless contrasted the magnitude of learning of those transitional probabilities that were common during both Phases (unchanged sequence parts) and those that were about to change in the Rewiring Phase (changed sequence parts). As expected, there was no difference between the two ($p = 0.568$, Cohen's $d = 0.080$), indicating that they were equally easy to learn (**Fig. 4**, light vs. dark grey bars). In the Rewiring Phase, however, we found smaller statistical learning for the changed sequence parts ($p < 0.001$, $d = 0.829$) compared to the unchanged sequence parts (**Fig. 4**, light grey vs. blue bars). This was apparent in the Implicit-Implicit ($p < 0.001$, $d = 1.425$) and Explicit-Explicit groups ($p = 0.008$, $d = 0.737$), but not in the Implicit-Explicit group ($p = 0.128$, $d = 0.406$). These results suggest that the Implicit-Explicit group was the most successful in adapting to the new statistical regularities, even to the extent that their rewired knowledge was not much different than that of the unchanged sequence parts.

On some of the trials, participants pressed a key that did not correspond to the stimulus. Some of these errors reflected anticipations of the most probable stimulus when the actual stimulus was a less probable one, termed as *anticipatory errors*. As two sequences were taught, we can measure anticipations of Sequence A's most probable transitions and those of Sequence B's (errors that could be regarded as anticipations of both sequences were not analysed). We compared the proportion of anticipatory errors to each other (anticipations of Sequence A vs. anticipations of Sequence B), and to a baseline proportion that could be expected by chance (16.67%, see **Fig. 5)**.

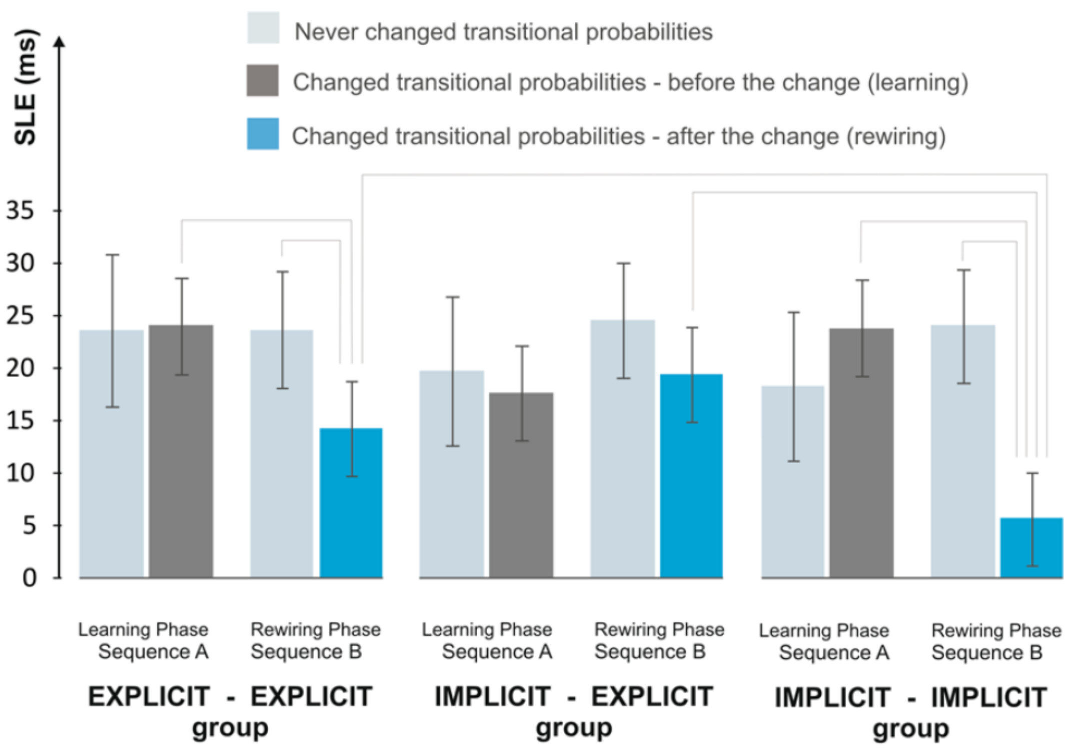

Figure 4. Statistical learning effect (SLE) in the Learning and Rewiring Phase. The magnitude of SLE indicates the difference of reaction times (RTs) given to frequent transitions (more probable stimuli) in contrast to infrequent transitions (less probable stimuli). Some of the transitions had constant frequency in the Learning Phase and Rewiring Phase (unchanged transitions, light grey bars), while other transitions swapped their frequency – previously infrequent transitions became frequent in the Rewiring Phase, and vice versa (changed transitions, dark grey bars – before the change, blue bars – after the change occurred).

As expected, the Learning Phase was dominated by anticipations of Sequence A (dark grey bars on **Fig. 5)**, while the Rewiring phase was dominated by anticipations of Sequence B (both $p < 0.001$, both $d > 1.061$, blue bars on **Fig. 5**). From another point of view, there were less anticipations of Sequence B in the Learning Phase than in the Rewiring Phase, and vice versa (both $p < 0.001$, both $d > 0.979$). This pattern of results was observed in all groups, although effect sizes were substantially smaller in the the Implicit-Implicit group (both $d < 0.672$) than in the other groups (all $d > 1.226$). Most importantly, anticipations of Sequence B in the Rewiring Phase (that indicate adaptation to the new sequence structure) were less pronounced in the Implicit-Implicit group than in the Implicit-Explicit group ($p = 0.047$, $d = 0.721$), while anticipations of Sequence A in the same Phase (indicating the continuing influence of the knowledge gained in the Learning Phase) were more pronounced in the Implicit-Implicit group than in the Implicit-Explicit and the Explicit-Explicit groups (both $p < 0.036$, both $d > 0.795$). The Implicit-Implicit group showed no significant difference in proportions of anticipating Sequence A and Sequence B during the Rewiring Phase ($p = 0.529$, $d = 0.225$), both being above chance level (see 95% CIs on **Fig. 5**). These results clearly point to the continuing influence of the no-longer valid statistical knowledge gained in the Learning Phase – that is, proactive interference – in the Implicit-Implicit group.

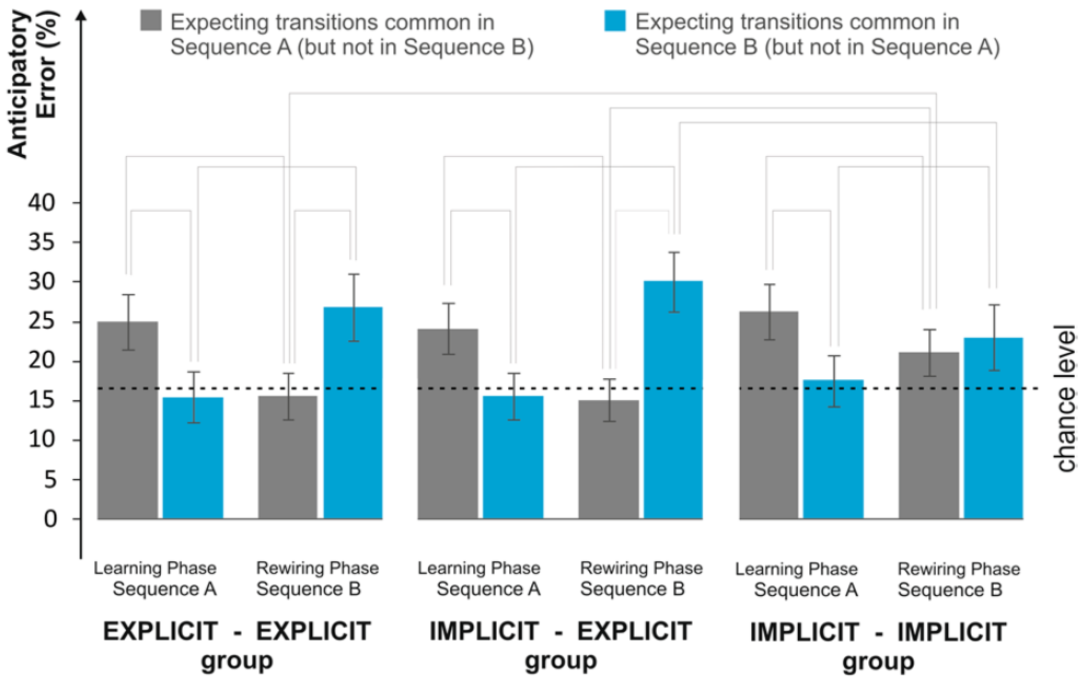

Figure 5. When a less probable stimulus came up, participants sometimes erroneously pressed the key corresponding to the most probable stimulus, termed as anticipatory errors. As two (partly) different sequences were taught, we differentiated between anticipations of Sequence A's most probable stimuli, and that of Sequence B's. Percentage of anticipatory errors of Sequence A (learned in the Learning Phase, grey bars) and Sequence B (learned in the Rewiring Phase, blue bars) over the two Phases, and chance level for anticipatory errors (dotted line) are shown. The solid lines connecting the bars indicate significant differences ($p < 0.05$). Error bars represent 95% confidence intervals (CIs).

Participants were retested on the third day for both sequences to test whether the first one became overwritten by the second one. Participants showed better performance on the transitions that were frequent in both the Learning and Rewiring Phases than on those that were frequent in only one of the Phases ($p = 0.003$, $d = 0.506$), which is not surprising given that the former ones were practiced almost twice as much. More importantly, better performance was expressed for Sequence B than for Sequence A ($p = 0.015$, $d = 0.404$). This pattern of results may indicate that statistical knowledge for Sequence A became partly overwritten by knowledge of Sequence B, showing retroactive interference which is beneficial for the rewiring process. No group differences were observed (ps > 0.303).

When performing Sequence A on the third day of the study, anticipations of Sequence A were more common than anticipations of Sequence B ($p = 0.004$, $d = 0.533$), and than what might have been expected by chance. When performing Sequence B, on the other hand, anticipations of Sequence B outnumbered anticipations of Sequence A ($p = 0.009$, $d = 0.503$), and were more numerous than expected by chance. From another point of view, anticipations of Sequence A were significantly more pronounced when performing Sequence A than when performing Sequence B, and vice versa (both $p < 0.003$, $d > 0.494$). This pattern of results indicate no proactive or retroactive interference effects, as participants were able to quickly adapt to changes in the statistical structure, and suggests that knowledge about the two statistical structures coexist and can be adaptively used in the appropriate situation. No group differences were observed ($p = 0.745$).

Discussion

In summary, we found successful rewiring of the acquired knowledge in all three experimental groups. In the Rewiring Phase the group that learned implicitly and rewired with the help of explicit cues (i.e., the Implicit-Explicit group) showed better performance than the other groups. In other words, explicit cues during the rewiring process led to faster adaptation to the changed reguralities. By the end of the rewiring period, all groups showed similar performance suggesting an efficient but slower rewiring in the Implicit-Implicit group as well. We also found evidence that the first learned sequence was accessible when needed, shown by sequence specific anticipatory errors in the Follow-up Phase, although the motor execution of it was not as fluent as the execution of the secondly learned sequence.

STUDY 4

In **Study 4** our goal was two-fold. First, we wanted to show that the ASRT task makes it possible to assess the learning of both second-order and third-order statistical structure without any modification to the task (just by refining the analysis methods), and also assess the question of pattern (rule) learning, i.e. whether participants learn about the alternating structure of the sequence in addition to its statistical properties. We have also compared the currently/typically used analysis methods with the proposed method (in terms of goodness of fit). Second, we assessed the psychometric properties of the task (both with the typical analysis methods and with the newly proposed method), and we suggested the application of a filter to lessen the impact of pre-existing (cognitive or biomechanical) biases to certain stimulus combinations which could result in artifacts in the learning scores.

Theoretical considerations

In the ASRT task, statistical properties of trials are confounded with trial type (e.g. whether a particular trials identity follows the pattern or is determined randomly). For this reason it is hard to separete the possible effects of the two factors, although the question was approached by creating categories that differentiated between three types of trials: random-ending high frequency triplets, pattern-ending high-frequency triplets and random-ending low frequency triplets (Janacsek et al., 2012; Kóbor et al., 2018; Nemeth, Janacsek, & Fiser, 2013; Schwartz et al., 2003; Simor et al., 2019).

In Study 4 we suggested to go deeper and assess four consecutive trials (i.e. quads) instead of three consequtive trials (triplets) as a basic unit of the analysis because high frequency triplets can further be divided into two subroups based on the N-3th trial; quads that are moderately frequent and quads that are very frequent. Furthermore, it gets possible to separate two types of statistical learning**, joint probability learning** (meaning that we learn that particular combinations are frequent and others are not) and **conditional probability learning** (meaning that a particular trial could either be very probable or less so given the N-3th- N-1th trials), as the two dissociates on the level of quads.

As a concrete manifestation of our proposal, we distinguised five models that couls serve as the basis of analysis of the Task. Model 1 only considers trial type (whether a particular trial is pattern or random); Model 2 only considers triplet level statistical information (whether a particular triplet is of high or low frequency); Model 3 considers both of the above; Model 4 considers quad level statitsical information, and finally, Model 5 considers both quad level statistics and trial type. Importantly, Model 1-3 has been routinely used in the literature, while Model 4-5 has been introduced in our paper; for the specific learning scores of the Models see **Table 1**.

Table 1. Specific Learning Scores of the different Models

	Specific learning scores that can be quantified with a Model			
Model 1	Trial type Learning			
Model 2	Triplet Learning			
Model 3	Triplet Learning,	Higher Order Learning,		Max Learning
Model 4	Triplet Learning,	Quad Learning,		Max Learning
Model 5	Triplet Learning,	Quad Learning,	Pattern Learning,	Max Learning

Confounding variables in the ASRT task

For some stimulus combinations, e.g. serial repetitions of the same stimuli, response facilitation is observed when contrasted with other combinations, e.g. an inconsistent pattern of alternations and repetitions. These so-called *sequential effects* (Remington, 1969) are evincible in random streams of stimuli, but also from reaction time tasks in which the conditional probabilities of stimuli vary, see (Kornblum, 1973). We have no solid idea of exactly which combinations should be relatively „easier" (facilitated) compared to others because this phenomenon has mostly been studied in binary-choice reaction time tasks (Kirby, 1976; Soetens, Boer, & Hueting, 1985;

Vervaeck & Boer, 1980; but see Lee, Beesley, & Livesey, 2016), those combinations being less numerous and less complex than the combinations in the ASRT task. In the absence of concrete expectations of how and to what extent sequential effects occur in the ASRT (and bearing in mind that ASRTs with different RSIs may differ in this regard), the wisest thing we can do is to ensure that the groups of trials that are to be contrasted in the ASRT (e.g. pattern vs. random trials or highly predictable vs. moderately/slightly predictable trials, etc.) belong to the same types of combinations with respect to local sequential effects ("easy" or "hard").

In this work, we propose the elimination of such biases on the level of quads. We created quad categories based on their abstract structure. Our notations were derived the following way: whatever the current stimulus was (position 1, 2, 3 or 4), it was denoted as *„a"*. If the previous stimulus was identical to the current one, it was also denoted as *„a"*, thus the combination of the two was denoted as *„aa"*. Otherwise, if the previous stimulus was different, the combination was denoted as *„ba"*. Going further, if the N-2th trial was identical to the N-1th or Nth trial, it was denoted with the same letter as the one that it was identical to (e.g. „aba" or „bba"); otherwise, it got the following letter from the alphabet (e.g. *„c"*). This way a quad that consistsed of four different stimuli was always denoted as *„dcba"* (irrespective of whether it was derived from 1-2-3-4, 3-1-4-2 or else). Only three out of 13 categories are counterbalanced across the groups of trials being compared within subjects (e.g. pattern vs. random trials in Model 1, or high frequency vs. low frequency triplets in Model 2, etc.) and across participants (i.e. any particular quad having an equal chance of belonging to either statistical category). These quad types are *dcba, cbba* and *acba.* Hereinafter, we will refer to this filtering method as *Quad Filtering*.

Methods

Participants

One hundred and eighty healthy young adults participated in the study, mean age M = 24.64 (SD = 4.11), Min_{age} = 18, Max_{age} = 48; 28 male/152 female. All participants had normal or corrected-to-normal vision and none of them reported a history of any neurological and/or psychiatric condition. All participants provided written informed consent before enrollment and received course credits for taking part in the experiment. The study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) in Hungary (Approval number: 30/2012) and by the research ethics committee of Eötvös Loránd University, Budapest, Hungary. The study was conducted in accordance with the Declaration of Helsinki.

Tasks and Procedure

The Alternating Serial Reaction Time (ASRT) task was used to measure statistical learning capabilities of individuals (J. H. Howard & Howard, 1997). The ASRT task consisted of 45 presentation blocks in total, with 85 stimulus presentations per block. After each of these training blocks, participants received feedback about their overall RT and accuracy for 5 seconds, and then they were given a 10-s rest before starting a new block. Each of the three sets of 15 training blocks constitutes a training session. Between training sessions, a longer (3–5 min) break was introduced.

Results

Comparison of the Models

Reaction Times - We analysed the same dataset five times (corresponding to the five models). We computed individual adjusted R^2-s for each epoch of each participant as a way of assessing the goodness of fit of each Model; since there were nine epochs, this resulted in nine values per participant. These were than averaged to yield a single value for everyone. The effect of different filtering methods was also taken into account by computing these effect sizes for each filtering type separately (No Filter, Triplet Filter – which is the typically used filter in the literature; and Quad Filter). The goodness of fits were then compared by a FILTER TYPE (3 levels: No Filter, Triplet Filter, Quad Filter) x MODEL (5 levels: Model 1 - Model 5) Repeated Measures ANOVA. Sphericity was assessed with Mauchly's Test, and if this precondition was not met, degrees of freedom were adjusted with the Greenhouse-Geisser method. Bonferroni-corrected post hoc tests were performed whenever the omnibus ANOVA showed significant main effects or interactions. Partial eta squared effect sizes are reported in line with significant main effects or interactions in the ANOVA.

The main effect of FILTER TYPE was significant, $F(1.553, 278.066) = 25.562$, $MSE < 0.001$, $p < 0.001$, $\eta_p^2 = 0.125$, indicating that, on average, the goodness of fits differed as a function of the filter used. Bonferroni corrected post hoc tests revealed that means of adjusted R squared values were highest with the Quad Filter and lowest with the Triplet Filter, all contrasts being significant ($p < 0.001$) except for the contrast No Filter vs. Quad Filter ($p = 0.569$). The main effect of MODEL was also significant, $F(1.384, 247.759) = 408.371$, $MSE < 0.001$, $p < 0.001$, $\eta_p^2 = 0.695$, indicating that model goodness of fits differed as a function of the Model used in the analysis, as values grew monotonicaly from Model 1 to Model 5. Bonferroni corrected post hoc tests revealed that all paired comparisons were significant (all $p < 0.001$). Finally, the interaction of FILTER TYPE x MODEL was also significant, $F(2.492, 446.058) = 11.122$, $MSE < 0.001$, $p < 0.001$, $\eta_p^2 = 0.058$, indicating that the monotonic growth of adjusted R squared values as a funtion of MODEL were not equivalent with the three filtering methods used. Bonferroni corrected post hoc tests revealed that each Model differed from all the others within each filtering method (all $p < 0.012$). The effect of the differing filters was also quite consistent with each Model, showing that both the No Filter condition and the Quad filter condition yielded higher fits than the Triplet Filter condition (all $p < 0.001$), the Quad Filter and No Filter condition not differing from each other in 4 out of 5 cases (all $p > 0.437$, except for Model 2 where $p = 0.006$). The results are shown on **Fig. 6a**).

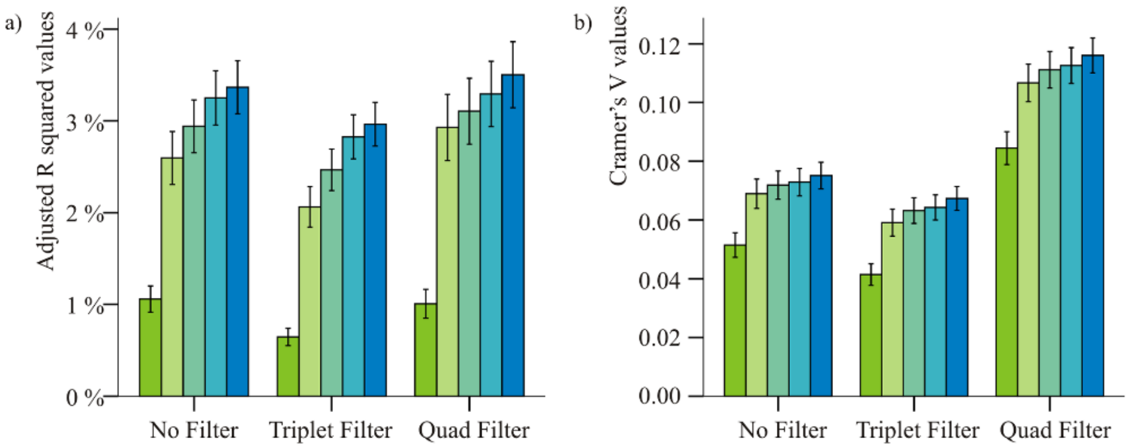

Figure 6. Goodness of fit of the different models within each filtering method. a) Individual Adjusted R2 values based on reaction times. Each Model differed from all the other Models within each filtering method (all $p < 0.012$). b) Individual Cramer's V values based on error data. Each Model differed from all the others within each filtering method, except for the differences Model3 vs. Model4 (no filter $p = 0.166$, triplet filter $p = 0.359$, quad filter $p = 0.261$). Error bars are 95% confidence intervals.

Errors - The goodness of fit of the different models were than calculated in the form of Cramer's V values (data from the nine epochs were collapsed into a single category due to the small number of errors) separately for each filtering method. To compare the obtained Cramer V values, we run a FILTER TYPE (3 levels: No Filter, Triplet Filter, Quad Filter) x MODEL (5 levels: Model 1 - Model 5) Repeated Measures ANOVA. Sphericity was assessed with Mauchly's Test, and if this precondition was not met, degrees of freedom were adjusted with the Greenhouse-Geisser method. Bonferroni-corrected post hoc tests were performed whenever the omnibus ANOVA showed significant main effects or interactions. Partial eta squared effect sizes are reported in line with significant main effects or interactions in the ANOVA.

The main effect of FILTER TYPE was significant, $F(1.472, 263.495) = 489.885$, $MSE = 0.002$, $p < 0.001$, $\eta_p^2 = 0.732$, indicating that, on average, the goodness of fits differed as a function of the filter used. Bonferroni corrected post hoc tests revealed that means of Cramer V values were highest with the Quad Filter and lowest with the Triplet Filter, all contrasts being significant ($p < 0.001$). The main effect of MODEL was also significant, $F(1.598, 286.124) = 281.264$, $MSE = 0.001$, $p < 0.001$, $\eta_p^2 = 0.611$, indicating that model goodness of fits differed as a function of the Model used in the analysis, values grew monotonicaly from Model 1 to Model 5. Bonferroni corrected post hoc tests revealed that all paired comparisons were significant (all $p < 0.001$) except for the difference between Model3 and Model4 ($p = 0.231$). Finally, the interaction of FILTER TYPE x MODEL was also significant, $F(1.747, 312.721) = 40.517$, $MSE < 0.001$, $p < 0.001$, $\eta_p^2 = 0.185$, indicating that the monotonic growing of adjusted R squared values as a funtion of MODEL were not equivalent with the three filtering methods used. Bonferroni corrected post hoc tests revealed that each Model differed from all the others within each filtering method, except for the differences Model3 vs. Model4 (no filter $p = 0.166$, triplet filter $p = 0.359$, quad filter $p = 0.261$). The effect of the differing filters was also quite consistent with each Model, showing all filtering methods differed from the rest (all $p < 0.001$). The results are shown on **Fig. 6b**).

Comparison of the Filters

To assess whether the filtering method had an effect on individual effect sizes of different types of learning that could be detected with the Models, we first run Repeated Measures ANOVA-s on the Cohen's d values obtained for all the possible learning measures of the five Models with FILTER (no filter, triplet filter, quad filter) as an independent variable. Filter had an effect in all cases (all $p < 0.001$, all $\eta_p^2 > 0.164$), except for the *pattern learning* measure of Model5, which remained unchanged ($p = 0.626$, $\eta_p^2 = 0.003$). In cases of significant omnibus ANOVAs, Bonferroni corrected post hoc tests were run. Triplet filtering (in contrast to no filtering) left some of the learning measures unaffected. In all the remaining cases individual effect sizes decreased as a result of triplet filtering (all $p < 0.048$). Quad filtering, on the other hand, resulted in mixed effects. It increased effect sizes obtained in the simple models Model1 and Model2, and in the more elaborated Models (3-5) it increased those effects that depicted higher order statistical learning measures (all $p < 0.001$). It is worth noting that some of these values not only increased but reversed their direction when applying the quad filter, leading to qualitatively different conclusions about learning.

Variability

To test the homogeneity of variances, Levene-test was applied on individual learning scores with FILTERING (No Filter, Triplet Filter, Quad Filter) as an independent variable (for this particular analysis treated as a between-subjects variable). According to the test, filtering had a significant effect on variances in most of the cases ($p < 0.032$); the overall pattern was that Quad Filtering resulted in higher variability. The exceptions were the Higher Order Learning score of Model 3; $F(2, 537) = 1.683$, $p = 0.187$; and the Quad Learning score in Model 4; $F(2, 537) = 1.977$, $p = 0.140$, in which cases filtering did not have an effect on the variability.

Does higher variability go in hand with lower reliability?

We calculated split-half reliability of all of the measures by randomly assigning each keypress to one of two categories; the individual effect sizes were then computed for both sets, and the correlation of the two values was computed. Our results showed that reliability indices dropped substantially when using the quad filter (e.g reliability of triplet level learning scores dropped from .691 to .556). This may be attributable to the possibility that pre-existing biases are a form of a systematic artifact (rather than noise), as noted earlier. Conversely, it is also possible that the drop is attributable to increased levels of noise (since fewer trials are analyzed with stricter filtering).

What is being learned in the ASRT?

Since Model 5 made it possible to assess four different types of learning (i.e. triplet level statistical learning, quad level statistical learing, pattern learning and overall leanring), it was of our interest to quantify how many of our participants showed evidence of particular learning types. Our analysis revealed that triplet level learning was apparent in 77% of the cases, quad learning in 12% of the cases, pattern learning in 5% of the cases and overall (mixed) learning in 87% of the cases.

Discussion

In this study, we discussed in detail the many possible information types that could be learned in the ASRT task (such as pattern learning and different levels of joint frequency learning and conditional probability learning), and our concerns that these types of learning are not sufficiently differentiated by the currently used analysis methods. Moreover, as we have shown, the learning measures that are typically extracted from data might be biased by pre-existing tendencies to certain stimulus combinations, indicating that the ASRT does not measure (only) what it supposed to. We provided a presentation of how different analysis methods and filtering methods result in different levels of artifacts and biases, a hopefully practical aid for the (re)interpretation of the results obtained with the task. We also proposed new analysis methods (with a somewhat new terminology) and a filtering method that eliminates at least some of the biases discussed so far and thus can be used in future studies (or for reanalyzing already existing datasets).

We believe that the ASRT task is a great tool for measuring implicit sequence learning and memory – it might even be more promising than we ever thought. However, in order to get more out of it, we need to improve our analysis methods and take possible confounding factors more seriously. In this paper, we provided a possible solution to these problems. Our results point to the ASRT being primarily a statistical learning task (at least in the short term), where triplet learning occurs for most of the participants but quad learning is the privilege of fewer. We have also shown that these results depend strongly on the filter being used, and for methodological reasons, we suggest the usage of the Quad Filter in the future.

GENERAL DISCUSSION OF OUR FINDINGS

Our studies aimed to gain a better understanding of implicit statistical learning phenomena, and about the diversity of findings of seemingly similar processes. We also emphasized the need to assess the psychometric properties of the measures of implicit statistical learning, since many findings – or null findings – could be a result methodological rather than theoretical issues. In this Dissertation, we addressed many of these topics, such as modality specificity in implicit learning; the (in)dependence from other cognitive domains, types of statistics that could be learned, and methodological considerations in the measurement.

Implicit learning – One or Many?

Modality specificity - In Study 1 and Study 2 we found that both the visual sequence and the motor sequence had been learned by participants, and although similar magnitudes of transfer have been observed for the two types of information when performance was assessed in a single session (without delay), the consolidation of the perceptual information seemed to be weaker than that of the motor information after a delay of 12 or 24 hours.

Importantly, the existence of perceptual statistical learning was inferred from the fact that – in spite of the interference that occurred because of the change in the motor sequence – learning score in the second phase was greater than zero. Two other possibilities could have led to the observed results. **1)** It is possible that no perceptual transfer occurred, only motor learning of the new (interfering) sequence. In Study 2, we quantified learning scores in the first two blocks of Session 1 (before the change in either the perceptual or the motor information occured), and in the first two blocks of Session 2 (i.e. immediately when the change occurred). We found significant learning effects in the latter but not in the former – these results indicate that the learning scores of Session 2 are transferred from Session 1, and not learned anew, since learning of the motor information was not this fast even without interfering information in Session 1. **2)** Second, it is also possible that the transfer is not perceptual in nature, but motor transfer, since the two interfering motor sequences had a few similar chunks that were frequent throughout (a possibility we did not address at the time of publication). However, we analyzed the data and found similar results even when controlling for this possibility (*unpublished results*).

In sum, it has been shown that both kinds of learning occur and that consolidation is different for the two types of learning. These results support the notion of multiple implicit statistical learning submodules rather than a single system that is responsible for all kinds of implicit learning phenomena (in line with Emberson et al., 2011; Li et al., 2018; Walk & Conway, 2016).

Independency from other cognitive abilities - In Study 3 we compared the implicit statistical learning and rewiring abilities of participants who were either given information about the embedded sequence in the ASRT or not. Importantly, even participants performing the explicit version of the task could only anticipate pattern trials explicitly – any statistical learning effects detected on random trials were implicit nevertheless.

We found that the implicit statistical learning was similar across groups when learning the first sequence, thus knowledge about pattern trials did not help participants learn the statistical properties of the task when assessed on random trials. However, when a new sequence was introduced, and proactive interference from the first learning episode had to be overcome, the explicit groups outperformed the implicit group (or more precisely, the implicit group needed more time to overcome the interference than the explicit groups).

One possibility is that those receiving explicit information about every second trial had a different attitude, or were less bored, than the implicit groups, and they might have concentrated more even on random trials, leading to better performance. As an alternative, it is also possible that telling participants about the embedded structure somehow emphasized the statistical structure of the sequence, too. It has been found, for example, that the exaggeration of some features of speech acts as a perceptual catalyst whereby it helps infants discriminate between similar inputs (Karzon, 1985), and that infant-directed speech (e.g. motherese, which also exaggerates important aspects of speech) is a more effective signal for learning phonetic categories than adult-directed speech (de Boer & Kuhl, 2003). Our results could have been similarly caused by the emphasis that explicitly shown pattern trials created.

The fact that knowledge about the sequence in the ASRT somehow affected the implicit processes indicate that implicit and explicit processes are not independent (for a similar conclusion, see Boyd & Winstein, 2003; Destrebecqz et al., 2005; Lagarde et al., 2002). Our results are in line with Boyd & Winstein (2003) who also found facilitation of implicit learning following explicit information in healthy adults –

however, they used a deterministic 10 elements long sequence, and hence from the point when the sequence was explicitly stated, each element of the sequence could have been consciously anticipated. In our case, however, trials where explicit anticipations could have been in play were eliminated. For this reason, we think that our results are stronger support for the notion of interacting systems, although the exact mechanism by which the interaction operates has to be identified.

Type of statistics – Does it matter? - It has been known that humans are capable of both conditional probability learning and joint frequency learning (J. H. Howard et al., 2008; Thiessen et al., 2013; Thiessen, 2017) and it has been suggested that they are the results of independent processes (Thiessen, 2017). In Study 4 we have shown that the ASRT task is adequate to distinguish between the two types of learning (or at least in showing which one dominates for a given individual), and that results point towards a greater dominance of conditional probability learning (13% vs. 5% of participants showing reliable pattern for conditional probability learning and joint frequency learning, respectively).

It has also been known that humans are capable of learning higher-order statistical structure (e.g. four consecutive trials – quads – or even higher levels) (Remillard, 2008, 2010, 2011), and that learning of higher-order information can be selectively impaired (in dyslexia: W. Du & Kelly, 2013; J. H. Howard et al., 2006; in Parkinson's disease: Smith & McDowall, 2004; in Schizophrenia: Schwartz et al., 2003; with age: J. H. Howard et al., 2007; D. V. Howard et al., 2004; Feeney et al., 2002; J. H. Howard & Howard, 1997; Urry et al., 2018). In Study 4 we have shown that – although traditionally only the level of triplets is being assessed – quad-level learning could also be quantified without any modification to the task, just by applying a different analysis method. By reanalyzing the huge amount of already existing ASRT datasets, we could get closer to understand the nature of quad learning, and whether triplet level learning and quad level learning are degrees of the same capability, or they dissociate within subjects. This could be assessed by verifying the correlation between the two in large samples; or by examining the learning curves, e.g. whether quad learing follows triplet level learning or it happens in parallel (for those who are sensitive to these statistics).

The psychometric properties of the ASRT task

Low reliability - Although the relatively low reliability of implicit compared to explicit measures has been acknowledged (Lebel & Paunonen, 2011), reliability indices are rarely included in implicit learning research (but see Siegelman & Frost, 2015). In Study 4 we have shown that the split-half reliability indices of ASRT learning scores vary from 0.02 to 0.84 depending on the type of the analysis (e.g. data grouping and filtering) and on type of learning that is being assessed (e.g. triplet level learning or quad-level learning). In general, triplet learning scores are more reliable than quad learning scores (~0.6 vs. ~0.4), and the reliability of pattern learning is the worst – downright unreliable (~0.15). By using a stricter filter to eliminate the effects of pre-existing biases, individual variability got higher and reliability indices typically got lower, underscoring that without the strict filtering, the performance of different participants is more similar than their true learning abilities – differences just don't show so much. Thus, there is a trade-off between validity and reliability, and this needs to be considered when deciding how to analyze our data.

Low individual variability - It has been assumed that implicit learning is robust and shows small inter- and intraindividual variability (Reber, 1993; Reber & Allen, 2000). Accordingly, individual differences in implicit cognition remain largely unexplored (Reber & Allen, 2000; but see Kaufman et al., 2010; and Kalra et al., 2019).

We add to this field of research by providing inter-individual variability indices for different learning scores in Study 4. We quantified the spread in data both in absolute (standard deviation) and in relative (coefficients of variation) terms; unfortunately, we are not aware of any standards by which we could tell whether the values that we found refute the hypothesis of „small individual variability" or not. We would need similar descriptions of individual variability in other fields of cognitive psychology (e.g. explicit learning capacity) to draw any conclusions.

Another important message of Study 4 is that the variability of performance is sometimes shaped by factors we do not intend to measure (artifacts; e.g. pre-existing biases); in the ASRT such biases make the detectable differences smaller – an effect that has implications from the interpretation of result to the theories of implicit learning.

Issues related to reaction-time based measures - It has been suggested that accuracy is a better measure than difference scores derived from reaction times (Urry et al., 2015, 2018). We did not directly test this hypothesis, but nevertheless contributed to this debate by showing that accuracy measures are less reliable than reaction time measures (Study 4, reliability indices). Also, generally speaking, we found reaction time based measures and error ratio based measures to point in one direction in our experiments, although sometimes complementing each other. For example in Study 3, by looking at reaction time based measures, we found that the implicit-implicit group showed reduced rewiring (compared to the other two groups). By looking at anticipatory errors, we found that this effect was due to the implicit-implicit group still expecting stimuli that were no longer probable, while the other two groups stopped expecting them shortly after the change in sequences. Thus, we were able to find a possible explanation for an effect that we also detected with reaction-time based measures.

In sum, we did not find unequivocal evidence that accuracy-based measures fare better than reaction-time based measures, and we even found that their reliability is substantially smaller. It is possible that accuracy-based measures in ASRT are actually prone to result in floor effects (contrary to the claims of Urry et al., 2015) because of the very high accuracy rates that are expected from participants in this task.

We also assessed the impact of pre-existing biases in serial reaction time measures (requiring serial motor responding). Such biases were acknowledged by Song et al. (2007a) in the context of ASRT, but, to our knowledge, their impact has not been systematically studied. In **Study 4,** we found that even the N-3rd stimulus has an effect on the reaction times measured on the Nth trial and that accounting for this (by using a filter that we named *quadfilter*) reverses the previously puzzling patterns observed for higher-order learning. Apart from affecting the magnitude of learning scores, higher individual variability could be detected by using this filter. This result indicate that the typically low variability on this particular task could be attributed to people's susceptibility to pre-existing biases rather than to their similarity in their statistical learning capacity. We would like to underscore the need to address this question in every task that requires serial motor responding since the artifact that is introduced by these biases seems to be substantial.

Conclusion

The research of implicit statistical learning (or implicit learning, in general) lead to very diverse – and sometimes contradictory – results. In order to understand the source of this variability, our duty is twofold: first, we need to improve the tasks that we use so that the results of measurements aid our theoretical understanding of implicit processes better; and second, based on our ever-expanding theoretical knowledge, we need to refine the tasks even more to narrow and specify their scope. Only this way could we get to the point where we know exactly what we intend to measure, and also have means to do it.

References

Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711). https://doi.org/10.1098/rstb.2016.0058

Arciuli, J., & Conway, C. M. (2018). The Promise—And Challenge—Of Statistical Learning for Elucidating Atypical Language Development. *Current Directions in Psychological Science*, *27*(6), 492–500. https://doi.org/10.1177/0963721418779977

Barnes, K. A., Howard, J. H., 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.

Barnes, K. A., Howard, J. H., 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., & Howard, D. V. (2007). Age-Related Differences in Implicit Learning of Subtle Third-Order Sequential Structure. *The Journals of Gerontology: Series B*, *62*(2), P98–P103. https://doi.org/10.1093/geronb/62.2.P98

Bergstrom, J. C. R., Howard, J. H., & Howard, D. V. (2012). Enhanced Implicit Sequence Learning in College-age Video Game Players and Musicians. *Applied Cognitive Psychology*, *26*(1), 91–96. https://doi.org/10.1002/acp.1800

Boyd, L. A., & Winstein, C. J. (2003). Impact of Explicit Information on Implicit Motor-Sequence Learning Following Middle Cerebral Artery Stroke. *Physical Therapy*, *83*(11), 976–989. https://doi.org/10.1093/ptj/83.11.976

Buchner, A., & Wippich, W. (2000). On the Reliability of Implicit and Explicit Memory Measures. *Cognitive Psychology*, *40*(3), 227–259. https://doi.org/10.1006/cogp.1999.0731

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(1), 24–39. https://doi.org/10.1037/0278-7393.31.1.24

Conway, C. M., & Christiansen, M. H. (2009). Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning. *European Journal of Cognitive Psychology*, *21*(4), 561–580. https://doi.org/10.1080/09541440802097951

Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(1), 189–202. https://doi.org/10.1037/0278-7393.19.1.189

de Boer, B., & Kuhl, P. K. (2003). Investigating the role of infant-directed speech with a computer model. *Acoustics Research Letters Online*, *4*(4), 129–134. https://doi.org/10.1121/1.1613311

Destrebecqz, A., Peigneux, P., Laureys, S., Degueldre, C., Del Fiore, G., Aerts, J., … Maquet, P. (2005). The neural correlates of implicit and explicit sequence learning: Interacting networks revealed by the process dissociation procedure. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *12*(5), 480–490. https://doi.org/10.1101/lm.95605

Dew, I. T. Z., & Cabeza, R. (2011). The porous boundaries between explicit and implicit memory: Behavioral and neural evidence. *Annals of the New York Academy of Sciences*, *1224*, 174–190. https://doi.org/10.1111/j.1749-6632.2010.05946.x

Dorfberger, S., Adi-Japha, E., & Karni, A. (2007). Reduced susceptibility to interference in the consolidation of motor memory before adolescence. *PloS One*, *2*(2), e240. https://doi.org/10.1371/journal.pone.0000240

Du, W., & Kelly, S. W. (2013). Implicit sequence learning in dyslexia: A within-sequence comparison of first- and higher-order information. *Annals of Dyslexia*, *63*(2), 154–170. https://doi.org/10.1007/s11881-012-0077-1

Du, Y., & Clark, J. E. (2017). New insights into statistical learning and chunk learning in implicit sequence acquisition. *Psychonomic Bulletin & Review*, *24*(4), 1225–1233. https://doi.org/10.3758/s13423-016-1193-4

E. Hick, W. (1952). On the rate of information gain. *Quarterly Journal of Experimental Psychology*, *4*, 11–26. https://doi.org/10.1080/17470215208416600

Edwards, J. R. (2001). Ten Difference Score Myths. *Organizational Research Methods*, *4*(3), 265–287. https://doi.org/10.1177/109442810143005

Emberson, L. L., Conway, C. M., & Christiansen, M. H. (2011). Timing is everything: Changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Quarterly Journal of Experimental Psychology (2006)*, *64*(5), 1021–1040. https://doi.org/10.1080/17470218.2010.538972

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.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality vs. modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125. https://doi.org/10.1016/j.tics.2014.12.010

Gao, J., Wong-Lin, K., Holmes, P., Simen, P., & Cohen, J. D. (2009). Sequential effects in two-choice reaction time tasks: Decomposition and synthesis of mechanisms.

Neural Computation, *21*(9), 2407–2436. https://doi.org/10.1162/neco.2009.09-08-866

Goedert, K. M., & Willingham, D. B. (2002). Patterns of interference in sequence learning and prism adaptation inconsistent with the consolidation hypothesis. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *9*(5), 279–292. https://doi.org/10.1101/lm.50102

Goodwin, L. D., & Leech, N. L. (2006). Understanding Correlation: Factors That Affect the Size of r. *The Journal of Experimental Education*, *74*(3), 249–266. https://doi.org/10.3200/JEXE.74.3.249-266

Grosjean, M., Rosenbaum, D. A., & Elsinger, C. (2001). Timing and reaction time. *Journal of Experimental Psychology. General*, *130*(2), 256–272.

Gucciardi, D. F., & Dimmock, J. A. (2008). Choking under pressure in sensorimotor skills: Conscious processing or depleted attentional resources? *Psychology of Sport and Exercise*, *9*(1), 45–59. https://doi.org/10.1016/j.psychsport.2006.10.007

Hallgató, E., Győri-Dani, D., Pekár, J., Janacsek, K., & Nemeth, D. (2013). The differential consolidation of perceptual and motor learning in skill acquisition. *Cortex*, *49*(4), 1073–1081. https://doi.org/10.1016/j.cortex.2012.01.002

Handa, A., Rhee, J., & Wright, D. L. (2016). The Structural Relationship Between Two Motor Sequences Practiced Close in Time Impacts Offline Facilitation. *Journal of Motor Behavior*, *48*(1), 47–56. https://doi.org/10.1080/00222895.2015.1035431

Hedenius, M., Persson, J., Tremblay, A., Adi-Japha, E., Veríssimo, J., Dye, C. D., … Ullman, M. T. (2011). Grammar predicts procedural learning and consolidation deficits in children with Specific Language Impairment. *Research in Developmental Disabilities*, *32*(6), 2362–2375. https://doi.org/10.1016/j.ridd.2011.07.026

Hedge, C., Powell, G., Bompas, A., Vivian-Griffiths, S., & Sumner, P. (2018). Low and variable correlation between reaction time costs and accuracy costs explained by accumulation models: Meta-analysis and simulations. *Psychological Bulletin*, *144*(11), 1200–1227. https://doi.org/10.1037/bul0000164

Heerey, E. A., & Velani, H. (2010). Implicit learning of social predictions. *Journal of Experimental Social Psychology*, *46*(3), 577–581. https://doi.org/10.1016/j.jesp.2010.01.003

Horvath, K., Torok, C., Pesthy, O., Nemeth, D., & Janacsek, K. (2018). Explicit instruction differentially affects subcomponents of procedural learning and consolidation. *BioRxiv*, 433243.

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. https://doi.org/10.1037//0882-7974.7.2.232

Howard, D. V., & Howard, J. H. (2001). When it does hurt to try: Adult age differences in the effects of instructions on implicit pattern learning. *Psychonomic Bulletin & Review*, *8*(4), 798–805. https://doi.org/10.3758/BF03196220

Howard, D. V., Howard, J. H., 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. https://doi.org/10.1037/0882-7974.19.1.79

Howard, J. H., & Howard, D. V. (1997). Age differences in implicit learning of higher order dependencies in serial patterns. *Psychology and Aging*, *12*(4), 634–656.

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. https://doi.org/10.1037/a0012797

Howard, J. H., Howard, D. V., Dennis, N. A., & Yankovich, H. (2007). Event Timing and age deficits in higher-order sequence learning. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *14*(6), 647–668. https://doi.org/10.1080/13825580601186635

Howard, J. H., Howard, D. V., Japikse, K. C., & Eden, G. F. (2006). Dyslexics are impaired on implicit higher-order sequence learning, but not on implicit spatial context learning. *Neuropsychologia*, *44*(7), 1131–1144. https://doi.org/10.1016/j.neuropsychologia.2005.10.015

Janacsek, K., Ambrus, G. G., Paulus, W., Antal, A., & Nemeth, D. (2015). Right hemisphere advantage in statistical learning: Evidence from a probabilistic sequence learning task. *Brain Stimulation*, *8*(2), 277–282. https://doi.org/10.1016/j.brs.2014.11.008

Janacsek, K., Borbély-Ipkovich, E., Nemeth, D., & Gonda, X. (2018). How can the depressed mind extract and remember predictive relationships of the environment? Evidence from implicit probabilistic sequence learning. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *81*, 17–24. https://doi.org/10.1016/j.pnpbp.2017.09.021

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Developmental Science*, *15*(4), 496–505. https://doi.org/10.1111/j.1467-7687.2012.01150.x

Japikse, K. C., Negash, S., Howard, J. H., & Howard, D. V. (2003). Intermanual transfer of procedural learning after extended practice of probabilistic sequences. *Experimental Brain Research*, *148*(1), 38–49. https://doi.org/10.1007/s00221-002-1264-9

Kalra, P. B., Gabrieli, J. D. E., & Finn, A. S. (2019). Evidence of stable individual differences in implicit learning. *Cognition*, *190*, 199–211. https://doi.org/10.1016/j.cognition.2019.05.007

Karzon, R. G. (1985). Discrimination of polysyllabic sequences by one- to four-month-old infants. *Journal of Experimental Child Psychology*, *39*(2), 326–342.

Kaufman, S. B., Deyoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*(3), 321–340. https://doi.org/10.1016/j.cognition.2010.05.011

Kirby, N. H. (1976). Sequential effects in two-choice reaction time: Automatic facilitation or subjective expectancy? *Journal of Experimental Psychology: Human Perception and Performance*, *2*(4), 567.

Kóbor, A., Takács, Á., Kardos, Z., Janacsek, K., Horváth, K., Csépe, V., & Nemeth, D. (2018). ERPs differentiate the sensitivity to statistical probabilities and the learning of sequential structures during procedural learning. *Biological Psychology*, *135*, 180–193. https://doi.org/10.1016/j.biopsycho.2018.04.001

Kornblum, S. (1973). Sequential effects in choice reaction time: A tutorial review. *Attention and Performance IV*, 259–288.

Lagarde, J., Li, L., Thon, B., Magill, R., & Erbani, E. (2002). Interactions between human explicit and implicit perceptual motor learning shown by kinematic variables. *Neuroscience Letters*, *327*(1), 66–70. https://doi.org/10.1016/s0304-3940(02)00380-4

Lebel, E. P., & Paunonen, S. V. (2011). Sexy but often unreliable: The impact of unreliability on the replicability of experimental findings with implicit measures. *Personality & Social Psychology Bulletin*, *37*(4), 570–583. https://doi.org/10.1177/0146167211400619

Lee, J. C., Beesley, T., & Livesey, E. J. (2016). Sequential effects and sequence learning in a three-choice serial reaction time task. *Acta Psychologica*, *170*, 168–176. https://doi.org/10.1016/j.actpsy.2016.08.004

Li, X., Zhao, X., Shi, W., Lu, Y., & Conway, C. M. (2018). Lack of Cross-Modal Effects in Dual-Modality Implicit Statistical Learning. *Frontiers in Psychology*, *9*, 146. https://doi.org/10.3389/fpsyg.2018.00146

Marvel, C. L., Schwartz, B. L., Howard, D. V., & Howard, J. H. (2005). Implicit learning of non-spatial sequences in schizophrenia. *Journal of the International Neuropsychological Society*, *11*(6), 659–667. https://doi.org/10.1017/S1355617705050861

Masters, R. S. W. (1992). Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology*, *83*(3), 343–358. https://doi.org/10.1111/j.2044-8295.1992.tb02446.x

Maxwell, J. P., Masters, R. S. W., & Eves, F. F. (2000). From novice to no know-how: A longitudinal study of implicit motor learning. *Journal of Sports Sciences*, *18*(2), 111–120. https://doi.org/10.1080/026404100365180

Negash, S., Boeve, B. F., Geda, Y. E., Smith, G. E., Knopman, D. S., Ivnik, R. J., … Petersen, R. C. (2007). Implicit Learning of Sequential Regularities and Spatial Contexts in Corticobasal Syndrome. *Neurocase*, *13*(3), 133–143. https://doi.org/10.1080/13554790701401852

Negash, S., Howard, D. V., Japikse, K. C., & Howard, J. H. (2003). Age-Related Differences in Implicit Learning of Non-Spatial Sequential Patterns. *Aging, Neuropsychology, and Cognition*, *10*(2), 108–121. https://doi.org/10.1076/anec.10.2.108.14462

Nemeth, D., Hallgató, E., Janacsek, K., Sándor, T., & Londe, Z. (2009). Perceptual and motor factors of implicit skill learning. *NeuroReport*, *20*(18), 1654. https://doi.org/10.1097/WNR.0b013e328333ba08

Nemeth, D., & Janacsek, K. (2011). The dynamics of implicit skill consolidation in young and elderly adults. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *66*(1), 15–22. https://doi.org/10.1093/geronb/gbq063

Nemeth, D., Janacsek, K., Balogh, V., Londe, Z., Mingesz, R., Fazekas, M., … Vetro, A. (2010). Learning in autism: Implicitly superb. *PloS One*, *5*(7), e11731. https://doi.org/10.1371/journal.pone.0011731

Nemeth, D., Janacsek, K., Csifcsak, G., Szvoboda, G., Howard, J. H., & Howard, D. V. (2011). Interference between Sentence Processing and Probabilistic Implicit Sequence Learning. *PLOS ONE*, *6*(3), e17577. https://doi.org/10.1371/journal.pone.0017577

Nemeth, D., Janacsek, K., & Fiser, J. (2013). Age-dependent and coordinated shift in performance between implicit and explicit skill learning. *Frontiers in Computational Neuroscience*, *7*. https://doi.org/10.3389/fncom.2013.00147

Nemeth, D., Janacsek, K., Király, K., Londe, Z., Németh, K., Fazekas, K., … Csányi, A. (2013). Probabilistic sequence learning in mild cognitive impairment. *Frontiers in Human Neuroscience*, *7*, 318. https://doi.org/10.3389/fnhum.2013.00318

Nemeth, D., Janacsek, K., Londe, Z., Ullman, M. T., Howard, D. V., & Howard, J. H. (2010). Sleep has no critical role in implicit motor sequence learning in young and old adults. *Experimental Brain Research*, *201*(2), 351–358. https://doi.org/10.1007/s00221-009-2024-x

Nemeth, D., Janacsek, K., Polner, B., & Kovacs, Z. A. (2013). Boosting Human Learning by Hypnosis. *Cerebral Cortex*, *23*(4), 801–805. https://doi.org/10.1093/cercor/bhs068

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

Norman, E., & Price, M. C. (2012). Social intuition as a form of implicit learning: Sequences of body movements are learned less explicitly than letter sequences.

Advances in Cognitive Psychology, *8*(2), 121–131. https://doi.org/10.2478/v10053-008-0109-x

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Reber, A. S., & Allen, R. (2000). Individual Differences in Implicit Learning Implications for the Evolution of Consciousness. In R. G. Kunzendorf & B. Wallace, *Individual differences in subjective experience: First-person constraints on theories of consciousness, subconsciousness, and self-consciousness*. Amsterdam: John Benjamins.

Remillard, G. (2008). Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies. *The Quarterly Journal of Experimental Psychology*, *61*(3), 400–424. https://doi.org/10.1080/17470210701210999

Remillard, G. (2010). Implicit learning of fifth- and sixth-order sequential probabilities. *Memory & Cognition*, *38*(7), 905–915. https://doi.org/10.3758/MC.38.7.905

Remillard, G. (2011). Pure perceptual-based learning of second-, third-, and fourth-order sequential probabilities. *Psychological Research*, *75*(4), 307–323. https://doi.org/10.1007/s00426-010-0309-0

Remington, R. J. (1969). Analysis of sequential effects on choice reaction times. *Journal of Experimental Psychology*, *82*(2), 250–257. https://doi.org/10.1037/h0028122

Sanchez, D. J., & Reber, P. J. (2013). Explicit pre-training instruction does not improve implicit perceptual-motor sequence learning. *Cognition*, *126*(3), 341–351. https://doi.org/10.1016/j.cognition.2012.11.006

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 501–518. https://doi.org/10.1037/0278-7393.13.3.501

Schwartz, B. L., Howard, D. V., Howard, J. H., Hovaguimian, A., & Deutsch, S. I. (2003). Implicit learning of visuospatial sequences in schizophrenia. *Neuropsychology*, *17*(3), 517–533.

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. https://doi.org/10.1016/j.jml.2015.02.001

Simor, P., Zavecz, Z., Horváth, K., Éltető, N., Török, C., Pesthy, O., … Nemeth, D. (2019). Deconstructing Procedural Memory: Different Learning Trajectories and Consolidation of Sequence and Statistical Learning. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02708

Smith, J. G., & McDowall, J. (2004). Impaired higher order implicit sequence learning on the verbal version of the serial reaction time task in patients with Parkinson’s disease. *Neuropsychology*, *18*(4), 679–691. https://doi.org/10.1037/0894-4105.18.4.679

Soetens, E., Boer, L. C., & Hueting, J. E. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 598.

Song, S., Howard, J. H., & Howard, D. V. (2007a). Implicit probabilistic sequence learning is independent of explicit awareness. *Learning & Memory*, *14*(3), 167–176. https://doi.org/10.1101/lm.437407

Song, S., Howard, J. H., & Howard, D. V. (2007b). Sleep does not benefit probabilistic motor sequence learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(46), 12475–12483. https://doi.org/10.1523/JNEUROSCI.2062-07.2007

Song, S., Howard, J. H., & Howard, D. V. (2008). Perceptual sequence learning in a serial reaction time task. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *189*(2), 145–158. https://doi.org/10.1007/s00221-008-1411-z

Song, S., Marks, B., Howard, J. H., & Howard, D. V. (2009). Evidence for parallel explicit and implicit sequence learning systems in older adults. *Behavioural Brain Research*, *196*(2), 328–332. https://doi.org/10.1016/j.bbr.2008.09.022

Stark-Inbar, A., Raza, M., Taylor, J. A., & Ivry, R. B. (2016). Individual differences in implicit motor learning: Task specificity in sensorimotor adaptation and sequence learning. *Journal of Neurophysiology*, *117*(1), 412–428.

Sun, R., Zhang, X., Slusarz, P., & Mathews, R. (2007). The interaction of implicit learning, explicit hypothesis testing learning and implicit-to-explicit knowledge extraction. *Neural Networks: The Official Journal of the International Neural Network Society*, *20*(1), 34–47. https://doi.org/10.1016/j.neunet.2006.07.002

Takács, Á., Kóbor, A., Chezan, J., Éltető, N., Tárnok, Z., Nemeth, D., … Janacsek, K. (2018). Is procedural memory enhanced in Tourette syndrome? Evidence from a sequence learning task. *Cortex*, *100*, 84–94. https://doi.org/10.1016/j.cortex.2017.08.037

Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1711). https://doi.org/10.1098/rstb.2016.0056

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin*, *139*(4), 792–814. https://doi.org/10.1037/a0030801

Török, B., Janacsek, K., Nagy, D. G., Orbán, G., & Nemeth, D. (2017). Measuring and filtering reactive inhibition is essential for assessing serial decision making and learning. *Journal of Experimental Psychology. General*, *146*(4), 529–542. https://doi.org/10.1037/xge0000288

Urry, K., Burns, N. R., & Baetu, I. (2015). Accuracy-based measures provide a better measure of sequence learning than reaction time-based measures. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01158

Urry, K., Burns, N. R., & Baetu, I. (2018). Age-related differences in sequence learning: Findings from two visuo-motor sequence learning tasks. *British Journal of Psychology*, *109*(4), 830–849. https://doi.org/10.1111/bjop.12299

Verneau, M., van der Kamp, J., Savelsbergh, G. J. P., & de Looze, M. P. (2015). Proactive and retroactive transfer of middle age adults in a sequential motor learning task. *Acta Psychologica*, *156*, 57–63. https://doi.org/10.1016/j.actpsy.2015.01.009

Vervaeck, K. R., & Boer, L. C. (1980). Sequential effects in two-choice reaction time: Subjective expectancy and automatic after-effect at short response-stimulus intervals. *Acta Psychologica*, *44*(2), 175–190.

Walk, A. M., & Conway, C. M. (2016). Cross-Domain Statistical–Sequential Dependencies Are Difficult to Learn. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00250

Walker, C. M., Walker, M. F., & Sunderland, A. (2003). Dressing after a Stroke: A Survey of Current Occupational Therapy Practice. *British Journal of Occupational Therapy*, *66*(6), 263–268. https://doi.org/10.1177/030802260306600605

Walker, M. P., Brakefield, T., Hobson, J. A., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature*, *425*(6958), 616–620. https://doi.org/10.1038/nature01930

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. https://doi.org/10.3758/BF03198552

Woods, D. L., Wyma, J. M., Yund, E. W., Herron, T. J., & Reed, B. (2015). Factors influencing the latency of simple reaction time. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00131

Yotsumoto, Y., Watanabe, T., Chang, L.-H., & Sasaki, Y. (2013). Consolidated learning can be susceptible to gradually-developing interference in prolonged motor learning. *Frontiers in Computational Neuroscience*, *7*, 69. https://doi.org/10.3389/fncom.2013.00069

List of publication that the dissertation is based upon

Nemeth, D., **Hallgató, E.**, Janacsek, K., Sándor, T., & Londe, Z. (2009). Perceptual and motor factors of implicit skill learning. NeuroReport, 20(18), 1654. https://doi.org/10.1097/WNR.0b013e328333ba08. Impact factor: 1.805.

Hallgató, E., Győri-Dani, D., Pekár, J., Janacsek, K., & Nemeth, D. (2013). The differential consolidation of perceptual and motor learning in skill acquisition. *Cortex*, *49*(4), 1073–1081. https://doi.org/10.1016/j.cortex.2012.01.002. Impact factor: 6.042.

Szegedi-Hallgató, E., Janacsek, K., Vékony, T., Tasi, L. A., Kerepes, L., Hompoth, E. A., … Németh, D. (2017). Explicit instructions and consolidation promote rewiring of automatic behaviors in the human mind. Scientific Reports, 7(1), 1–7. https://doi.org/10.1038/s41598-017-04500-3. Impact factor: 4.122.

Szegedi-Hallgató, E., Janacsek, K., & Nemeth, D. (2019). Different levels of statistical learning—Hidden potentials of sequence learning tasks. *PloS One*, *14*(9), e0221966. https://doi.org/10.1371/journal.pone.0221966. Impact factor: 2.776 (*in 2018*)

Cumulative impact factor of the studies: **14.745**

Note. Each co-author has granted permission for the given publication to be included in the current dissertation.

List of publications directly not used in the dissertation

Bálint, A., Hompoth, E. A., Kerepes, L., Tasi, L. A., Vékony, T., & **Hallgató, E.** (2013). Áthuzalozás—A szokások megváltoztatásának modellezése szekvenciatanulással. In Szegedi Pszichológiai Tanulmányok 2013 (pp. 7–18). Szeged: SZEK JGYF Kiadó.

Belucz, J., Bertók, I. B., Csomós, D., Farkas, D., Léh, N., & **Hallgató, E.** (2014). A kontextusfüggő implicit tanulás modalitásainak vizsgálata: Perceptuális és motoros faktorok. Impulzus - Szegedi Pszichológiai Tanulmányok, 1(1), 108–123.

Horváth, M. D., **Szegedi-Hallgató, E.**, & Csábi, E. (2018). Vizuális ingerek figyelemelterelő hatásának vizsgálata vezetést szimuláló feladat során. Magyar Pszichológiai Szemle, 73(3), 357–372. https://doi.org/10.1556/0016.2018.006

Huszár, T., Makra, E., & **Hallgató, E.** (2010). A tabuszavak hatása lexikális döntési helyzetben. Psychiatria Hungarica, 25(6), 525–537.

Krajcsi, A., & **Hallgató, E.** (2012). Fejlődési diszkalkulia diagnózisa felnőtteknél—Az Aritmetikai Képességek Kognitív Fejlődése teszt. Gyógypedagógiai Szemle, 4(4–5), 330–342.

Lukács, G., Gula, B., **Szegedi-Hallgató, E.,** & Csifcsák, G. (2017). Association-based Concealed Information Test: A novel reaction time-based deception detection method. Journal of Applied Research in Memory and Cognition, 6(3), 283–294.

Lukács, G., Huszár, K., & **Hallgató, E.** (2014). Implicit szekvencia tanulás: Inter-stimulus intervallum és szubjektív élmény. Impulzus - Szegedi Pszichológiai Tanulmányok, 1(1), 124–141.

Szabó Hangya, L., Szentes, B., & **Hallgató, E.** (2013). Placebohatás implicit tanulás esetén. In Szegedi Pszichológiai Tanulmányok 2013 (pp. 151–161). Szeged: SZEK JGYF Kiadó.

Szalóki, S., & **Hallgató, E.** (2015). A fotóolvasás mint előfeszítési jelenség vizsgálata. Impulzus - Szegedi Pszichológiai Tanulmányok, 2(1).